Murat Yilmaz
Jörg Niemann
Paul Clarke
Richard Messnarz (Eds.)

Communications in Computer and Information Science 1251

Systems, Software and Services Process Improvement

27th European Conference, EuroSPI 2020
Düsseldorf, Germany, September 9–11, 2020
Proceedings

Springer

Communications in Computer and Information Science 1251

More information about this series at http://www.springer.com/series/7899

Murat Yilmaz · Jörg Niemann ·
Paul Clarke · Richard Messnarz (Eds.)

Systems, Software and Services Process Improvement

27th European Conference, EuroSPI 2020 Düsseldorf, Germany, September 9–11, 2020 Proceedings

Springer

Editors
Murat Yilmaz
School of Computing
Dublin City University
Dublin, Ireland

Paul Clarke
School of Computing
Dublin City University
Dublin, Ireland

Jörg Niemann
University of Applied Sciences Düsseldorf
Düsseldorf, Germany

Richard Messnarz
I.S.C.N. GesmbH
Graz, Steiermark, Austria

ISSN 1865-0929 ISSN 1865-0937 (electronic)
Communications in Computer and Information Science
ISBN 978-3-030-56440-7 ISBN 978-3-030-56441-4 (eBook)
https://doi.org/10.1007/978-3-030-56441-4

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This textbook comprises the proceedings of the 27th International Conference on Systems, Software and Services Process Improvement (EuroSPI 2020), held during September 9–11, 2020, in Düsseldorf, Germany. The conference was partially held virtually due to the COVID-19 pandemic.

Conferences were held in Dublin (Ireland) in 1994, in Vienna (Austria) in 1995, in Budapest (Hungary) in 1997, in Gothenburg (Sweden) in 1998, in Pori (Finland) in 1999, in Copenhagen (Denmark) in 2000, in Limerick (Ireland) in 2001, in Nuremberg (Germany) in 2002, in Graz (Austria) in 2003, in Trondheim (Norway) in 2004, in Budapest (Hungary) in 2005, in Joensuu (Finland) in 2006, in Potsdam (Germany) in 2007, in Dublin (Ireland) in 2008, in Alcala (Spain) in 2009, in Grenoble (France) in 2010, in Roskilde (Denmark) in 2011, in Vienna (Austria) in 2012, Dundalk (Ireland) in 2013, in Luxembourg in 2014, in Ankara (Turkey) 2015, in Graz (Austria) 2016, in Ostrava (Czech Republic) 2017, in Bilboa (Spain) in 2018, in Edinburgh (UK) in 2019, and in Düsseldorf (Germany) in 2020.

EuroSPI is an initiative with the following major action lines http://www.eurospi.net:

- Establishing an annual EuroSPI conference supported by software process improvement networks from different EU countries.
- Establishing a social media strategy with groups from LinkedIn, Facebook, and Twitter, as well as online statements, speeches, and keynotes on YouTube, and a set of proceedings and recommended books.
- Establishing an effective team of national representatives (from each EU country) growing step by step into more European countries.
- Establishing a European Qualification Framework for a pool of professions related to SPI and management. This is supported by European certificates and examination systems.

EuroSPI has a cooperation with the EU Blueprint for Automotive project DRIVES (2018–2021) where leaders in automotive industry discuss and present skills for the Europe 2030 strategy in the automotive sector.

EuroSPI also has a cooperation with the EU Blueprint for Battery Systems ALBATTS (2020–2023) where leaders in the industry discuss and present skills for the creation of a battery production in Europe for cars, ships, planes, industry plants, etc.

EuroSPI established the SPI Manifesto (SPI = Systems, Software and Services Process Improvement), a set of social media groups including a selection of presentations and keynotes freely available on YouTube, and access to role-based job qualifications through the European Certification and Qualification Association (www.ecqa.org).

From 2013 onwards new communities (cybersecurity, Internet of Things, Agile) joined EuroSPI, the letter S encompassing System, Software, Service, Safety, and

Security and the letter I encompassing Improvement, Innovation, and Infrastructure (Internet of Things).

In 2019, our dear friend and long-term EuroSPI Conference Series Editor, Prof. Rory O'Connor of Dublin City University and Lero – the Science Foundation Ireland Research Centre for Software, passed away. His knowledge, wit, and energy are missed greatly and the committee has, in collaboration with ISCN, ASQ, and Lero, established the Rory O'Connor Award for Research Excellence. On an annual basis, the individual presenting the highest quality work to the conference audience, especially in areas of major importance to our field, will be awarded this honor. The award was inaugurated at EuroSPI 2019 and we look forward to recognizing our finest research work and our excellent researchers with this award.

A typical characterization of EuroSPI is reflected in the following statement: "… the biggest value of EuroSPI lies in its function as a European knowledge and experience exchange mechanism for SPI and innovation."

Since its beginning in 1994 in Dublin, the EuroSPI initiative has outlined that there is not a single silver bullet with which to solve SPI issues, but that you need to understand a combination of different SPI methods and approaches to achieve concrete benefits. Therefore, each proceedings volume covers a variety of different topics, and at the conference we discuss potential synergies and the combined use of such methods and approaches. These proceedings contain a visionary paper by a h-index > 40 researcher, and 17 selected research papers under the following headings:

- Section I: SPI Manifesto and Improvement Strategies
- Section II: SPI and Emerging Software and Systems Engineering Paradigms
- Section III: SPI and Standards and Safety and Security Norms
- Section IV: SPI and Team Performance & Agile & Innovation
- Section V: SPI and Agile
- Section VI: Selected Thematic Workshop Papers

The visionary paper talks about fertilization in software engineering and how future methods and developments will impact software engineering in the next few years. Section I presents two papers addressing manifesto and improvement strategies in SPI. Section II presents three papers related to SPI and emerging software and system engineering paradigms. Section III presents four papers dealing with SPI's standards, safety, and security norms. Section IV presents two papers exploring the issues of SPI and team performance in agile software development and innovation. Section V presents five papers exploring SPI and agile.

Section VI presents selected keynotes from EuroSPI workshops concerning the future of SPI. From 2010 onwards EuroSPI invites recognized key researchers to publish new future directions of SPI. These key messages are discussed in interactive workshops and help create SPI communities based on new topics. The first set of papers relates to the Emerging SW Engineering Paradigms workshop and explores upcoming new developments like AI and big data. The second collection of papers relate to the topic of "Digitalization of Industry, Infrastructure and E-Mobility." The third collection of papers surrounds the topic of "Good and Bad Practices in Improvement for SPI." The fourth collection relates to the topic of "Functional Safety and Cybersecurity" and addresses best practices from the automotive industry to cope with cybersecurity and functional safety. The fifth collection addresses experiences with “Agile and Lean” and examines a series of success factors and examples of being lean and agile. The sixth collection of workshop papers addresses the topic of “Standards and Assessment Models” and examines different ISO standards – assessment models are introduced, explained, and discussed. The seventh collection of papers addresses “Team Skills and Diversity Strategies” and examines a variety of organizational and human factors related to SPI. The eight set of papers address “Recent Innovations” or new service innovation models. The ninth collection of papers discusses the use of “Virtual Reality” in different application environments.

September 2020

Murat Yilmaz
Jörg Niemann
Paul Clarke
Richard Messnarz

Recommended Further Reading

In [1] the proceedings of three EuroSPI conferences were integrated into one book, which was edited by 30 experts in Europe. The proceedings of EuroSPI 2005 to 2019 have been published by Springer in [2–16], respectively.

References

1. Messnarz, R., Tully, C. (eds.): Better Software Practice for Business Benefit – Principles and Experience, 409 pages. IEEE Computer Society Press, Los Alamitos (1999).
2. Richardson, I., Abrahamsson, P., Messnarz, R. (eds.): Software Process Improvement. LNCS, vol. 3792, p. 213. Springer, Heidelberg (2005).
3. Richardson, I., Runeson, P., Messnarz, R. (eds.): Software Process Improvement. LNCS, vol. 4257, pp. 11–13. Springer, Heidelberg (2006).
4. Abrahamsson, P., Baddoo, N., Margaria, T., Messnarz, R. (eds.): Software Process Improvement. LNCS, vol. 4764, pp. 1–6. Springer, Heidelberg (2007).
5. O'Connor, R.V., Baddoo, N., Smolander, K., Messnarz, R. (eds): Software Process Improvement. CCIS, vol. 16, Springer, Heidelberg (2008).
6. O'Connor, R.V., Baddoo, N., Gallego C., Rejas Muslera R., Smolander, K., Messnarz, R. (eds): Software Process Improvement. CCIS, vol. 42, Springer, Heidelberg (2009).
7. Riel A., O'Connor, R.V. Tichkiewitch S., Messnarz, R. (eds): Software, System, and Service Process Improvement. CCIS, vol. 99, Springer, Heidelberg (2010).
8. O'Connor, R., Pries-Heje, J. and Messnarz R., Systems, Software and Services Process Improvement, CCIS Vol. 172, Springer-Verlag, (2011).
9. Winkler, D., O'Connor, R.V. and Messnarz R. (Eds), Systems, Software and Services Process Improvement, CCIS 301, Springer-Verlag, (2012).
10. McCaffery, F., O'Connor, R.V. and Messnarz R. (Eds), Systems, Software and Services Process Improvement, CCIS 364, Springer-Verlag, (2013).
11. Barafort, B., O'Connor, R.V. and Messnarz R. (Eds), Systems, Software and Services Process Improvement, CCIS 425, Springer-Verlag, (2014).
12. O'Connor, R.V. Akkaya, M., Kemaneci K., Yilmaz, M., Poth, A. and Messnarz R. (Eds), Systems, Software and Services Process Improvement, CCIS 543, Springer-Verlag, (2015).
13. Kreiner, C., Poth., A., O'Connor, R.V., and Messnarz R. (Eds), Systems, Software and Services Process Improvement, CCIS 633, Springer-Verlag, (2016).
14. Stolfa, J, Stolfa, S., O'Connor, R.V., and Messnarz R. (Eds), Systems, Software and Services Process Improvement, CCIS 633, Springer-Verlag, (2017).

15. Larrucea, X., Santamaria, I., O'Connor, R.V., Messnarz, R. (Eds), Systems, Software and Services Process Improvement, CCIS Vol. 896, Springer-Verlag, (2018).
16. Walker A., O'Connor, R.V., Messnarz, R. (Eds), Systems, Software and Services Process Improvement, CCIS Vol. 1060, Springer-Verlag, (2019).

Acknowledgements

Some contributions published in this book have been funded with support from the European Commission. European projects (supporting ECQA and EuroSPI) contributed to this Springer book, including DRIVES – BLUEPRINT Project (591988-EPP-1-2017-1-CZ-EPPKA2-SSA-B), OpenInnotrain (H2020-MSCA-RISE-2018, exchange of researchers), ProHeritage (785211 – Pro Heritage – H2020-EE-2016-2017), ALBATTS – BLUEPRINT Project (612675-EPP-1-2019-1-SE-EPPKA2-SSA-B), and ECEPE Erasmus+ Project (2019-1-CZ01-KA203-061430).

In this case the publications reflect the views only of the author(s), and the Commission cannot be held responsible for any use, which may be made of the information contained therein.

This work was supported, in part, by Science Foundation Ireland grant 13/RC/2094 and co-funded under the European Regional Development Fund through the Southern & Eastern Regional Operational Programme by Lero – the Science Foundation Ireland Research Centre for Software (www.lero.ie).

Funded by the
Erasmus+ programme
of the European Union

Organization

General Chair and Workshop Chair

Richard Messnarz	ISCN GesmbH, Austria

General Co-chair

Micheal Mac an Airchinnigh	ISCN, Ireland

Scientific Chairs

Murat Yilmaz	Dublin City University, Ireland
Paul Clarke	Dublin City University, Ireland

Organization Chairs

Richard Messnarz	ISCN GesmbH, Austria
Andreas Riel	Grenoble INP, France
Damjan Ekert	ISCN GesmbH, Austria
Tobias Zehetner	ISCN GesmbH, Austria
Laura Aschbacher	ISCN GesmbH, Austria

Local Organization Chairs

Jörg Niemann	University of Applied Sciences Düsseldorf, Germany
Claudia Fussenecker	University of Applied Sciences Düsseldorf, Germany

W1 Emerging Software Engineering Paradigms Co-chairs

Murat Yilmaz	Dublin City University, Ireland
Paul Carke	Dublin City University, Ireland
Ricardo Colomo Palacios	Ostfold University College, Norway

WS2 Digitalization of Industry, Infrastructure and E-Mobility Workshop Co-chairs

Peter Dolejsi	The European Automobile Manufacturers Association (ACEA), Belgium
Jakub Stolfa	VSB Ostrava, Czech Republic
Svatopluk Stolfa	VSB Ostrava, Czech Republic

Andreas Riel	Grenoble INP, France
Michael Reiner	University of Applied Sciences Krems, Austria
Georg Macher	TU Graz, Austria
Richard Messnarz	ISCN GesmbH, Austria

WS3 Good and Bad Practices in Improvement Workshop Co-chairs

Eva Breske	Bosch Engineering, Germany
Tomas Schweigert	ExpleoGroup, Germany
So Norimatsu	JASPIC, Japan

WS4 Functional Safety and Cybersecurity Workshop Co-chairs

Alexander Much	Elektrobit, Germany
Miklos Biro	SCCH, Austria
Richard Messnarz	ISCN GesmbH, Austria

WS5 Experiences with Agile and Lean Workshop Co-chairs

Alexander Poth	Volkswagen AG, Germany
Susumu Sasabe	JUSE, Japan
Antonia Mas	University of the Balearic Islands, Spain

WS6 Standards and Assessment Models Workshop Co-chairs

Gerhard Griessnig	AVL, Austria
Klaudia Dussa Zieger	IMBUS, Germany
Timo Varkoi	FiSMA, Finland

WS7 Team Skills and Diversity Strategies Workshop Co-chairs

Gabriele Sauberer	Termnet, Austria
Mirna Munoz	CIMAT, Mexico

WS 8: Recent Innovations Workshop Co-chairs

Thomas Fehlmann	Euro Project Office AG, Switzerland
Andreas Riel	Grenoble INP, France

WS 9: Virtual Reality Co-chairs

Michael Reiner	University of Applied Sciences Krems, Austria
Jörg Niemann	University of Applied Sciences Düsseldorf, Germany
Christian Reimann	University of Applied Sciences Dortmund, Germany
Philip Wogart	VR/AR Association & Miami Ad School Europe Hamburg, Germany

Board Members

EuroSPI Board Members represent centers or networks of SPI excellence having extensive experience with SPI. The board members collaborate with different European SPINS (Software Process Improvement Networks). The following have been members of the conference board for a significant period:

- Richard Messnarz, ISCN, Austria
- Micheal Mac an Airchinnigh, ISCN, Ireland
- Paul Clarke, Dublin City University, Ireland
- Gabriele Sauberer, TermNet, Austria
- Jörg Niemann, University of Applied Sciences Düsseldorf, Germany
- Andreas Riel, Grenoble Institute of Technology, France
- Miklós Biró, Software Competence Center Hagenberg GmbH, Johannes Kepler Universität Linz, Austria
- Ricardo Colomo-Palacios, Ostfold University, Norway
- Georg Macher, Graz University of Technology, Austria
- Michael Reiner, IMC FH Krems, University of Applied Sciences, Austria
- Murat Yilmaz, Dublin City University, Ireland
- Jakub Stolfa, VSB Ostrava, Czech Republic

EuroSPI Scientific and Industry Program Committee

EuroSPI established an International Committee of selected well-known experts in SPI who are willing to be mentioned in the program and to review a set of papers each year. The list below represents the Research and Industry Program Committee members. EuroSPI also has a separate Industrial Program Committee responsible for the industry/experience contributions.

EuroSPI 2020 Scientific Program Committee

Biro Miklos	John von Neumann Computer Society, Hungary
Calvo-Manzano Villalon Jose A.	Politechnic University of Madrid (UPM), Spain
Clarke Paul	Dublin City University, Ireland
Colomo Palacios Ricardo	Ostfold University College, Norway
Dobaj Jürgen	TU Graz, Austria
Fussenecker Claudia	University of Applied Sciences Düsseldorf, Germany
Georgiadou Elli	Independent, UK
Gokalp Ebru	Baskent University, Turkey
Gomez Alvarez Maria Clara	Universidad de Medellin, Colombia
Gulec Ulas	Hasan Kalyoncu University, Turkey
Hirz Mario	TU Graz, Austria
Krisper Michael	TU Graz, Austria
Landes Dieter	University of Coburg, Germany

Mac An Airchinnigh Micheal	ISCN, Ireland
Macher Georg	TU Graz, Austria
Macmahon Silvana Togneri	Dublin City University, Ireland
Makkar Samer	VALEO, Egypt
Martin Helmut	Virtual Vehicle, Austria
Martins Paula	University of the Algarve, Portugal
Matthies Christoph	Hasso Plattner Institute, Germany
Mas Antonia	University of the Balearic Islands, Spain
Mayer Nicolas	Luxembourg Institute of Science and Technology (LIST), Luxembourg
Mesquida Calafat Antoni Lluís	University of the Balearic Islands, Spain
Munoz Mirna	CIMAT, Unidad Zacatecas, Mexico
Niemann Jörg	University of Applied Sciences Düsseldorf, Germany
Paul Alexander	University of Applied Sciences Düsseldorf, Germany
Regan Gilbert	Dundalk Institute of Technology, Ireland
Riel Andreas	Grenoble INP, France
Rodic Miran	University of Maribor, Slovenia
San Feliu Tomas	Politechnic University of Madrid (UPM), Spain
Schöning Marius	University of Applied Sciences Düsseldorf, Germany
Sechser Bernhard	Process Fellows, Germany
Stolfa Jakub	VSB Ostrava, Czech Republic
Stolfa Svatopluk	VSB Ostrava, Czech Republic
Treacy Ceara	Dundalk Institute of Technology, Ireland
Vinter Otto	Roskilde University, Denmark
Winkler Dietmar	University of Technology Vienna, Austria
Wolski Marcin	Poznan University of Technology, Supercomputing and Networking Center, Poland
Yilmaz Murat	Dublin City University, Ireland

EuroAsiaSPI 2020 Industrial Program Committee

Barafort Beatrix	Luxembourg Institute of Science and Technology (LIST), Luxembourg
Breske Eva	Bosch Engineering GmbH, Germany
Daughtrey Taz	American Society for Quality, USA
Dreves Rainer	Continental Corporation, Germany
Dussa-Zieger Klaudia	IMBUS, Germany
Ekert Damjan	ISCN GesmbH, Slovenia
Fehlmann Thomas	Euro Project Office AG, Switzerland
Griessnig Gerhard	AVL LIST GMBH, Austria
Ito Masao	Nil Software Corporation, Japan
Johansen Jorn	Whitebox, Denmark
Kaynak Onur	ASELSAN, Turkey
Keskin Kaynak Ilgi	ASELSAN, Turkey

Larrucea Uriarte Xabier	Tecnalia Spain
Lindermuth Peter	Magna Powertrain, Austria
Mayer Nicolas	Luxembourg Institute of Science and Technology (LIST) Luxembourg
Mandic Irenka	Magna Powertrain, Austria
Messnarz Richard	ISCN, Austria
Morgenstern Jens	Independent, Germany
Much Alexander	Elektrobit Automotive GmbH, Germany
Nevalainen Risto	Falconleader, Finland
Norimatsu So	JASPIC, Japan
Peisl Tom	University of Applied Sciences Munich, Germany
Poth Alexander	Volkswagen AG, Germany
Reiner Michael	IMC Krems, Austria
Sasabe Susumu	JUSE, Japan
Sauberer Gabriele	TermNet, Austria
Schweigert Tomas	ExpleoGroup Germany
Sechser Bernhard	Process Fellows GmbH, Germany
Spork Gunther	Magna Powertrain, Austria
Stefanova Pavlova Maria	CITT Global, Bulgaria
Steger Bernhardt	ISCN GesmbH, Austria
Varkoi Timo	SPINET, Finland
Wegner Thomas	ZF Friedrichshafen AG, Germany

Contents

Emerging Software Engineering Paradigms

Digitalisation of Industry, Infrastructure and E-Mobility

Visionary Paper

Cross Fertilization in Software Engineering

Ricardo Colomo-Palacios

Østfold University College, Halden, Norway
ricardo.colomo-palacios@hiof.no

Abstract. Software engineering is related to a set of disciplines both inside and outside computing. One of the aspects to consider in the development of a discipline is cross fertilization. In this paper, author reviews the cross fertilization produced by related disciplines in Software Engineering. The influences come from Computer Engineering and Computer Science inside computing outside this field quality and project management, naming just a few of them. More in particular by focusing on specific technologies, author will overview bidirectional relationships between two of the most promising technologies nowadays namely: blockchain and machine learning.

Keywords: Blockchain · Machine learning · Software engineering

1 Introduction

According to the Merriam-Webster dictionary, cross-fertilization is the interchange or interaction between different ideas, cultures, or categories especially of a broadening or productive nature. Cambridge dictionary defines the term as the mixing of the ideas, customs, etc. of different places or groups of people to produce a better result. Finally, Collins dictionary defines the term as the interaction or interchange between two or more cultures, fields of activity or knowledge, or the like, that is mutually beneficial and productive. In the two latter definitions, we find the inspiration for this paper: exploring the benefits from the associations between software engineering and some other fields inside and outside computing.

Specialization is one of the current trends in scientific work [58]. However, there are also voices warning on the risk that high specialization will lead to knowledge silos and lack of understanding or concern about works conducted in other areas, while others underlining that specialization conducted in a non-isolated way is a foundation for cross-fertilization [32].

One of the main aspects of analyzing cross fertilization is the disciplines hybridization. Coming from social sciences, Dogan and Pahre presented a set of publications from their seminal work [15] to underline the fact that the fragmentation of disciplines leaves research gaps that could be filled by means of hybridization. Dogan, the first of the couple recently underlined that hybridization appears in all fields [14].

Literature reported tons of papers on similar concepts like intradisciplinary, multidisciplinary, interdisciplinary or transdisciplinary and their effects on science. There are works devoted to analyze and compare these terms e.g. [3]. In general, authors reported several benefits from collaboration among different disciplines and fields [16, 35].

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 3–13, 2020.
https://doi.org/10.1007/978-3-030-56441-4_1

Focusing on Software engineering, the discipline is itself highly influenced by several disciplines outside computing. Software engineering presents, according to the Software Engineering Body of Knowledge (SWEBOK), strong interfaces with Management, Economics, Mathematics and Engineering fields [1]. In the joint initiative of ACM-IEEE Curriculum Guidelines for Undergraduate Degree Programs in Software Engineering [5, 33], Software Engineering also presents connections with computing, engineering, mathematics and statistics, psychology and social sciences and management. Apart from curricular efforts and bodies of knowledge, the cross-fertilization or cross pollination has been objective of study in professional magazines as well. Not in vain, twenty years ago, IEEE Software devoted a special issue to Benefits and Applications of Cross-Pollination in software engineering [39]. In this effort, authors reported several efforts and calls for a full-duplex exchange of results in a continuous way. More recently, in the same publication, a special issue devoted to Software Engineering in Society provides an overview of the connections, not necessarily in both directions, between software engineering and health, physical sciences, environmental sciences, social sciences, management, economics, computing and engineering, security, safety and privacy, policing, manufacturing, engineering emerging cyber physical systems or arts [26].

In any case, while the relationship with the rest of the computing fields is quite evident, the contact with engineering has been an arena for discussion, mostly coming from the licensing of software engineering [28, 46]. In any case, the similarities and differences with more traditional engineering are explained in the IEEE Curriculum Guidelines for Undergraduate Degree Programs in Software Engineering [5, 33], as follows:

- Software is abstract and invisible.
- Software presents static and dynamic properties alike.
- Software is complex itself in terms of its organization.
- No universal measures of quality exist for assessing a software product.
- The manufacturing cycle for software products involves the needs of distribution mechanisms also.
- Software does not wear out (maintenance is a key activity).

Mathematics & Statistics is seen as one of the foundations of the discipline according to SWEBOK [1]. In this initiative, authors claim that this area "helps software engineers comprehend this logic, which in turn is translated into programming language code". Other authors pointed out that mathematics provides a scientific basis for the discipline, leading to a deeper understanding of the development process and backing up its methods by means of mathematical techniques [9]. While there are voices in the literature underlining the fact that some software engineers do their work without applying any mathematics, good or correct software engineering is a quite difficult task to accomplish and mathematical foundations are a key to face new scenarios in which statistics and mathematics will likely play the main role e.g. big data or machine learning projects [17].

The interface between Software Engineering and Psychology is also quite established. Psychologists have been studying the behavioral aspects of the discipline since the fifties [12]. Currently, software engineering literature is full of influences of different fields of study inside psychology like developmental psychology [21], cross-cultural studies [22, 43], personality studies [41, 54, 62], emotions [44, 49, 60] or motivation [20, 50], naming just some of the most important connections.

According to SWEBOK [1], Software engineering management is the application of management activities—planning, coordinating, measuring, monitoring, controlling, and reporting—to ensure that software products and software engineering services are delivered efficiently, effectively, and to the benefits of stakeholders. Influences in the software engineering field have been pervasive and constant from the very beginning of the discipline.

Although Management and psychology are part of social sciences, the interfaces between software engineering and this branch of science are quite important. Focusing on research methods, qualitative methods used in software engineering e.g. grounded theory, surveys or case studies are normally taken from social sciences [18, 42]. Other authors from the literature have also underlined the similarities of software engineering with social sciences, given the focus on subjects [25].

In this paper, the approach taken by the author is based on the analysis of the connections of Software Engineering with other computing disciplines and more specifically with research fields that are currently in different stages in the hype, namely, blockchain and machine learning.

The remaining of this document is structured as follows. Section 2 includes the intersections between machine learning and software engineering. In Sect. 3 author reviews Blockchain and its implications with software engineering in both directions. Finally, Sect. 4 wraps up the paper and presents main insights and future prospects.

2 Machine Learning

The popularity of Machine Learning is quite high. Not in vain, a query at Google Scholar including both terms produces more than 150,000 results by March 2020. Machine Learning alone as a search string is producing more than 3 million results in this academic search engine.

Machine learning is aimed to answer these two questions [24], firstly, how can we build computer-based systems that automatically improve through experience? And secondly, what are the fundamental statistical-computational-information-theoretic laws that govern all learning systems, including computers, humans, and organizations? As a field of study, Machine Learning combines several disciplines including statistics, mathematics, engineering, biology, neuroscience and computer science [38]. Machine Learning applications work and optimize their performance using example data from past experiences [2]. It starts with the definition of a predictive model based on some configuration parameters and from that, the learning comes in the execution of a computer program to optimize the model in a predictive (make predictions), descriptive (gain knowledge from data) or combined way. In other words, the machine learning focus is on making computers modify or adapt their actions in the search of better accuracy (measured by how well actions reflect the correct ones) [38].

Machine learning is based on algorithms and their capacity to learn a model. Learning techniques are often divided into supervised, unsupervised, reinforcement and evolutionary learning. In supervised learning, both examples of inputs and outputs are provided and the task of the algorithm is constructing a mapping from one to the other [53]. In unsupervised learning, there are not available sample correct responses and the task is processing the input data in the search of underlying patterns and categories. Reinforcement learning combines both the approaches [38], there is information whether an answer is wrong but not how to correct it, so it is needed to evaluate input–output pairs and hence discover, through trial and error, the optimal outputs for each input [53]. Finally, evolutionary learning is an approach inspired by biology and natural evolution. In this approach, the algorithm maintains a population of candidates, which are compared with the output. Then, through multiple generations of variation, selection, and reproduction, the population adapts to the selection criterion (the relative distance from the desired outcome) and produces fitter solutions [23].

To provide a structured review of the influence of the field, author used Google Scholar to search for the 10 knowledge areas defined in the 2004 edition of SWEBOK. Author decided to exclude the new five knowledge areas present in SWEBOK V3 (Software engineering professional practice, Software engineering economics, Computing foundations, Mathematical foundations and Engineering foundations), given their lack of specificity for tools.

Table 1. Number of hits of machine learning per SWEBOK 2004 knowledge areas

SWEBOK knowledge area	# Google scholar results
1 Software requirements	7760
2 Software design	18900
3 Software construction	1140
4 Software testing	15900
5 Software maintenance	12400
6 Software configuration management	833
7 Software engineering management	399
8 Software engineering process	1380
9 Software engineering models and methods	33
10 Software quality	16100

While it is true that the labeling of the knowledge areas could influence in deep the number of results (e.g. Software construction) and that a simple query is not able to give a full description of the influence, it is also unquestionable that some knowledge areas are more affected than others and that, overall, the penetration of machine learning in software engineering is not superficial. In order to see the evolution of this influence, Table 2 includes the results of the ten knowledge areas in the period 2011–2019.

Table 2. Evolution on the number of hits of machine learning per SWEBOK 2004 knowledge areas

KA	2011	2012	2013	2014	2015	2016	2017	2018	2019
1	258	304	362	497	510	745	805	1130	1190
2	709	812	960	1080	1250	1450	1750	2270	2780
3	43	54	47	67	66	71	81	103	141
4	521	631	788	964	1170	1400	1640	2130	2700
5	442	541	607	791	951	1100	1300	1650	2010
6	38	40	35	69	77	56	69	67	72
7	23	26	22	43	34	17	14	24	37
8	67	62	74	103	78	113	105	129	136
9	0	0	0	3	2	0	2	4	11
10	590	721	801	976	1190	1390	1680	2220	2630
	2691	3191	3696	4593	5328	6342	7446	9727	11707

A review of the last nine years shows a significant increase in the number of works published and available in Google Scholar with yearly increments of around 20% in almost all the knowledge areas.

The influence of Machine Learning applications in Software Engineering practice is quite apparent and the number of papers devoted to the topic is vast. The popularity of the so called ML4SE (Machine Learning for Software Engineering) has driven to a good set of courses, conferences and workshops devoted to the field. The popularity of the subject leads further to a set of tertiary studies on the topic devoted to aspects like effort estimation [57], software fault prediction [36], software optimization [40] or code smell detection [6], citing just some of the relevant examples.

If we switch direction towards SE4ML (Software engineering for Machine Learning), the road ahead was described back in 2018 [27] as a result of the First Symposium on Software Engineering for Machine Learning Applications. In their statement, authors claim that Machine Learning systems are difficult to test and verify, given that Machine Learning based applications are built on rules inferred from training data. Another relevant initiative is the International Workshop on Machine Learning and Software Engineering in Symbiosis. In this workshop, Software Engineering and Machine Learning communities were encouraged to work together to solve the critical aspects of assuring the quality of artificial intelligence and software systems. As a result of their discussions, they stated that the combined knowledge of Software engineering and Machine Learning is required to answer the key questions regarding the integration of Machine Learning pipelines into software development processes and identifying the desired new roles to address respective challenges.

3 Blockchain

Blockchain is a technology that because of its potentials could be ubiquitous [31]. Blockchain technology was first implemented in Bitcoin by Nakamoto back in 2008. From the technological standpoint, it consists of a sequence of blocks, each of which holds a complete list of transaction records like the conventional public ledger [61]. Thus, Blockchain is seen as a distributed ledger technology (DLT) that supports collaborative processes by means of a shared, distributed and trusted dataset implementing also point-to-point transmission, multi-node collective maintenance, consensus mechanisms and encryption algorithms [7]. Blockchain benefits include aspects like decentralization, persistency, anonymity and auditability [65]. However, blockchain is not a perfect technology and several challenges and limitations have been pointed out. Concerns on energy consumption, privacy, scalability and connectivity, among others, are reported in the literature. A recent and in deep review on challenges and issues is available at [29].

However, Blockchain has evolved in its features, and as a consequence, several generations of blockchain is developed. Blockchain 1.0 is the seminal Blockchain attached to crypto-currencies applications. Blockchain 2.0 started with Ethereum back in 2013 and provides a wider range of application scenarios by using the distributed ledger of blockchain to record, confirm and transfer various forms of contracts and properties [55]. This new generation includes smart contracts as one of its main features, a recent review on the topic can be found at [64]. Blockchain 3.0 includes a vast array of applications including art, health, and science, among others [19]. Blockchain 3.0 is aimed to enable interoperability [52] and increase network speed. It also incorporates features like immutability, transparency and no need for intermediaries, obtained by the blockchain trustless decentralization to other systems which are built on top of blockchain technology [13]. Finally, there is a new generation in sight [4], Blockchain 4.0 includes artificial intelligence as part of the platform, reducing the need of human management by enabling functions to make decisions and act on systems.

Given the current hype, the relationship between Blockchain and software engineering is quite broad in both directions. Although by early March 2020 a simple query in Google Scholar with both terms just retrieves 564 results, the cross fertilization present in both fields is a growing research area. Moreover, it is also true that there is an increasing interest in the development of a dedicated field inside software engineering for blockchain-oriented applications, so called Blockchain-based software engineering [8, 47]. Examples of this influence can be found in aspects like designing [37], architecting [59], modelling [48], programming [10, 11] or testing [45].

Conversely, there are also influences of blockchain in the software engineering research field. Blockchain has been used to improve the integrity of the software development process [63], to complement agile practices [34], service composition [56], to enable distributed teams [51] or as a support for collaborative software teams [30].

As presented in the case of Machine Learning, Table 3 presents the total number of hits per knowledge area in Blockchain and Table 4 presents results of the ten knowledge areas in the period 2011–2019.

Table 3. Number of hits of blockchain per SWEBOK 2004 knowledge areas

SWEBOK knowledge area	# Google scholar results
1 Software requirements	346
2 Software design	886
3 Software construction	52
4 Software testing	598
5 Software maintenance	379
6 Software configuration management	53
7 Software engineering management	8
8 Software engineering process	65
9 Software engineering models and methods	2
10 Software quality	694

As underlined in the case of Table 1, there are knowledge areas that initially could receive less attention because of its labelling or its nature. In any case, champions are again, requirements, testing, maintenance quality and overall design.

Table 4. Evolution on the number of hits of blockchain per SWEBOK 2004 knowledge areas

KA	2011	2012	2013	2014	2015	2016	2017	2018	2019
1	0	2	1	0	5	7	33	104	144
2	3	6	14	15	11	28	99	214	375
3	0	0	0	0	1	1	3	8	23
4	3	1	5	6	9	12	57	142	266
5	2	0	1	5	7	12	42	85	167
6	1	0	1	4	3	1	4	4	16
7	0	0	0	0	0	0	2	1	5
8	0	0	2	0	2	2	9	14	20
9	0	0	0	1	0	0	0	0	1
10	2	2	3	6	5	12	56	168	345
	11	11	27	37	43	75	305	740	1362

Regarding the evolution over time, in the case of Blockchain, the progression has accelerated from 2017 with an extensive increment of around 200% yearly. This could be rooted in the novelty of the technology that in the early 2010s was in the Blockchain 1.0 phase and expanded its applications just some years ago.

To sum up the interaction of blockchain and machine learning, in the near future, both the generalization of blockchain-based solutions and the advance on the software engineering practices will lead to a more mature hybrid field and a more intense cross-fertilization.

4 Conclusions

In this paper, author presented the cross-fertilization between software engineering and two different knowledge areas: Machine Learning and Blockchain. Being both the topics among the technologies in the hype, it is unquestionable the higher influence and repercussion of Machine Learning in Software Engineering and vice versa. However, Blockchain technologies are beginning to be more mature and their interchange with the Software Engineering field is also increasing in deep, specially from 2018.

Both Machine learning and Blockchain fields are advancing and new prospects will impact Software Engineering in the next coming years. For instance, in the Machine Learning field, Adaptive machine learning, as underlined by Gartner, will provide a plus to machine learning-based systems. Adaptive machine learning is about retraining ML models when they are in their runtime environment frequently. This will impact, in deep, not only Software Engineering tools but also will need new software engineering approach to govern this continuous training. Regarding Blockchain, the new wave 4.0 that includes artificial intelligence as part of the platform will be an opportunity to expand the influence of the technology on Software engineering practices and also to connect blockchain-based methods with Software engineering processes.

References

1. Abran, A., Fairley, D.: SWEBOK: GUIDE to the Software Engineering Body of Knowledge Version 3. IEEE Computer Society, Los Alamitos (2014)
2. Alpaydin, E.: Introduction to Machine Learning. MIT Press, Cambridge (2020)
3. Alvargonzález, D.: Multidisciplinarity, interdisciplinarity, transdisciplinarity, and the sciences. Int. Stud. Philos. Sci. **25**(4), 387–403 (2011). https://doi.org/10.1080/02698595.2011.623366
4. Angelis, J., Ribeiro da Silva, E.: Blockchain adoption: a value driver perspective. Bus. Horiz. **62**(3), 307–314 (2019). https://doi.org/10.1016/j.bushor.2018.12.001
5. Ardis, M., et al.: SE 2014: curriculum guidelines for undergraduate degree programs in software engineering. Computer **48**(11), 106–109 (2015). https://doi.org/10.1109/MC.2015.345
6. Azeem, M.I., et al.: Machine learning techniques for code smell detection: a systematic literature review and meta-analysis. Inf. Softw. Technol. **108**, 115–138 (2019). https://doi.org/10.1016/j.infsof.2018.12.009
7. Bai, C., Sarkis, J.: A supply chain transparency and sustainability technology appraisal model for blockchain technology. Int. J. Prod. Res. **58**, 2142–2162 (2020). https://doi.org/10.1080/00207543.2019.1708989
8. Beller, M., Hejderup, J.: Blockchain-based software engineering. In: Proceedings of the 41st International Conference on Software Engineering: New Ideas and Emerging Results, pp. 53–56. IEEE Press, Montreal (2019). https://doi.org/10.1109/ICSE-NIER.2019.00022
9. Broy, M.: Mathematics of software engineering. In: Möller, B. (ed.) MPC 1995. LNCS, vol. 947, pp. 18–48. Springer, Heidelberg (1995). https://doi.org/10.1007/3-540-60117-1_3
10. Chakraborty, P., et al.: Understanding the software development practices of blockchain projects: a survey. In: Proceedings of the 12th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement, pp. 1–10. Association for Computing Machinery, Oulu (2018). https://doi.org/10.1145/3239235.3240298

11. Coblenz, M.: Obsidian: a safer blockchain programming language. In: 2017 IEEE/ACM 39th International Conference on Software Engineering Companion (ICSE-C). pp. 97–99 (2017). https://doi.org/10.1109/ICSE-C.2017.150
12. Curtis, B.: Fifteen years of psychology in software engineering: individual differences and cognitive science. In: Proceedings of the 7th International Conference on Software Engineering, pp. 97–106. IEEE Press, Orlando (1984)
13. Di Francesco Maesa, D., Mori, P.: Blockchain 3.0 applications survey. J. Parallel Distrib. Comput. **138**, 99–114 (2020). https://doi.org/10.1016/j.jpdc.2019.12.019
14. Dogan, M.: Creative Marginality: Innovation at the Intersections of Social Sciences. Routledge, Abingdon (2019)
15. Dogan, M., Pahre, R.: Fragmentation and recombination of the social sciences. Stud. Comp. Int. Dev. **24**(2), 56–72 (1989). https://doi.org/10.1007/BF02687172
16. Domik, G., Fischer, G.: Coping with complex real-world problems: strategies for developing the competency of transdisciplinary collaboration. In: Reynolds, N., Turcsányi-Szabó, M. (eds.) KCKS 2010. IAICT, vol. 324, pp. 90–101. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-15378-5_9
17. Dougherty, J.P.: MATH COUNTS: where mathematics meets software engineering. ACM Inroads **8**(3), 13–15 (2017). https://doi.org/10.1145/3123734
18. Dybå, T., et al.: Qualitative research in software engineering. Empir. Softw. Eng. **16**(4), 425–429 (2011). https://doi.org/10.1007/s10664-011-9163-y
19. Efanov, D., Roschin, P.: The all-pervasiveness of the blockchain technology. Procedia Comput. Sci. **123**, 116–121 (2018). https://doi.org/10.1016/j.procs.2018.01.019
20. França, A.C.C., et al.: Motivation in software engineering industrial practice: a cross-case analysis of two software organisations. Inf. Softw. Technol. **56**(1), 79–101 (2014). https://doi.org/10.1016/j.infsof.2013.06.006
21. Gren, L., et al.: The perceived effects of group developmental psychology training on agile software development teams. IEEE Softw. (2019). https://doi.org/10.1109/MS.2019.2955675
22. Hoda, R., et al.: Socio-cultural challenges in global software engineering education. IEEE Trans. Educ. **60**(3), 173–182 (2017). https://doi.org/10.1109/TE.2016.2624742
23. Hu, T., et al.: An evolutionary learning and network approach to identifying key metabolites for osteoarthritis. PLOS Comput. Biol. **14**(3), e1005986 (2018). https://doi.org/10.1371/journal.pcbi.1005986
24. Jordan, M.I., Mitchell, T.M.: Machine learning: trends, perspectives, and prospects. Science **349**(6245), 255–260 (2015). https://doi.org/10.1126/science.aaa8415
25. Juristo, N., Moreno, A.M.: Basics of Software Engineering Experimentation. Springer, Heidelberg (2013)
26. Kazman, R., Pasquale, L.: Software engineering in society. IEEE Softw. **37**(1), 7–9 (2020). https://doi.org/10.1109/MS.2019.2949322
27. Khomh, F., et al.: Software engineering for machine-learning applications: the road ahead. IEEE Softw. **35**(5), 81–84 (2018). https://doi.org/10.1109/MS.2018.3571224
28. Knight, J.C., Leveson, N.G.: Should software engineers be licensed? Commun. ACM **45**(11), 87–90 (2002)
29. Kolb, J., et al.: Core concepts, challenges, and future directions in blockchain: a centralized tutorial. ACM Comput. Surv. **53**(1), 9:1–9:39 (2020). https://doi.org/10.1145/3366370
30. Król, M., et al.: ChainSoft: collaborative software development using smart contracts. In: Proceedings of the 1st Workshop on Cryptocurrencies and Blockchains for Distributed Systems, pp. 1–6. ACM, New York (2018). https://doi.org/10.1145/3211933.3211934

31. Lao, L., et al.: A survey of IoT applications in blockchain systems: architecture, consensus, and traffic modeling. ACM Comput. Surv. **53**(1), 18:1–18:32 (2020). https://doi.org/10.1145/3372136
32. Leahey, E., Reikowsky, R.C.: Research specialization and collaboration patterns in sociology. Soc. Stud. Sci. **38**(3), 425–440 (2008). https://doi.org/10.1177/0306312707086190
33. LeBlanc, R.J., Sobel, A.: Software Engineering 2014: Curriculum Guidelines for Undergraduate Degree Programs in Software Engineering. IEEE Computer Society, Los Alamitos (2014)
34. Lenarduzzi, V., et al.: Blockchain applications for agile methodologies. In: Proceedings of the 19th International Conference on Agile Software Development: Companion, pp. 1–3. Association for Computing Machinery, Porto (2018). https://doi.org/10.1145/3234152.3234155
35. Madni, A.M.: Transdisciplinarity: reaching beyond disciplines to find connections. J. Integr. Des. Process Sci. **11**(1), 1–11 (2007)
36. Malhotra, R.: A systematic review of machine learning techniques for software fault prediction. Appl. Soft Comput. **27**, 504–518 (2015). https://doi.org/10.1016/j.asoc.2014.11.023
37. Marchesi, M., et al.: An agile software engineering method to design blockchain applications. In: Proceedings of the 14th Central and Eastern European Software Engineering Conference Russia, pp. 1–8. Association for Computing Machinery, Moscow (2018). https://doi.org/10.1145/3290621.3290627
38. Marsland, S.: Machine Learning: An Algorithmic Perspective, 2nd edn. CRC Press, Boca Raton (2015)
39. Matsubara, T., Ebert, C.: Guest editor's introduction: benefits and applications of cross-pollination. IEEE Softw. **17**(1), 24 (2000)
40. Memeti, S., Pllana, S., Binotto, A., Kołodziej, J., Brandic, I.: Using meta-heuristics and machine learning for software optimization of parallel computing systems: a systematic literature review. Computing **101**(8), 893–936 (2018). https://doi.org/10.1007/s00607-018-0614-9
41. Mendes, F.F., et al.: The relationship between personality and decision-making: a systematic literature review. Inf. Softw. Technol. **111**, 50–71 (2019). https://doi.org/10.1016/j.infsof.2019.03.010
42. Méndez Fernández, D., Passoth, J.-H.: Empirical software engineering: from discipline to interdiscipline. J. Syst. Softw. **148**, 170–179 (2019). https://doi.org/10.1016/j.jss.2018.11.019
43. Niazi, M., et al.: Software process improvement barriers: a cross-cultural comparison. Inf. Softw. Technol. **52**(11), 1204–1216 (2010). https://doi.org/10.1016/j.infsof.2010.06.005
44. Novielli, N., Serebrenik, A.: Sentiment and emotion in software engineering. IEEE Softw. **36**(5), 6–23 (2019). https://doi.org/10.1109/MS.2019.2924013
45. Parizi, R.M., et al.: Empirical vulnerability analysis of automated smart contracts security testing on blockchains. In: Proceedings of the 28th Annual International Conference on Computer Science and Software Engineering, pp. 103–113. IBM Corp., Markham (2018)
46. Parnas, D.: Software engineering - missing in action: a personal perspective. Computer **44**(10), 54–58 (2011). https://doi.org/10.1109/MC.2011.268
47. Porru, S., et al.: Blockchain-oriented software engineering: challenges and new directions. In: 2017 IEEE/ACM 39th International Conference on Software Engineering Companion (ICSE-C), pp. 169–171 (2017). https://doi.org/10.1109/ICSE-C.2017.142

48. Rocha, H., Ducasse, S.: Preliminary steps towards modeling blockchain oriented software. In: Proceedings of the 1st International Workshop on Emerging Trends in Software Engineering for Blockchain, pp. 52–57. ACM, New York (2018). https://doi.org/10.1145/3194113.3194123
49. Sánchez-Gordón, M., Colomo-Palacios, R.: Taking the emotional pulse of software engineering—a systematic literature review of empirical studies. Inf. Softw. Technol. **115**, 23–43 (2019). https://doi.org/10.1016/j.infsof.2019.08.002
50. Sharp, H., et al.: Models of motivation in software engineering. Inf. Softw. Technol. **51**(1), 219–233 (2009). https://doi.org/10.1016/j.infsof.2008.05.009
51. Singi, K., et al.: Compliance adherence in distributed software delivery: a blockchain approach. In: 2018 IEEE/ACM 13th International Conference on Global Software Engineering (ICGSE), pp. 126–127 (2018)
52. Siris, V.A., et al.: Interledger approaches. IEEE Access **7**, 89948–89966 (2019). https://doi.org/10.1109/ACCESS.2019.2926880
53. Smith, A.J.: Applications of the self-organising map to reinforcement learning. Neural Netw. **15**(8), 1107–1124 (2002). https://doi.org/10.1016/S0893-6080(02)00083-7
54. Soomro, A.B., et al.: The effect of software engineers' personality traits on team climate and performance: a systematic literature review. Inf. Softw. Technol. **73**, 52–65 (2016). https://doi.org/10.1016/j.infsof.2016.01.006
55. Wang, N., et al.: When energy trading meets blockchain in electrical power system: the state of the art. Appl. Sci. **9**(8), 1561 (2019). https://doi.org/10.3390/app9081561
56. Wang, P., et al.: QoS-aware service composition using blockchain-based smart contracts. In: Proceedings of the 40th International Conference on Software Engineering: Companion Proceedings, pp. 296–297. ACM, New York (2018). https://doi.org/10.1145/3183440.3194978
57. Wen, J., et al.: Systematic literature review of machine learning based software development effort estimation models. Inf. Softw. Technol. **54**(1), 41–59 (2012). https://doi.org/10.1016/j.infsof.2011.09.002
58. Wenger, E.: Communities of Practice: Learning, Meaning, and Identity. Cambridge University Press, Cambridge (1999)
59. Wessling, F., Gruhn, V.: Engineering software architectures of blockchain-oriented applications. In: 2018 IEEE International Conference on Software Architecture Companion (ICSA-C), pp. 45–46 (2018). https://doi.org/10.1109/ICSA-C.2018.00019
60. Wrobel, M.R.: Applicability of emotion recognition and induction methods to study the behavior of programmers. Appl. Sci. **8**(3), 323 (2018). https://doi.org/10.3390/app8030323
61. Xie, S., et al.: Blockchain for cloud exchange: a survey. Comput. Electr. Eng. **81**, 106526 (2020). https://doi.org/10.1016/j.compeleceng.2019.106526
62. Yilmaz, M., et al.: An examination of personality traits and how they impact on software development teams. Inf. Softw. Technol. **86**, 101–122 (2017). https://doi.org/10.1016/j.infsof.2017.01.005
63. Yilmaz, M., Tasel, S., Tuzun, E., Gulec, U., O'Connor, R.V., Clarke, P.M.: Applying blockchain to improve the integrity of the software development process. In: Walker, A., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 260–271. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_20
64. Zheng, Z., et al.: An overview on smart contracts: challenges, advances and platforms. Future Gener. Comput. Syst. **105**, 475–491 (2020). https://doi.org/10.1016/j.future.2019.12.019
65. Zheng, Z., et al.: Blockchain challenges and opportunities: a survey. Int. J. Web Grid Serv. **14**(4), 352–375 (2018). https://doi.org/10.1504/IJWGS.2018.095647

SPI Manifesto and Improvement Strategies

Expanding the Tactical Level in ISO/IEC 33014 to Deal with a Broader Set of Change Initiatives

Peter H. Carstensen[1] and Otto Vinter[1,2(✉)]

[1] Department of Computer Science, University of Copenhagen, 2200 Copenhagen N, Denmark
peter.h.carstensen@di.ku.dk

[2] Department of People and Technology, Roskilde University, 4000 Roskilde, Denmark
otv@ruc.dk, vinter@ottovinter.dk

Abstract. The ISO/IEC 33014 is an excellent guide for planning and implementing process improvement initiatives. The standard, however, only deals with improvement initiatives that are based on assessments of an organization's processes. For the standard to be capable of dealing with a broader set of change initiatives the tactical level in the standard needs to be expanded with a new perspective. In this paper we present a proposal for such an expansion. The central part of this expansion is the introduction of a change-charter that defines the directions, policies and guidelines for these change initiatives. We propose that the charter be built on the nine change-aspects that we have previously aggregated through a comprehensive literature study of prominent change management literature.

Keywords: Tactical approach to change · Change charter · Change-charter · ISO/IEC 33014 · Change plan development · Action plans · Aspects of change · Change-aspects · Change strategies · Change initiatives

1 Introduction

Performing changes in an organization is a challenge irrespective of whether you introduce a new IT-product, perform an organizational implementation as a result of a development project, or want to change the structures, processes and routines in the organization. They are all about changing people's way of working and consequently inherently complex, difficult and often prone to failure.

The ISO/IEC 33014 standard (Guide for Process Improvement) [1] "provides guidance on using process assessment as part of a complete framework for performing process improvement as part of a continual improvement activity."

The standard introduces a framework, a process, methods and guides that aim at strengthening an organization's ability to carry out improvements. The framework operates with three levels of process improvement: Strategic, tactical and operational. The complete framework is shown in Fig. 1.

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 17–29, 2020.
https://doi.org/10.1007/978-3-030-56441-4_2

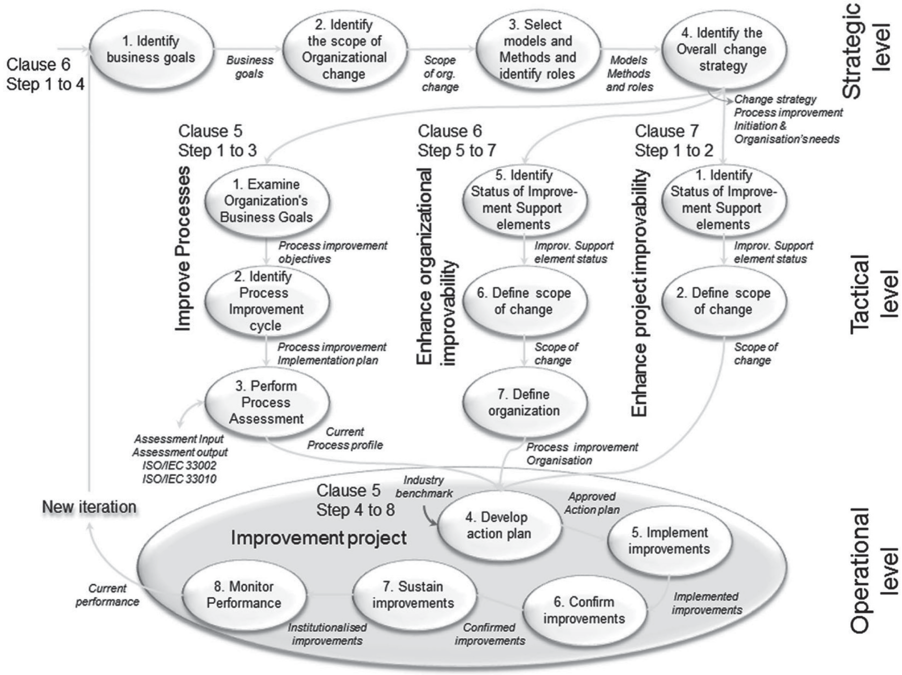


Fig. 1. Overview of the ISO/IEC 33014 guide for process improvement (Fig. 1 in [1]).

At the strategic level the standard advises an organization to define its business goals, scope of organizational change, models/methods/roles involved, and identify the change strategy applicable for specific situation/context of the organization.

At the tactical level the standard proposes three improvement perspectives: Improve processes; enhance organizational improvability; and enhance project improvability. The perspective that an organization selects depends on whether the organization wishes to initiate a process improvement programme, enhance the overall improvability of the organization, or ensure a process improvement project's success.

At the operational level the improvement projects are planned, implemented and monitored. After this a new cycle of improvements can be performed initiated from the strategic level – constituting a continual improvement activity. The standard thus represents a significant contribution to how improvement activities should be organized, planned and implemented.

However, the focus of the standard is on improvements that have been selected exclusively based on assessments of the organization's processes (clause 5 step 3, clause 6 step 5, and clause 7 step 1 in Fig. 1).

We find that the applicability of the standard could be significantly enhanced if expanded to deal with other types of changes in an organization than those based on assessments. We therefore propose an additional perspective to the tactical level in the ISO/IEC 33014 standard and this perspective will be elaborated on in this paper.

Most needs for changes in an organization arise from many other sources than an assessment of processes. For example the needs can arise from changes in the market, changes in laws and regulations, or new technology requiring new ways of working or new business models. These changes will often lead to a need for planning and implementing changes in the organization structures, values streams, products, services, tool implementations, internal processes or ways of working together.

For all these types of changes it could be advantageous for an organization to follow the steps at the strategic level of the standard (clause 6 step 1–4 in Fig. 1). Also the planning and implementing of a change could gain from following the steps at the operational level (clause 5 step 4–8 in Fig. 1).

However, there is little (or no) help for management or change-team at the tactical level of the standard to deal with these other types of change initiatives. This is the reason why an additional tactical perspective is needed. But which steps should this perspective consist of and how should the work at this level be structured?

We propose the perspective to consist of 3 steps and introduce the concept of a charter for change initiatives where the directions, policies and guidelines for the detailed planning and execution of the change are defined. Furthermore we propose that this charter is structured around the nine change-aspects that we have previously aggregated through a comprehensive literature study of prominent change management literature. The nine change-aspects are listed in Table 1.

The rest of the paper is organized as follows: Sect. 3 is central as it presents the proposed expansion of the tactical level in the ISO/IEC 33014 standard. But first in Sect. 2 we describe the research approach we have applied. In Sect. 4 we provide an overview of related research and practice, which can be seen as alternatives to our approach. Finally we conclude and propose future work.

2 Research Approach

Our original research idea was to deliver to management and change-teams a set of recommendations for action that they could build their change plans on. The recommendations should be based on what a series of prominent authors of literature on (change) management have found essential. The main approach to this part of our work could be understood as analytic induction.

We began by extracting recommended actions from the literature behind two of the ten ISO/IEC 33014 strategies [1] (Optionality and Specialist-driven). We read the texts end-to-end and extracted statements by the author(s) that seemed characteristic for the strategy. The extracted recommendations were kept as close to the original statements in the texts as possible. We looked into the resulting recommended actions and from this we initially identified eight central aspects that evidently need careful considerations when planning and executing changes. We coined these central aspects of change: change-aspects. We then hypothesized that the eight change-aspects would be applicable also to the recommendations for the remaining eight change strategies. We discussed each change-aspect and defined them properly. We quickly realized that the change-aspects were not completely orthogonal, which was never our intension. To validate the hypothesis we repeated the same process for first two more of the ten strategies (Production-organized and Socializing) and then the remaining six strategies. The extracted recommendations from the literature were now analyzed whether they could be allocated to the identified change-aspects. We found this rather easy, which seemed to confirm the validity and applicability of the change-aspects.

In total the extraction of recommended actions from the literature has until now resulted in a gross list of more than 700 recommendations. Of these we have selected those recommendations that were most clearly indicative of each the ten strategies. This selection resulted in a total of 257 recommendations covering all of the change-aspects with 37 to 72 recommendations for each of the ten overall change strategies. In addition to this we also found 79 recommendations that could be applied to many strategies. The literature review is still ongoing (although rather few new titles are added), so the numbers might increase slightly over the next year.

The eight change-aspects were documented in [2] and published more widely in [3]. For a detailed description of the change-aspects please consult our paper [4].

Following this we evaluated the change-aspects and their relevance, usefulness and applicability through a number of sources. We discussed it with our research partners involved in the research project in which the ten strategies were originally identified. We had a number of masters students specializing in change management apply the change-aspects on their projects on different "real-life" change situations in their organizations. Finally we presented our change-aspects for change management practitioners at company seminars, experience exchange meetings etc. These evaluations have resulted in small changes in the understanding and definitions of some of the change-aspects.

Lately we have decided to split one of the original change-aspects (Methods & Techniques) into two because it was too broadly defined. The two change-aspects that replace it are: Attitude & Behavior and Maintaining Focus. The splitting was based on a renewed review of the existing recommendations looking for patterns. The resulting nine change-aspects are listed in Table 1 along with a short definition.

Following the work of identifying and defining the change-aspects, we have worked on how the change-aspects are best integrated into "real life" change processes in practice. This has been done by studying suggestions and models in existing approaches (e.g. the ISO 33014) and discussions with practitioners on how the change-aspects could be integrated into their approaches and ways of working with changes on a daily basis.

Through these discussions and trial applications we have come to realize that the change-aspects not only could be used to assist in the detailed planning, but could also be used proactively by management and change-teams to structure the work on setting the direction of a change i.e. be instrumental to the tactical level.

Looking at the collected recommendations from literature it became clear to us that many of them actually were more relevant at this level. This has led us to propose the use of the change-aspects to structure the work of setting directions, policies and guidelines for a change. Ultimately this has led to our proposal for expanding the tactical level of the ISO/IEC 33014 standard.

Table 1. The nine change-aspects aggregated from the change management literature.

Change-aspect	Short definition
Attitude and behavior	The attitude and behavior that should be exercised or demonstrated by the change manager/team during the execution of the change
Communication	Types of information that should be communicated, to/by whom, when/through which channels
Competences and training	The competences management and change-team should have before the work on the change is initiated, or which should be built up on the way
Culture	The culture that should be established/changed/supported in the organization to secure the success of the desired change
Decision-making	Decisions and commitments it is important to make and achieve before and during the work on the change, who should make them, when, and within what scope they can be made
Knowledge acquisition	The knowledge that should be obtained/gained/collected before and during the work on the change
Maintaining focus	How to maintain the direction and progress of the change focusing on objectives, outcomes and other results
Organizing	How the change-team as well as all involved or affected by it should be organized/structured, and their roles in the change defined
Processes and plans	Concrete processes and plans that management and change-team establish for the execution of the whole and/or specific parts of the work

3 The Proposed Expansion to the Tactical Level in ISO/IEC 33014

As mentioned above the focus of the ISO/IEC 33014 [1] standard is on process improvements that have been selected exclusively based on assessments of the organization's processes.

Employing assessments of processes is a good way of selecting and prioritizing the improvement work in an organization. However, most change initiatives arise from other sources. We therefore propose to expand the tactical level in the standard with a perspective for improving change initiatives from other sources than those based on assessments (i.e. a new clause 8 in the standard).

As these change initiatives have not been defined and prioritized by an assessment we propose that management and change-team instead develop a charter for the change. This change-charter should set the direction and establish a set of policies and guidelines of the change to be planned and implemented.

To assist in developing this change-charter we propose to structure it using the nine change-aspects that we have aggregated from the study of change management literature (see Sect. 2). They represent a categorization of the most widely recognized recommendations for action by major authors writing about how to organize and execute changes in a wide range of organizational contexts and change situations.

The proposed expansion to the tactical level in the ISO/IEC 33014 standard consists of the following steps:

Step 1 – Identify status of change initiative
Step 2 – Set directions for the change
Step 3 – Define scope of change – and what to change

In Fig. 2 we have inserted the steps in the structure of the standard. The steps are described in the following subsections.

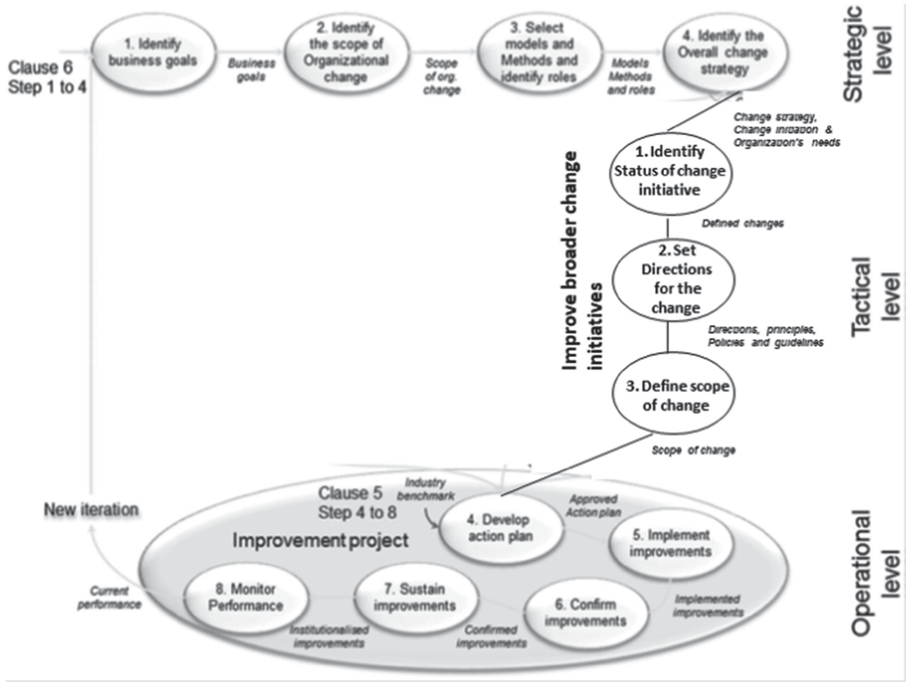


Fig. 2. The main steps in the expansion to the target level in ISO/IEC 33014 for dealing with the improvement of a broader set of change initiatives.

3.1 Step 1 – Identify Status of Change Initiative

Defining the changes needed i.e. the vision for the future is the responsibility of (top level) management. The change initiative should be aligned with the business goals and within the scope of the organizational change (clause 6 step 1 and 2 in Fig. 2). Models, methods, roles and the overall change strategy must have been identified (clause 6 step 3 and 4 in Fig. 2).

The step must be performed by management before embarking on the route to planning and implementing the change initiative. It is important that the desired end-state (after the changes have been implemented) is defined clearly, i.e. which specific goals to achieve.

3.2 Step 2 – Set Directions for the Change

After the status of the change initiative has been identified, management must elaborate on the vision, intent and idea of the change initiative making it clear how the changes to be initiated best meet the business needs of the organization, how the overall change strategies should be applied, and how the execution of the changes should be approached in the organization.

In order to assist management in setting the directions for the detailed planning and implementation of the change we propose that a charter for the change be developed. The structure of this change-charter should be based on the nine change-aspects that we identified and validated as described above in Sect. 2.

For each of the change-aspects listed in Table 1 management must discuss how they and the change-team will ensure that this change-aspect is properly addressed in the subsequent steps at the operational level. Through this discussion they should set the direction; establish policies and guidelines for the detailed planning and implementation of the change; i.e. develop the contents of the change-charter.

In Table 2 is shown how a partly completed hypothetical change-charter might look giving examples of directions for each aspect. The examples presented in Table 2 are taken from the recommendations we found in the change management literature and should therefore be replaced in practice by detailed and contextualized directions for the particular organization and change initiative.

Table 2. Partly completed hypothetical change-charter with examples from literature for each of the nine change-aspects.

Change-aspect	Examples of directions, policies and guidelines
Attitude and behavior	– We will act with benevolence to both victims and survivors e.g. through voluntary separation and generous compensation ([5] p. 614) – We will co-ordinate, advise and manage instead of keeping control centrally ([6] p. 39) – We will create situations in which curiosity and need are driving forces for improvement ([7] p. 91) – We will treat …
Communication	– We will develop and publish clear documented corporate beliefs and purpose - a mission statement ([8] p. 36) – We will establish and maintain a shared vision ([7] p. 96) – We will meet the employees face-to-face and communicate our vision ([5] p. 612) – We will inform …
Competences and training	– We will develop leadership skills among middle managers to install new values in them ([5] p. 620) – We will develop “cookbooks”, manuals and train employees in the standardized routines ([9] p. 40) – We will train change agents to become competent in interpersonal inquiry ([5] p. 608) – We will ensure …

(*continued*)

Table 2. (*continued*)

Change-aspect	Examples of directions, policies and guidelines
Culture	– We will encourage individuals to establish improvement goals for themselves and their groups ([8] p. 19) – We will establish a sense of urgency to gain the needed cooperation for the change ([10] p. 37) – We will take steps to reduce the incidence of political behavior and potential conflicts ([11] p. 79) – We will engage in …
Decision-making	– We will base our management decisions on a long term philosophy, even at the expense of short term financial goals ([12] p. 5) – We will drive responsibility downwards to set-off self-organization and innovation ([11] p. 282) – We will give the employees access to relevant information to encourage their participation in decision making ([13] p. 171) – We will delegate …
Knowledge acquisition	– We will ensure that the impact of existing initiatives are evaluated and considered carefully before starting another series of improvement initiatives ([8] p. 159) – We will identify the process owners who are responsible for the business process ([14] p. 119) – We will identify change enablers (e.g. technological opportunities, constraining technological factors or constraining human factors) ([8] p. 48) – We will assess …
Maintaining focus	– We will set standards or targets for performance, or expected outcomes; and take corrective action to remove any deviations from these ([11] p. 78) – We will ensure that people assigned to the change activities are not pulled away from these responsibilities ([15] p. 77) – We will apply risk management in order to identify, carefully evaluate and effectively mitigate risks involved in the change ([16] p. 13) – We will check …
Organizing	– We will ensure that leadership of the change belongs to one small group of people typically located at the top of the hierarchy ([5] p. 605) – We will create loosely coupled organizations, where the experimenting units are highly buffered (separated) from the exploiting units ([17] p. 247) – We will ensure that senior management are on the board and play an active role in the change teams ([15] p. 74) – We will structure …
Processes and plans	– We will ensure control at all stages of planning and operationalization of the strategy ([18] p. 55) – We will find the optimal rate of change rather than the fastest ([19] p. 62) – We will overhaul processes which create a vicious cycle of overload, stress, burnout and low morale ([5] p. 612) – We will execute …

3.3 Step 3 – Define Scope of Change – and What to Change

After the identification of the status of the change initiative has been performed and the directions set, scoping has to take place together with (top level) management. Detailed scope, budget, time and resource issues must be decided on by management.

It is furthermore important to identify the people to be involved in the change (change manager/team) or affected by the change (target audience), and which role they will play in the execution of the change initiative.

After this step has been completed detailed planning and implementation of the changes can take place at the operational level.

4 What Have Others Done

Change processes and change management have been researched in many settings, and there exists huge amounts of literature on strategies for change and recommendations on how to organize and conduct strategic changes. Among the most widely known and acknowledged are: Kotter focusing on an eight point plan for a change [10]; Hammer and Champy arguing for reengineering the whole corporation [20]; Mintzberg's overall conceptual frameworks for understanding and changing different organization structures [9]; and Mintzberg and co-authors furthermore suggested a set of "strategy schools" approaching the strategy formation process as: a conception, a negotiation, a transformation, or as being a formal, analytical, visionary, mental, emergent, collective or reactive process [18]. Other central contributions to the change literature are Senge's approach to change organizations through socializing and learning [19]; and Huy focusing on structuring and sequencing of strategies over time [5]. As mentioned previously these widely acknowledged scholars have – together with a number of other authors – been the key starting point for our research.

However, most of these influential works can be regarded as representing, one specific approach: Either the focus remains on the strategic level of the process, or the authors provide a series of detailed recommendations for how to structure and undertake the change processes at the operational level. The tactical level is more or less absent in most of the change literature.

Few authors provide recommendations and suggestions for how to establish a set of directions, policies and guidelines. A number of layered or step-based models for change activities exist, e.g. Kotter's famous eight steps [10] and Lewin's unfreeze-change-freeze model [21]. They are relevant and useful as overall models for change at a strategical level.

One of the few attempts in the literature to address the tactical level of change work are Balogun and her co-authors [22, 23] that provide a diagnostic three layer framework called the "change kaleidoscope" for identifying appropriate "design choices". Besides an organizational strategic change context referring to the broader strategic analysis they suggest eight essential features of the change context and six dimensions of choices for the change agent. Hereby they provide a framework for understanding and reflecting upon the change process and the change agent role. In terms of a strategic, a tactical and an operational level of the change initiatives, Balogun and co-authors have explicitly focused on the strategical level in their “organizational strategic change context”. The eight essential features and the six dimensions of choices focus on themes for the change context and for the role of the change agent, and both of these address the operational level rather than the tactical level of the change activities.

Also the guidelines in the original SPICE standard (ISO/IEC 15504-4) [24] do not mention a strategic or tactical level. It is mainly concerned with guidelines for the operational level. The strategic and tactical levels have later been introduced in the updated SPICE standard found in the ISO/IEC 33000 family of standards, i.e. in ISO/IEC 33014 [1].

The ISO/IEC 33014 standard in focus here contributes by defining fairly specific steps at the tactical level. But the three perspectives (see Fig. 1) all take their point of departure in process improvement based on assessments, and hence the steps focus on how to measure and assess the ability to improve.

Next to the academic literature lots of individual consultants, bloggers, and consultancy companies have provided their suggestions and recommendations. An example is strategy + business that provides ten principles for leading change management [25] focusing on culture, how to involve all layers in the organization, how to engage and lead etc. All ten principles are relevant and useful for many change settings, but they do mainly address our change-aspects of Attitude & Behavior; Culture; and Organizing. Throughout the literature are also mentioned aspects like: motivating people and sharing the vision. But there is still a need for also focusing on the other change-aspects that we have identified.

Many of the large consultancy companies have developed their own “standard processes” for how to prepare, plan and conduct changes in their customer organizations. A company like McKinsey presents their approach as a four step process of Aspire, Asses, Architect and Act [26]. Accenture describes their service on change management as to help their customers based upon “proven tools and methods” [27]. For most of these company-specific approaches the details and specific suggestions for standards and processes are kept as business secrets. Hence there is very little help and support for how set up directions and policies for the operational change work, i.e. how to support the tactical level of the change process.

In our studies of the change literature and from studying changes, change work and change processes in the “real world” clearly shows that a large number of changes have their origin from other sources, e.g. from changes in the market, changes in laws and regulations, or new technology or business models requiring new ways of working. Our proposed perspective at the tactical level of the standard should be seen as an expansion of the application for the ISO/IEC 33014 standard to deal with this broader set of changes.

We conclude from this overview of what others have done that it is our overall impression that most of the change literature presents strategies and approaches primarily at a strategic or operational level of recommendations. It focuses primarily on the overall approach to be employed in the planning and execution of the change.

Our approach has been to aim at suggesting how the nine change-aspects can be applied when setting out the directions for the change initiative, i.e. at the tactical level formulate policies and guidelines for the planning and implementation at the operational level.

The nine change-aspects we have established can be seen as overall headings for areas that must be taken into consideration when planning and conducting the actual change work (at the operational level), i.e. support the tactical level of change work.

5 Conclusion and Further Work

This paper has presented a proposal for expanding the tactical level in the ISO/IEC 33014 standard [1] to deal with a broader set of change initiatives. Focus has been on defining steps that can cover and support changes that do not have their main focus – or point of departure – from a process improvement perspective. We have briefly argued that this is the case for most change initiatives stemming e.g. from changes in the market, changes in laws and regulations and new technology.

Central to our suggested expansion to the tactical level is the concept of a change-charter for change initiatives, and the application of the nine change-aspects as a conceptual framework for structuring this change-charter and supporting the dialogue with top management when changes are initiated and directions, policies and guidelines for the operational level are established.

The nine change-aspects have been evaluated via interactions with a number of change managers and consultancy companies working with practical change planning and management in real work settings.

However, we still need further research on whether the change-aspects can be used by managers and change-teams in practice, and whether there are types of changes where they cannot be applied. This could lead to more detailed steps in the process.

Therefore much further testing and evaluation in real-life settings and in different types of change situations must be performed. It is our plan to do this through a number of close collaborations with different types of companies and organizations.

Part of our future research and evaluation work will also be to find out how the recommendations we have identified in literature can be provided to management and change-teams in a proper and useful way, and how the recommendations are best integrated with the steps at the tactical level of change initiatives.

To summarize, we are confident that our expansion to the tactical level dealing with other types of changes than those based on assessments is needed. We are also confident that the change-aspects we have found in literature are useful and applicable for assisting in developing change-charters for change initiatives in practice.

Acknowledgments. We wish to thank Professor Jan Pries-Heje (Roskilde University) and Partner Jørn Johansen (Whitebox) for thoughtful discussions and review of the change-aspects. We also wish to thank the change management master students at Roskilde University for their contribution in the validation process.

The two authors of this paper are listed in alphabetical order only. The paper is a joint effort.

References

1. ISO: ISO/IEC/TR 33014. Information technology–Process assessment–Guide for process improvement, Geneva, Switzerland (2013)
2. Carstensen, P.H., Vinter, O.: Aspects you should consider in your action plan when implementing an improvement strategy. In: Mas, A., Mesquida, A., O'Connor, R.V., Rout, T., Dorling, A. (eds.) SPICE 2017. CCIS, vol. 770, pp. 467–480. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-67383-7_34
3. Carstensen, P.H., Vinter, O.: Developing action plans based on strategy – aspects to consider. Softw. Qual. Prof. J. **20**(2), 4–15 (2018)
4. Carstensen, P.H., Vinter, O.: Eight aspects of actions in improvement plans. In: Larrucea, X., Santamaria, I., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2018. CCIS, vol. 896, pp. 147–158. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-97925-0_12
5. Huy, Q.N.: Time, temporal capability, and planned change. Acad. Manag. Rev. **26**(4), 601–623 (2001)
6. Ciborra, C.U., et al.: From Control to Drift. The Dynamics of Corporate Information Infrastructures. Oxford University Press, Oxford (2000)
7. Andersen, C.V., Krath, F., Krukow, L., Mathiasssen, L., Pries-Heje, J.: The grass root effort. In: Mathiassen et al. (eds.) Improving Software Organizations - From Principles to Practice, Addison-Wesley, Reading (2001)
8. Oakland, J.S.: TQM – Text with Cases, 3rd edn. Butterworth-Heinemann, Burlington (2003)
9. Mintzberg, H.: Structure in Fives - Designing Effective Organizations. Prentice-Hall, Englewood Cliffs (1983)
10. Kotter, J.P.: Leading Change. Harward Business Review Press, Boston (2012)
11. Stacey, R.: Strategic Management and Organizational Dynamics – The Challenge of Complexity, 6th edn. Pearson Educational Ltd., Essex (2011)
12. Liker, J.K.: The Toyota Way. Tata McGraw-Hill, New York (2004)
13. Kensing, F., Blomberg, J.: Participatory design: issues and concerns. Comput. Support. Coop. Work **7**(3–4), 167–185 (1998)
14. Grover, V., Jeong, S.R., Kettinger, W.J., Teng, J.T.C.: The implementation of business process reengineering. J. Manag. Inf. Syst. **12**(1), 109–144 (1995)
15. Pande, P.S., Holpp, L.: What is Six Sigma? McGraw Hill, New York (2000)
16. SPI Manifesto: Editors Jan Pries-Heje, Roskilde University, Jørn Johansen, DELTA Axiom (2010). www.iscn.com/Images/SPI_Manifesto_A.1.2.2010.pdf
17. Benner, M., Tushman, M.: Exploitation, exploration, and process management: the productivity dilemma revisited. Acad. Manag. Rev. **28**(2), 238–256 (2003)
18. Mintzberg, H., Ahlstrand, B., Lampel, J.: Strategy Safari: A Guided Tour Through the Wilds of Strategic Management. Financial Times, 2nd edn. Prentice-Hall, London (2009)
19. Senge, P.M.: The Fifth Discipline: The Art and Practice of the Learning Organization. Doubleday, New York (1990)
20. Hammer, M., Champy, J.: Reengineering the Corporation: A Manifesto For Business Revolution. Harper Business, New York (1993)
21. Lewin, K.: Frontiers in group dynamics: concept, method and reality in social science; social equilibria and social change. Hum. Relat. **1**, 5–41 (1947)
22. Balogun, J., Hailey, V.H., Gustafsson, S.: Exploring Strategic Change. Pearson, Harlow (2015)
23. Balogun, J., Hailey, V.H.: Devising context sensitive approaches to change: the example of Glaxo welcome. Long Range Plann. J. **35**, 153–178 (2002)

24. ISO: ISO/IEC/TR 15504-4. Information technology–Process assessment–Guidance on use for process improvement and process capability determination, Geneva, Switzerland (2013)
25. www.strategy-business.com
26. www.youtube.com/watch?v=k69i_yAhEcQ
27. www.accenture.com

Best Practices for Software Maturity Improvement: A GÉANT Case Study

Bartosz Walter[1(✉)], Branko Marović[2], Ivan Garnizov[3], Marcin Wolski[1], and Andrijana Todosijevic[4]

[1] PSNC, Poznań, Poland
{bartek.walter,marcin.wolski}@man.poznan.pl
[2] University of Belgrade, Belgrade, Serbia
branko.marovic@rcub.bg.ac.rs
[3] Friedrich-Alexander-University of Erlangen-Nürnberg, Erlangen, Germany
ivan.garnizov@fau.de
[4] AMRES, Belgrade, Serbia
andrijana.todosijevic@amres.ac.rs

Abstract. Maturity models for software indicate the key areas that contribute to quality improvements. They usually combine technical, organisational and human aspects relevant for effective software development, to focus the efforts and draw the direction for optimisations. In this paper, we present the process of defining best practices that support the GÉANT Software Maturity Model (GSMM), aligned to the needs of a distributed, innovation-driven, pan-European organisation. Based on the identification of specific goals relevant for GÉANT and a preliminary maturity assessment, we created a catalogue of best practices that help the software teams to attain the goals defined in the GSMM.

Keywords: Maturity evaluation · Best practices · Software process improvement · SPI

1 Introduction

Managing software process improvement endeavours is a complex challenge that usually involves the effort of several teams and individuals. In particular, it is the case for large organisations focused on innovation, with an established culture of diversity and openness [14]. The process improvement usually requires identification of factors relevant in a given context, setting attainable objectives,

This work is part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 856726 (GN4-3).
The scientific/academic work is financed from financial resources for science in the years 2019–2022 granted for the realisation of the international project co-financed by the Polish Ministry of Science and Higher Education.

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 30–41, 2020.
https://doi.org/10.1007/978-3-030-56441-4_3

defining metrics for tracking the progress, but also coordinating the efforts in various areas: technical, human and organisational.

The concept of maturity, which captures the capability of an organisation to deliver high-quality products, is widely accepted as an effective method of improving the processes. Maturity models provide frameworks capturing the essential dimensions of quality that are relevant in a given context, to set objectives and propose methods of addressing them. Several models have been proposed for software development, both generic [6,12] and tailored [9].

However, defining the model is only one side of the coin. Apart from identifying the goals and defining the metrics, the subject teams and individuals also need guidance, support and actionable recommendations on how to work to attain the objectives. Although it is quite feasible to directly implement the model in uniform and hierarchical organisations, for internally diversified and independent structures it could be a difficult task.

This also is the case of GÉANT, a pan-European organisation, established and funded under several EU programmes for the development and operation of a fast, reliable networking environment for research and education, which offers software-based services to various end-users, including students and researchers. It involves many independent software teams that are free to define their processes and obliged only to adhere to the common organisation-wide recommendations.

In this paper, we report on how a custom maturity model for GÉANT could be supported by a catalogue of best practices. The catalogue guides on how the specific objectives of the model could be addressed and implemented by the GÉANT software teams.

The paper consists of seven sections. In Sect. 2 we report a literature overview; in Sect. 3 we shortly introduce the GÉANT organisation and its specifics concerning software development. In Sect. 4 we present the entire process of maturity improvement, from defining the model to constructing the catalogue of best practices. Next, in Sect. 5 we present how the best practices are described, formatted and presented in a catalogue. Section 6 reports the early results of the evaluation, and the Sect. 7 provides concluding remarks and the summary.

2 Related Work

Maturity of software organisations is a topic widely explored in literature. Several software-related maturity models have been developed that refer to specific areas and scopes, e.g. software process capability models (CMMI) [12], software analysis [8], operational management [13] or business process management [5]. Although maturity is usually related to traditional methods of software development, the concept of maturity has been also tailored for agile approaches: Agile Maturity Model (AMM) [7] defines levels of agility, which address the common agile practices and values starting from basic ones, e.g., planning and requirements management, up to managing uncertainty and defect prevention.

A number of maturity models have been also proposed in EU-funded projects. They address various areas that could be partially relevant in a software-related context, e.g., communication [11] or selected education techniques[1]. However, they usually focus on a single, selected dimension of the project.

As a consequence, although numerous models exist, they still need to be merged, customized or redefined to reflect the specific requirements and settings and to embrace all areas relevant for software development. To respond to this, in previous papers [14] we presented a preliminary version of the GSMM, a maturity model dedicated for the GÉANT organisation, along with recommendations on how to define models and implement them [15].

Effective implementation of the objectives and goals defined in maturity models requires also adequate guidance and recommendations. They could take the form of best practices that are well-founded on both the experience and the existing knowledge, and are applicable in the relevant context. This approach is widely adopted in software engineering, e.g., in SWEBOK [3]. Catalogues of such practices dedicated to specific areas have been proposed by various authors. Gamma et al. [10] expressed the collected experience in designing object-oriented software as a set of design patterns. They documented key design templates and presented them in the form of a catalogue of abstract structured recommendations. Also, anti-patterns have been defined, capturing practices that should be avoided [4]. Similar efforts have been undertaken also in several other software-related areas, e.g., testing, documentation etc.

Ambler [1], based on his observations, emphasized the importance of the context in the analysis of best practices. He argued that most practices are not applicable in all cases, and they needed to be either adapted before being applied or to be implemented only in specific environments.

We believe there is still a need for presenting custom, organisation-specific models for improving software maturity, supported by structured, practice-originated experience.

3 Background

GÉANT is a pan-European project focused on the development and maintenance of e-infrastructure and services for the research and education community. It operates the backbone network and associated services interconnecting national research and education networks (NRENs) across Europe and enables their collaboration. Also, it is a distributed, innovation-oriented organisation involving participants from many countries and organisations that develop and maintain network-based products and services, frequently based on dedicated or customized software. GÉANT portfolio comprises currently 30 software projects: some are used directly by GÉANT; some more are shared or used by NRENs; yet, others contribute to wider open-source communities.

Members of the software teams have specific working arrangements: they simultaneously work for GÉANT and their native organisations, can be simultaneously involved in several projects, are geographically distributed, and are

[1] https://embed.eadtu.eu/.

placed in different cultural and professional backgrounds. The teams share the common software development framework provided by GÉANT, but are allowed to choose and customize their processes, methodologies and approaches. Their developments are focused on innovative and often prototypical applications for the high-performance network, based on the novel and often federated services.

A previous analysis of the GÉANT software development practices [16] showed that there is a need within the GÉANT software development community for optimisation of the software development processes. This need could be addressed by providing the software teams with guidance on adopting and using software development methodologies and practices effectively and efficiently.

The motivation for the establishment of a software maturity model for GÉANT was to determine properly the improvements schema for software teams. Moreover, the practical appliance of maturity model would align the improvement effort with governance frameworks, commonly approved models, and with industry practices. The apparent attractiveness and popularity of the maturity model in the management of the software process and, more broadly, IT management, has contributed to our effort to develop a maturity model specifically for software process improvement (SPI) within GÉANT, while respecting the seven suggested requirements for the development of maturity models [2].

The GÉANT Software Maturity Model (GSMM) has been designed to achieve two primary goals: (1) to capture key practices that already help the teams to successfully deliver software, and (2) to identify areas for further improvements that could be applied by the teams [14,15].

The extracted practices can then be shared, adapted and applied by the teams to streamline and align software development processes and governance. The leading factors considered include the lasting nature of GÉANT, its products and services, distributed nature of conducted collaborations and the existence of many practices that have been established for some time.

The resulting model consists of categorised software engineering topics and processes into five key thematic areas, referred to as target areas (TA), namely: requirements engineering; design and implementation; software maintenance; quality assurance; and team organisation. These target areas were elaborated by providing specific content to the maturity model, with each one consisting of several specific goals (SGs) that capture sub-objectives and related activities.

4 Process of Defining the Best Practices

Best practices are commonly accepted procedures that aim at accomplishing certain objectives. They are applicable in a given context. They are considered as the gold standard for attaining specific objectives.

The process of defining the best practices comprised the four main steps (see Fig. 1):

1. Defining the software maturity model which would determine our improvements schema;

2. Preparing the questionnaire which would be consistent with the structure of the software maturity model;
3. Interviewing selected GÉANT software teams to collect both qualitative and quantitative information and opinions on how the specific goals are addressed;
4. Determining the set of Common Best Practices (CBPs), based on the results of the survey, literature review and own observations made by the software management team.

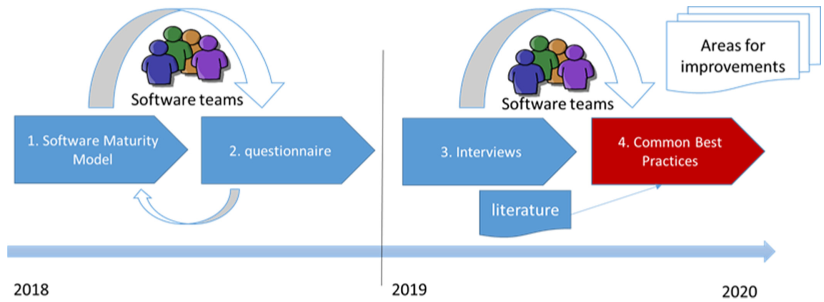


Fig. 1. The process of defining best practices adopted for GSMM

Additionally, the interviews highlighted areas for potential improvements in teams.

The first two steps (defining the software maturity model and preparing the questionnaire) have been accomplished within a few iterations: the model, its underlying conceptual framework, as well as the questionnaire, were iteratively developed in cooperation with several software teams. The other two steps (3–4) were accomplished sequentially.

4.1 Defining the Maturity Model and Questionnaire

Maturity models are widely applied managerial instruments used for the evaluation and improvement of organisational practices and processes. The GSMM focuses on software development processes within GÉANT, considering the particular constraints in which the software teams operate. As a result, the GSMM identifies 29 Specific Goals (SGs) grouped into five Target Areas (TAs), which are essential for effective software development in GÉANT. The goals indicate objectives that need to be addressed by the software teams in the technical, organisational and human domain.

The process of defining and implementing the GSMM, its elements, produced outputs and related actions and enhancements are presented in Fig. 2. The process of identification of specific elements of the GSMM and their further refinements is conducted iteratively, based on external and internal sources of the domain knowledge. In particular, a number of pilot interviews with the teams helped to identify key areas that are relevant in GÉANT, and further to decompose them into individual objectives. What is important, the identification process was not limited to the activities directly related to software development. We were also interested in capturing organisational, communication and human perspectives that are relevant for software teams.

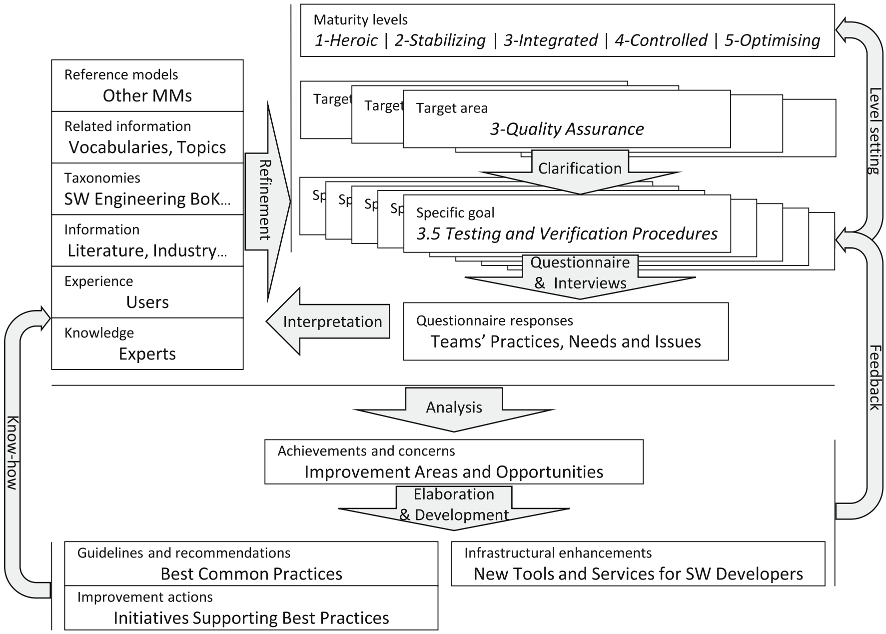


Fig. 2. The overview of GÉANT software maturity model development

This model has been elaborated in close cooperation with selected software teams, by applying iterative refinements, improvements and collecting feedback. Specifically, the process included the following phases:

1. **Defining a preliminary version of GSMM**. Based on the observation of software teams, pilot interviews with the teams and a literature review, an initial version of GSMM was drafted.
2. **Validating the preliminary version**. The preliminary version of GSMM was presented to representatives of selected teams and the executive managers responsible for supporting software development in GÉANT, to collect early feedback that could be immediately addressed.
3. **Revising the GSMM**. Based on the collected feedback, the core elements of GSMM have been revised to better align it with the needs and practice of the teams.
4. **Formalising the GSMM**. In parallel to that, the GSMM has been formalised into a framework by defining its elements. As a result, it can also be adapted to other application domains, beyond software development.

4.2 Data Collection for the Common Best Practices

To identify and capture the practices applied by the SWD teams, as well as the teams' expectations concerning maturity, we created a questionnaire for GSMM and used as a blueprint. The questionnaire covers us meant to extract comments from the teams on each GSMM-specific goal by asking a number of closed, open and rating questions. The questions were formulated to elicit useful explanations on whether some aspects are covered or significant, how the evaluated team addresses each specific goal, what resources and approaches are used, what outputs are produced, and in which areas support may be needed. In total, we interviewed 12 teams, a representative subset of the active software teams with varied sizes.

This resulted in an extensive but targeted and streamlined process of data collection, in which the participants were able to effectively and purposefully capture their teams' software development and management practices, but during which it was also possible to capture the teams' attitude and perception of the specific topics.

5 Common Best Practices

The GSMM, presented in Sect. 4, identifies the core areas and goals that contribute to the effective software development in GÉANT. However, software teams need not only the targets but also the guidance on how to organise the efforts towards meeting these them, considering the specific context of the organisation. To address this, we created a catalogue of Common Best Practices (CBPs). They have been identified based on three sources of information:

– practices currently applied by the software development teams, extracted during the questionnaire interviews (see Sect. 4),
– recommendations provided in the literature and case studies concerning similar process optimisation efforts,
– direct or indirect observations made by the software management team.

The catalogue includes recommendations showing how the particular GSMM specific goals could be addressed, considering the constraints and opportunities present in GÉANT. Therefore, the best practices balance between the need for providing detailed operational guidance and giving a general direction. Both of these are extremes that would result in missing the objective of providing the teams with effective guidance that they could adapt and implement in their own practice.

5.1 Template of a Best Practice

Description of a best practice includes a large volume of diversified information of various nature and format, which can hinder its readability. To make the practices more accessible for the team members and, consequently, facilitate the adoption of practices, we decided to define a template which presents the data in a structured way. Each practice is presented as a set of attributes that describe key elements of the practice. The template comprises the following attributes:

- **Objective** that the practice is expected to address;
- **Applicability**, describing types of projects or their phases, in which the practice could be effectively applied;
- **Context** that captures a specific setting (a set of technical, managerial or organisational constraints), for which the practice was identified and for which it is recommended; The practice can be used in other settings, but it may require additional validation;
- **Addressed elements in GSMM**, linking the practice with the SGs in the GSMM;
- **Prerequisites**, listing the condition necessary to apply the specific practice;
- **Recommendation**, outlining actions that should be undertaken to meet the SG; It includes high-level directional advice on *what* to do, with lower-level details on *how* to reach the goal.
- **Risks**, describing possible risk factors and their consequences in case of misusing the practice;
- **Related practices**, indicating other practices that could be applied in a similar context;
- **Origin**, providing details on how the practice originated.

5.2 Example

As an example, we present a practice related to managing stakeholders that belongs to the Requirements Engineering target area. It deals with the problem of identification of relevant organisations, teams and individuals, who would affect the project or are interested in its outcomes, and properly addressing their needs and expectations.

- **Objective**: Identify relevant stakeholders that can contribute to the project or have an impact on it
- **Context**: The practice applies to all projects.
- **Addressed elements in GSMM**: RE-1. Identification and overall management of stakeholders
- **Prerequisites**: none
- **Recommendation(s)**
  1. Identify an initial group of stakeholders
     (a) Consider teams, NRENs or individuals that could be affected or could impact the project.
     (b) Look for similarities to other projects, either previous or current.
     (c) Look for a dominant stakeholder, who is mostly interested in the outcome of the project
  2. Maintain (update) the group of stakeholders
     (a) Publish the list of stakeholders and their representatives

  (b) Periodically update (involve and retire) the group of stakeholders
  (c) Apply snowballing to identify new stakeholders
  (d) Categorise the stakeholders according to their relevance for the project
- **Risks**:
  1. The identified group does not include all relevant stakeholders
     (a) The project may be subject to tensions, sudden changes or drifting.
     (b) The decisions could be made/affected by people not officially involved in the project.
     (c) The project would be not driven by stakeholders, but rather by the project team.
  2. Group of stakeholders is not properly updated
     (a) The group may not reflect the actual balance of interests.
- **Related practices**: BP-A-2: Create a strategy to communicate with stakeholders.
- **Origin**: This practice has been defined based on the survey, supported by the observation by the software management team.

The recommendations do not provide direct instructions for the teams on how to proceed in a step-by-step manner, but rather give directional guidance. It includes advice concerning the factors and issues that facilitate addressing the respective goal in the GSMM and that should be considered by the team. As a result, the recommendations presented in a practice are a trade-off between the desire to deliver actionable procedures on one hand and the necessary abstractness on the other. The software teams are expected to analyse and adapt the recommendations to their local context.

The risk factors capture possible consequences of applying the practice improperly. In this case, they are mostly related to issues of stakeholders identification and management, including prioritisation and identification of relationships among them.

This specific practice is closely related to the process of defining the communication policy and maintaining contact with stakeholders. These two practices should be used together to ensure the maximum benefit from their implementation.

The entire catalogue is available for the GÉANT staff and currently includes 24 practices divided into five target areas that directly correspond to areas in the GSMM. Each practice addresses one or more SGs defined in the GSMM and, currently, all the goals are covered.

6 How to Identify Key Areas for Improvements?

One of the goals for analysis was to identify the areas that deserve most attention and effort. The data collected so helped in the selection of topics that should be promoted, worked on and supported through the improvement incentives. Below, we provide comments on them:

- First, we focused on the areas in which the declared knowledge and satisfaction of the software teams was the most diversified. For those items, we can rely on the immediately available internal expertise of some teams, which reduces the effort needed for process improvement. Such harmonisation of processes among teams could additionally strengthen their collaboration.
- Another issue refers to areas for which the survey scores were generally low or mediocre. These can be interpreted two-fold: either (1) the survey has identified a relevant topic that has been underestimated and has not been covered adequately, or (2) the topic is considered irrelevant by developers. To determine the actual status of these topics, teams and their leaders should be consulted. For that reason, it is useful at this point to establish a regular mechanism for collecting feedback from software teams. The apparent difficulty associated with these topics is the insufficient internal experience, so it may be necessary to look for external expertise and support to achieve the expected improvements.
- The third group of interest includes the areas with uniformly high marks. Here, we can expect relatively small improvements, even if the collected data is not completely accurate. High grades given by all teams may indicate that the reached level cannot or should not be further improved.

The instruments established to collect feedback from software teams should not be used just to address the dilemmas related to the selection of improvement areas. They could be also applied to get a response about the ongoing improvement initiatives and track the metrics that are relevant for the implementation of the GSMM. However, conducting extensive interview-based surveys, like the one described in Sect. 4, requires significant effort; at next stages, they could be supported with simpler and frequently run online questionnaires that focus on the key elements.

Optionally, this approach could be taken even further by associating the indicators with maturity levels. Maturity is assessed against key capability and practice characteristics linked with the selected target areas at each level of the model. Maturity levels are typically established using a five-point Likert scale where the higher the level, the higher is the level of maturity. These progressive levels guide the planning and development of roadmaps. Currently, there is no need for GÉANT to develop such a far-reaching maturity model. A future move in this direction would require reaching beyond GÉANT into an even larger base of software teams to capture the indicators and link their values to maturity levels. Given the number of potential indicators and certain subjectivity in associating them with maturity levels, it is necessary to draw from the opinions and attitudes of a larger community to get non-disputable criteria. Although other related standardisation processes and schemes and maturity models could additionally support this development, it is crucial to establish and employ a simple, inclusive and quality process in which the indicators and levels are produced and challenged. Again, a series of simple and low-effort online surveys could be used to iteratively establish and maintain the set of used indicators and maturity criteria.

7 Summary and Conclusions

The presented approach aims to establish an adaptive and practical framework for improving the quality and maturity of practices within GÉANT. It is primarily aimed at improving coordination of software process improvement efforts, fostering the collaboration of software teams, and supporting the entire software development life cycle with best practices.

In particular, the catalogue of best practices does not only help the teams to address and implement the objectives set in GSMM but also allows them to adapt the recommended activities to their contextual constraints. As such, it fits well into the SPI Manifesto[2] that emphasizes the need for embedding the maturity improvement efforts in practice and the actual needs of software teams. We believe that the catalogue of best practices will become a live, ever-growing toolbox of recommendations that would provide substantial guidance for the teams, but also receive updates from them.

References

1. Ambler, S.: Questioning best practices for software development: practices are contextual, Never Best (2011). http://www.ambysoft.com/essays/bestPractices.html. Accessed Apr 2020
2. Becker, J., Knackstedt, R., Pöppelbuß, J.: Developing maturity models for IT management. Bus. Inf. Syst. Eng. **1**(3), 213–222 (2009). https://doi.org/10.1007/s12599-009-0044-5
3. Bourque, P., Fairley, R.E.: Guide to the Software Engineering Body of Knowledge (SWEBOK(R)): Version 3.0, 3rd edn. IEEE Computer Society Press, Washington, DC (2014)
4. Brown, W.J., Malveau, R.C., McCormick, H.W.S., Mowbray, T.J.: AntiPatterns: Refactoring Software, Architectures, and Projects in Crisis: Refactoring Software, Architecture and Projects in Crisis, 1st edn. Wiley, New York (1998)
5. de Bruin, T., Rosemann, M.: Using the Delphi technique to identify BPM capability areas. In: ACIS 2007 Proceedings - 18th Australasian Conference on Information Systems, January 2007
6. Burnstein, I., Suwanassart, T., Carlson, R.: Developing a testing maturity model for software test process evaluation and improvement. In: Proceedings International Test Conference 1996. Test and Design Validity, pp. 581–589, October 1996. https://doi.org/10.1109/TEST.1996.557106
7. Chetankumar, P., Ramachandran, M.: Agile maturity model (AMM): a software process improvement framework for agile software development practices. Int. J. Softw. Eng. **2**, 3–28 (2009)
8. Covey, R., Hixon, D.: The creation and use of an analysis capability maturity model (ACMM). Technical report, Aerospace Corp. El Segundo CA Lab Operations (2005)
9. Fontana, R., Meyer Jr., V., Reinehr, S., Malucelli, A.: Progressive outcomes: a framework for maturing in agile software development. J. Syst. Softw. **102**, 88–108 (2015). https://doi.org/10.1016/j.jss.2014.12.032

[2] https://2019.eurospi.net/index.php/manifesto.

10. Gamma, E., Helm, R., Johnson, R., Vlissides, J.M.: Design Patterns: Elements of Reusable Object-Oriented Software, 1st edn. Addison-Wesley Professional, Reading (1994)
11. Muszynska, K.: Communication maturity model for organizations realizing EU projects. Procedia Comput. Sci. **126**, 2184–2193 (2018). https://doi.org/10.1016/j.procs.2018.07.230
12. Paulk, M., Curtis, B., Chrissis, M., Weber, C.V.: Capability maturity model for software, version 1.1, January 1993
13. Renken, J.: Developing an IS/ICT management capability maturity framework. In: Proceedings of the 2004 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists on IT Research in Developing Countries, pp. 53–62. South African Institute for Computer Scientists and Information Technologists (2004)
14. Stanisavljevic, Z., Walter, B., Vukasovic, M., Todosijevic, A., Labedzki, M., Wolski, M.: GÉANT software maturity model. In: 2018 26th Telecommunications Forum (TELFOR), pp. 420–425, November 2018
15. Walter, B., Wolski, M., Stanisavljević, Ž., Todosijević, A.: Designing a maturity model for a distributed software organization. An experience report. In: Walker, A., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 123–135. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_10
16. Wolski, M., et al.: Deliverable D5.3 - analysis of requirements for software management. Technical report (2017). https://www.geant.org/Projects/GEANT_Project_GN4/deliverables/D5-3_Analysis-of-Requirements-for-Software-Management.pdf

SPI and Emerging Software and Systems Engineering Paradigms

Visualization, Monitoring and Control Techniques for Use in Scrum Software Development: An Analytic Hierarchy Process Approach

Nesib Tekin[1], Mehmet Kosa[2], Murat Yilmaz[3,4(✉)], Paul Clarke[3,4], and Vahid Garousi[5]

[1] Turkish Defense Industries Research and Development Institute – TUBİTAK SAGE, Ankara, Turkey
nesib.tekin@tubitak.gov.tr
[2] Cognitive Science & Artificial Intelligence (CS&AI) Department, Tilburg University, Tilburg, Netherlands
m.kosa@tilburguniversity.edu
[3] Lero, The Science Foundation Ireland Research Centre for Software, Limerick, Ireland
[4] Dublin City University, Dublin, Ireland
{murat.yilmaz,Paul.M.Clarke}@dcu.ie
[5] Queen's University Belfast, Belfast, UK
V.Garousi@qub.ac.uk

Abstract. Scrum is the most widely used agile development framework that guides the development process with its ability to create customer-valued software artifacts iteratively and incrementally, whilst seeking best practices to provide continuous measurement during production. However, measuring success in Scrum can be a challenging endeavor. In particular, it is hard to select the best fitting agile metrics during consecutive Scrum sprints. The goal of this industrial case study was to utilize a systematic selection process for identifying the appropriate scrum metrics tools addon component within the TÜBİTAK SAGE software development group. Moreover, the distribution of software developers' preferences of process metrics were analyzed according to their characteristic features and defense industry structure, and are presented using various distribution charts. Finally, alternatives to the software development process measurement component, which was integrated into the agile software process tool employed by the TÜBİTAK SAGE software development group, were efficiently determined by using the Analytic Hierarchy Process approach. Among the options discussed, our results suggest that the Actionable Agile Addon scored the highest followed by the Screenful Addon. The present study presents a rigorous approach that ultimately have improved community participation in metric planning, implementation and monitoring, thus moving towards sustainable software development goals.

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 45–57, 2020.
https://doi.org/10.1007/978-3-030-56441-4_4

Keywords: Software measurement component · Software process metrics tool · Scrum · AHP · Software component selection · Industrial case study

1 Introduction

The notion of quality has become crucial across all engineering disciplines, and can be assessed through various forms of measurement, which is a crucial element to all scientific and engineering activity. Therefore, software engineering is no exception as it is essentially an engineering discipline. According to Pfleeger [1], software measurement is an inseparable part of software development and maintenance. Throughout the software development lifecycle, measurement processes should be used effectively in order to evaluate the quality, as well as the improvement [2] and performance of the product [3]. Today, performance measurement has become a key feature in the development of successful software engineering applications [4], and productivity [5].

The SPI manifesto suggests creating a learning organization [6]. This promotes the fact that metrics within the software development process should provide valuable information about the product under development in such organizations. One of the essential goals of the software metric is the aim to eliminate human-factor uncertainties within the software measurement process [7]. Since the collected metric data reflects the problem in the development process, a company can use such data to formulate regulatory actions and to improve the software development process [8]. There are many different software metrics (e.g., lines of code, code complexity, cycle time and velocity, mean time to recover) within the software development domain. In the literature, different approaches have been applied in classifying these metrics. Lee and Chang's [3] classification related to the metrics necessary for software measurement based on software quality. They divided software quality metrics into five groups: product quality; in-process quality; testing quality; maintenance quality; and, customer satisfaction quality. Another grouping can be expressed as: commercial perspective; significance perspective; observation perspective; measurement perspective; and, software development perspective. Commercial perspective includes technical metrics, defect metrics, end-user satisfaction metrics, warranty metrics, and reputation metrics. Significance perspective includes core and non-core metrics. Observation perspective contains primitive and computed metrics. Measurement perspective involves direct and indirect metrics [9]. Software development perspective includes process metrics, product metrics, test metrics, maintenance metrics, and subjective metrics [4]. However, Lee and Chang [3] pointed out that a metric can be included in one or more of these categories. Scrum is one such modern agile software development framework that is widely used and known. The most important change in the Scrum approach is the transfer of responsibilities and decision mechanisms that previously sat with administrative staff to the actual software development teams [10]. Therefore, an important purpose of Scrum metrics is to help the software team and their managers to monitor the business development process, as well as business quality, productivity, predictability, the health of the product and the team [11].

Organizations generally choose software-based development process tools in order to more easily control comprehensive system developments. Various software development process tools are widely used by companies to manage the Scrum process. In fact, tools have been developed for specific market areas that contain several add-ons. Various developers have created adaptive applications for these software development process tools, and companies can purchase and integrate these applications into their own software development process tools as an add-on. The variety of software process metrics increases daily. Commonly used process metrics have mostly been integrated into process tools as default. However, companies can also expand the scope of these process metrics by purchasing add-ons based on their specific business needs. The most commonly used Scrum problem and project tracking software tool contains a total of nine software process metrics by default. However, there are also 66 metrics add-ons available on the market. Their content, functionality, technical specifications, websites, usage training, and demonstrations are also available in the market.

The aim of the current study is to select the software development process measurement component which includes the appropriate software development process metrics by using the AHP method. Also, software project requirements and software developers' contributions are also considered in the selection process. The remaining part of the paper proceeds as follows: The next section of this paper reviews the literature. Third section elaborates on the research methodology. The fourth section discusses and summarizes the results. Finally, conclusion section concludes the study.

2 Background

Agile development practices enable self-development and ultimately embrace the human experience [12]. Scrum methodology consists of iterative and incremental sprint structures, and software development roles [13] with the target determined prior to a sprint being started. Scrum metrics indicate whether or not the target of the sprint can be accomplished. The most fundamental example is where the result of a sprint presents new functionality of a product. The Scrum team can monitor quantitative evaluation of the work, success rates of the sprint, and the maturity level of the team by using Scrum metrics. Sprint's success rate is an important starting point for adaptation and inspection [14].

2.1 Importance of Choosing the Right Metrics

Pfleeger and Fitzgerald [15] conducted a detailed study about selecting the right software metrics toolkit. They underlined that requirements, process, and maintainability are important factors which should be considered during the software metric selection process. Additionally, they defined that customization and coordination of the toolkit can be time-consuming activities. Card and Glass [16] stated that data collection and analysis in the calculation of metrics increased the project cost by 7–8% in a software engineering laboratory at the University of Maryland. While the cost of metrics in the project and data collection are important, accurate data collection is vital in order to generate accurate metrics. Spending both in terms of time and cost for unnecessary metrics should be avoided.

Research on software measurements continues to be a hot topic today. Although the benefits of using of software measurement are well known, problems in their practice are still ongoing. More than 80% of software measurement attempts fail within the first 18 months [17]. The most common explanation for this is the difficulty in understanding and using metrics. According to Fenton and Neil [18], metrics are not used effectively within the decision-making process, because they do not have reliability in terms of validity.

2.2 Importance of Software Process Metrics

Kitchenham et al. [19] mentioned that more sensitive and considerate behavior is required in the measurement process. Ptleeger [1] emphasized a very strong motto that better decisions can be made on various subjects by measuring the past and changing the future. This situation requires the recruitment, training and building of a team based on software quality. In particular, to improve coordination and collaboration between software development teams, Huawei Turkey Research and Development Center proposed a version Hoshin Matrix and mapped software metrics and organizational goals [20].

Measurement and competition are factors that encourage people to improve on their previous performance in a situational context [21]. Software measurements help to understand how a project performs according to its goals, and provides information about the state of an organization compared to some previous period [22]. Pfleeger and Fitzgerald [15] explained that a software metrics toolkit should be determined related to the needs of the software being developed. Also, when choosing metrics for a software project development process, three concerns should be taken into consideration: (i) maturity of the software development process; (ii) availability of measurement data; and, (iii) the project management requirements. However, several drawbacks were revealed when the selection process for a metrics toolkit were determined by only the project managers. These disadvantages were that metric customization is both time-consuming and difficult in terms of coordinating the usage of the metrics. Paulish and Cartelon [8] suggested that an evaluation process should be realized prior to starting a software process improvement program. The researchers indicated that the evaluation can be a powerful method that addresses priorities and consensus within an organization that focuses on improvement. According to Ebert et al. [22], metric creators should not act independently from metrics users. They highlighted that users of metrics should comprehend the fundamental information about the measurement process of the software.

2.3 Analytic Hierarchy Process (AHP) Framework

The Analytic Hierarchy Process (AHP) approach, which is one of the multi-criteria decision-making techniques, integrates different types of criteria into a hierarchical structure, enabling the evaluation of each alternative [23]. There have been a limited number of studies using multi-criteria decision-making methodologies such as AHP related to the selection of software metrics and their tools. Sharma et al. [24] used the AHP method, which is one of the multi-criteria decision-making method, for selecting software related to the project management process. Sagar et al. [25] proposed the fuzzy AHP method for selecting the most reusable software component such as plug-ins. They determined criteria that pointed to reusability as adaptability, availability, interface complexity, customizability, and understandability by reviewing the related literature. Ömürbek et al. [26], the aim was to select a project management program which could be used in software development based on using AHP and the TOPSIS methods.

In a study by Al-Qutaish et al. [27], the AHP method was used to manage the selection process of open-source software according to the ISO 9126-1 standard related to a set of six quality characteristics: functionality; reliability (R); usability; efficiency; maintainability; and, portability. Zaidan et al. [28] presented open-source electronic medical record software packages selection approach based on the AHP and TOPSIS methods. The results of the study showed that GNUmed and OpenEMR software had the most high-ranking scores when compared to other open-source EMR software packages.

All in all, the software process metrics in improving software quality is important to understand, as well as the cost of applying unnecessary metrics. Moreover, end-users themselves should be included in the metric selection, efficiency improvement and related activities [29]. However, software development process metrics are presented as a generic structure by the software process tracking tool. Also, there are several metrics add-ons on offer for the Scrum process metrics as sub-group kit products. It is a difficult and complex issue deciding which add-ons to use because the process is known to be affected by multiple criteria such as the type of software being developed, the development process model, the experience of the software developers, the domain of the project, and its duration. Therefore, a systematic selection process is needed as a solution to this multi-criteria problem.

In the current study, the selection of add-ons which include the software process metrics was performed using the AHP method. The purpose of the current study is to examine the selection of components that offer software development process metrics based on the AHP method. In this context, the weighting of selection criteria and the evaluation of alternatives were conducted in order to select add-ons that include software process metrics for an industrial case study in the TÜBİTAK SAGE software development group.

3 Research Method

First, the issue of appropriate software development process measurement component selection was determined. Then, a survey was applied to the study group. After that, the AHP method was implemented with the contribution of the expert group. The priorities of both criteria and alternatives were then evaluated. Finally, the highest priority alternative was selected as the final decision. Figure 1 illustrates each step of the case study process.

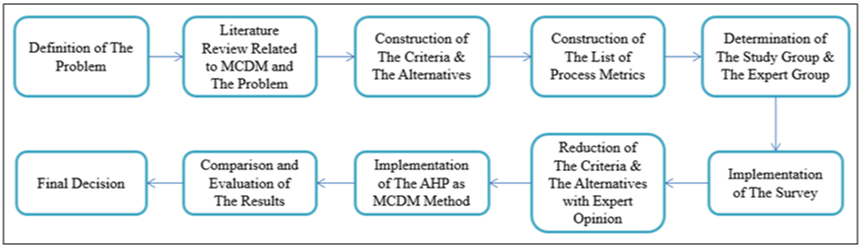


Fig. 1. Case study process steps

A list of software process metrics that are deemed appropriate to the Scrum software development process was constructed based on:

– The tool of the Scrum issue and the project tracking software, which has nine software process metrics by default.
– Various Scrum process metrics that can be integrated into the tool as an add-on to the Scrum issue and project tracking software.
– How users can customize the Scrum process metrics with the help of any available add-ons.

Study group: The study group consists of project software developers who take responsibility as developers of the software during the development process. This group was selected because they understand the corporate culture, the software development process, are aware of the things they are doing well or any problems, and have significant experience in the software development domain. Expert group: The expert group consists of employees who have at least 5 years' experience in software development. Group members have titles such as Team Leader, Unit Manager, Coordinator, and/or Chief Scientist. Also, this group has the primary decision-making responsibility for the software development process.

A software development process meeting was organized in the unit where the study was conducted. The participants were informed about the process metrics survey. Both the study group and the expert group participated in the meeting, and were asked to review the list of software development process metrics and to report their preferences through completion of the survey. Customized process metrics could be suggested by the participants when completing the survey as indicated at the first meeting with the study group. In addition, the participants were informed that demographic data about their years of work experience,

experience domains, graduation departments, and their working units would be collected within the survey.

The survey was prepared using Google Forms as a template to create the survey. Additionally, data collected via Google Forms could later be converted into graphical format. The survey schedule was announced to the participants 7 days in advance of its application. The results were grouped and associated with the demographic data collected from the participants. The findings obtained from the survey were then presented graphically to the expert group. Reduction of the criteria and the alternatives was conducted by the expert group according to the results of the survey.

In determining the software development process metric add-ons, consideration was given to the ability that add-ons could be integrated to the Scrum software development process tool used by the company at which this industrial case study was applied. The market for the Scrum issue and project tracking software was then examined, and a list of 66 products was identified as potential add-on applications. Each product was then examined according to the following:

– Technical specifications and capabilities (types and number of process metrics);
– Usage rate in the market;
– Scoring rate in the market.

The alternative add-on products were compared and their functionalities considered based on requirements suggested for software developers' process metrics. Each alternative product was then examined according to the following:

– Technical specifications that meet the requirements;
– Accessibility and availability;
– Includes visualization and/or report mechanisms.

The literature included several criteria for metrics tools. It was observed that software process metrics tools aim to increase market sales by highlighting these criteria, and that important criteria were emphasized within the technical documents of each metric tool (see Table 1).

While evaluating the software process metrics, robustness, simplicity and cost-effectiveness were seen as the key parameters. This is because the processes are defined according to systematic and powerful procedures. Each process was found to have been described quite simply. Finally, it was emphasized that the processes should be cost-effective. The following criterion group was then used in evaluating the tools offered for the software development process metrics.

The expert group performed the reduction process by considering the results of the survey. After that, the criteria and alternatives list in the AHP method were finalized. Reliability of the reduction operation was assessed by using Cronbach's alpha method. It is represented by a numerical value between 0 and 1 in terms of a reliability scale.

Table 1. Software process metrics add-ons evaluation criteria

Criteria	Meaning
Relevance	(to collect process metrics) Reflects the relationship between process metrics and the metric tool
Experience	Degree to which the metric tool has been used and recognized
Correctness	Assessment based on a tool's objectivity, justness and precision. Objectivity: input and results cannot be easily influenced. Justness: no specific result should form part of the metric tool. Precision: tool measures to a precise degree
Practicality	Required and relevant to both development and improvement
Feasibility usability	Assessment of a tool based on three conditions: Understandability of all formulae in the tool should be high; data collection should be easy; and, evaluation of the metric's results should be convenient
Functionality	The metric tool should meet technical requirements, with a high number of essential process metrics, advanced level strength of visualization and reporting mechanism
Adaptability portability	The metric tool can be integrated to the process methodology. It should be considered as portable and easy to integrate in a minimal time to the software development process management tool

4 Results

The study group consists of 28 software developers. Figure 1 shows the working units of the study group members. The study group members work in four different units: Embedded Software, Software Simulation, Software Testing, and Software Architectural Design. There most common working group for the study

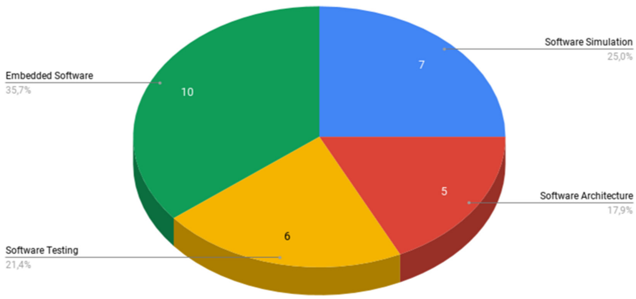


Fig. 2. Study group by working unit

group members was Embedded Software (n = 10, 35.7%), and the least common was Software Architectural Design (n = 5, 17.9%) (Fig. 2).

Cumulative Flow Diagram was the most selected software development process metric, and was selected by 15 (53.6%) of the study group members. Burndown Chart, Velocity Chart, and Burnup Chart were the options selected by more than 10 of the participants. The Blocked Issues Chart and Contribution Chart are marked with an asterisk symbol ("*") because they were additional options added by the participants. Potentially Deliverable Scope and Feature and Epic Progress were each selected by only one study group member. Most software development process measurement components were observed to contain the Cumulative Flow Diagram and Burndown Chart metric options. In other words, the majority of the selections made by the study group can be covered by the software development process metric providers (Fig. 3).

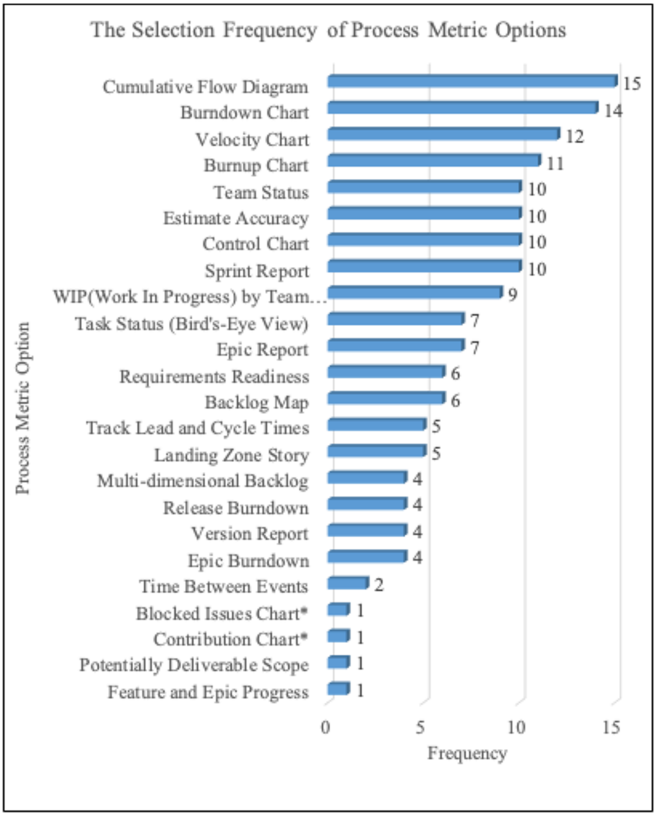


Fig. 3. Selection frequency of process metric options

The add-on products that can be integrated into the software development process management tool used by the TÜBİTAK SAGE software development group were evaluated. It is known that the global marketplace of the tool includes 66 add-on products for software process metrics. Some of these add-on products are considered primitive or are highly specific products that offer only a few metrics, whilst some offer a more comprehensive process metrics service. The tool currently used for software development at the company in this industrial case study provides nine common process metrics. Considering the varied suggestions made by the study group members, it is understood that the current tool contains an insufficient variety of process metrics. Moreover, it is seen that advanced add-on products that offer preferred process metrics were seen as useful by the study group members. As a result, user reviews in the market, user ratings, and software development experts' opinions identified alternatives for four add-on products. The alternatives were determined from feedback of the expert group. Here are the four alternatives (i) Actionable Agile[1], (ii) SenseAdapt Agile[2], (iii) Screenful[3], (iv) Predictable Agile[4] with the determined criteia (1) Feasibility & Usability, (2) Functionality, (3) Relevance, and (4) Experience.

The hierarchical structure was established between the criteria and the alternatives. Figure 4 presents the AHP hierarchical structure of the software development process measurement component selection. It shows the hierarchical structure of the criteria and alternatives determined for the selection of software development process metric add-ons.

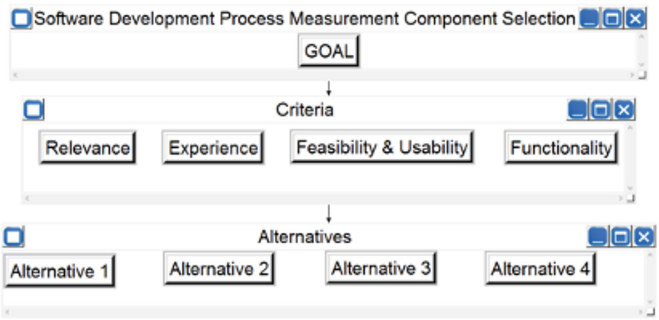


Fig. 4. AHP hierarchical structure of component selection

The decision matrix was created by using the Super Decisions program. The Relevance criterion had the highest weight value of .492, followed by the Experience criterion with a weight value of .225, the Feasibility and Usability criterion with a weight value of .170, and the Functionality criterion had the lowest weight value of .112. As a result of the matrix, weights for each of the alternatives were obtained. Alternative-1 had the highest weight with a normalized value of .402,

[1] https://actionableagile.com/.
[2] https://ripplerock.atlassian.net/wiki/spaces/SEN/overview.
[3] https://screenful.com/tour.
[4] http://agilemontecarlo.com/.

Alternative-3 was second with a normalized value of .236, Alternative-2 was third with a normalized value of .188, and Alternative-4 had the lowest weight with a normalized value of .173.

In the current study, the software development process metrics preferred by the software development group of the company in this industrial case study were reflected. The preferred metrics were considered as being suitable for the characteristic structures of the study groups, and also to the company. Moreover, the most selected metrics such as Cumulative Flow Diagram and Burndown Chart were found to be commonly provided within all software process measurement components. In other words, the study group was not considered to be contradictory in their selections. When evaluating the criteria by the expert group, it can be considered that the study group's preferences were taken into consideration as the Relevance criterion showed a significant value. It can therefore be said that the expert group were afforded sufficient time for the evaluation process so as to produce accurate and representative results, having reflected objective evaluations at the end of the process.

5 Conclusions

In the industrial case study conducted, the AHP method was applied, resulting in the Relevance criterion determined as the highest priority by the software development team. It was considered by the study group's members that the institution, the projects, and the software development team proposed metrics that were deemed suitable for the characteristic structures pertinent to the situation being assessed. Thus, the decision-making expert group examined the graphical results of the collected process metrics.

Here, it can be said that the expert group members were affected in accordance with the need for the *Relevance* criterion. In addition, it can be said that information such as comments, scores and the use of percentages from those with relevant experience may have been the reason for giving second priority to the *Experience* criterion. In addition, it was determined that the study group members gave almost equal priority to both the Functionality and the Feasibility and Usability criteria.

Expert-4, who considered the Functionality criterion to be of a low priority, verbally stated that "It is more important to use the proportion of features than the high functionality of the vehicle." An example of this would be that some of the tools used by the company have been upgraded with 100 new features; yet these new features were not being used, whilst the basic functions were continued to be used instead.

It was also determined that Alternative-1 was prioritized as the most preferred with 40.259% following the systemic evaluative process of the available alternatives. It was observed that the expert group had difficulty in applying discrete thinking at the beginning of the evaluation process. Once this problem had been identified, it was determined that the expert group had successfully completed this process when an appropriate time and place were considered for them to ponder their decisions.

The industrial case study of the current study has proven that multi-criteria decision-making process, and in particular the AHP method, can be used effectively for selecting software metric components within the software domain. In the future, it is considered that the current study may serve as an example of the AHP method being preferred when systematic decision-making is required in the field of software development. In addition, it is considered that the contribution of the software development team to the decision-making process positively supports the use of the selected metric components. The use of appropriate software development components increases the ability of companies to compete in the future.

Acknowledgments. This research work is supported, in part, by TUBITAK SAGE. During the research, we have received help and generous support from many people, to all of whom we would like to express our gratitude.

References

1. Pfleeger, S.L.: Software metrics: progress after 25 years? IEEE Softw. **25**, 32–34 (2008)
2. Yilmaz, M.: Observed effects of software processes change in three software firms: industrial exploratory case study. Pamukkale Univ. Muh. Bilim. Derg. **25**, 240–246 (2019)
3. Lee, M.C., Chang, T.: Software measurement and software metrics in software quality. Int. J. Softw. Eng. Appl. **7**, 15–34 (2013)
4. Farooq, S.U., Quadri, S., Ahmad, N.: Software measurements and metrics: role in effective software testing. Int. J. Eng. Sci. Technol. **3**, 671–680 (2011)
5. Yilmaz, M., O'Connor, R.V., Clarke, P.: Effective social productivity measurements during software development-an empirical study. Int. J. Softw. Eng. Knowl. Eng. **26**, 457–490 (2016)
6. Manifesto: Software Process Improvement (SPI). https://2019.eurospi.net/images/eurospi/spi_manifesto.pdf. Accessed 3 Sept 2019
7. Yilmaz, M.: A software process engineering approach to understanding software productivity and team personality characteristics: an empirical investigation. Ph.D. thesis, Dublin City University (2013)
8. Paulish, D.J., Carleton, A.D.: Case studies of software-process-improvement measurement. Computer **27**, 50–57 (1994)
9. Yilmaz, M., O'Connor, R.: Social capital as a determinant factor of software development productivity: an empirical study using structural equation modeling. Int. J. Hum. Cap. Inf. Technol. Prof. (IJHCITP) **3**, 40–62 (2012)
10. Yilmaz, M., O'Connor, R.V.: A scrumban integrated gamification approach to guide software process improvement: a Turkish case study. Tehnički Vjesn. **23**, 237–245 (2016)
11. Jones, C.: A Guide to Selecting Software Measures and Metrics. CRC Press, Boca Raton (2017)
12. Yilmaz, M., Atasoy, B., O'Connor, R.V., Martens, J.-B., Clarke, P.: Software developer's journey. In: Kreiner, C., O'Connor, R.V., Poth, A., Messnarz, R. (eds.) EuroSPI 2016. CCIS, vol. 633, pp. 203–211. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-44817-6_16

13. Yilmaz, M., O'Connor, R.V., Clarke, P.: A systematic approach to the comparison of roles in the software development processes. In: Mas, A., Mesquida, A., Rout, T., O'Connor, R.V., Dorling, A. (eds.) SPICE 2012. CCIS, vol. 290, pp. 198–209. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-30439-2_18
14. Wan, J., Zhu, Y., Zeng, M., et al.: Case study on critical success factors of running Scrum. J. Softw. Eng. Appl. **6**, 59 (2013)
15. Pfleeger, S.L., Fitzgerald Jr., J.: Software metrics tool kit: support for selection, collection and analysis. Inf. Softw. Technol. **33**, 477–482 (1991)
16. Card, D.N., Glass, R.L.: Measuring Software Design Quality. Prentice-Hall, Inc., Upper Saddle River (1990)
17. Wallace, L.G., Sheetz, S.D.: The adoption of software measures: a technology acceptance model (TAM) perspective. Inf. Manag. **51**, 249–259 (2014)
18. Fenton, N.E., Neil, M.: Software metrics: roadmap. In: Proceedings of the Conference on the Future of Software Engineering, pp. 357–370 (2000)
19. Kitchenham, B., Pfleeger, S.L., Fenton, N.: Towards a framework for software measurement validation. IEEE Trans. Softw. Eng. **21**, 929–944 (1995)
20. Akarsu, Z., Ozgun, M., Kuru, Y., Yilmaz, M.: Using adapted version of Hoshin matrix for selection of agile software development processes. In: 12th Turkish National Software Engineering Conference (UYMS 2018), pp. 1–11 (2018)
21. Marks, G., O'Connor, R.V., Yilmaz, M., Clarke, P.: An ISO/IEC 12207 perspective on software development process adaptation. Softw. Qual. Prof. **20**, 48–58 (2018)
22. Ebert, C., Bundschuh, M., Dumke, R., Schmietendorf, A.: Making metrics a success-the business perspective. In: Ebert, C., Bundschuh, M., Dumke, R., Schmietendorf, A. (eds.) Best Practices in Software Measurement: How to Use Metrics to Improve Project and Process Performance, pp. 9–34. Springer, Heidelberg (2005). https://doi.org/10.1007/3-540-26734-4_2
23. Saaty, T.L.: What is the analytic hierarchy process? In: Mitra, G., Greenberg, H.J., Lootsma, F.A., Rijkaert, M.J., Zimmermann, H.J. (eds.) Mathematical Models for Decision Support. NATO ASI Series, vol. 48, pp. 109–121. Springer, Heidelberg (1988). https://doi.org/10.1007/978-3-642-83555-1_5
24. Sharma, N.K., Gupta, S., Sharma, V.: A comparatively study for project management software selection using analytic hierarchy process method. Int. J. Eng. Manag. Res. (IJEMR) **5**, 188–195 (2015)
25. Sagar, S., Mathur, P., Sharma, A.: Multi-criteria selection of software components using fuzzy-AHP approach. Int. J. Innov. Comput. Inf. Control **11**, 1045–1058 (2015)
26. Ömürbek, N., Makas, Y., Ömürbek, V.: AHP ve TOPSIS yöntemleri ile kurumsal proje yönetim yazilimi seçimi. Süleyman Demirel Üniversitesi Sosyal BilimlerEnstitüsü Dergisi, pp. 59–83 (2015)
27. Al-Qutaish, R.E., Muhairat, M.I., Al-Kasasbeh, B.M., Al-Kasasbeh, B.: The analytical hierarchy process as a tool to select open source software. In: Proceedings of the 8th WSEAS International Conference on Software Engineering, Parallel and Distributed Systems, Cambridge, UK, pp. 39–44 (2009)
28. Zaidan, A.A., Zaidan, B.B., Al-Haiqi, A., Kiah, M.L.M., Hussain, M., Abdulnabi, M.: Evaluation and selection of open-source EMR software packages based on integrated AHP and TOPSIS. J. Biomed. Inform. **53**, 390–404 (2015)
29. Orgun, P., Gungor, D., Kuru, Y., Metin, O., Yilmaz, M.: Software development overall efficiency improvement in a CMMI level 5 organization within the scope of a case study. In: Uluslararası Bilgisayar Bilimleri ve Mhendisliği Konferansı (UBMK 2018), pp. 1–12 (2018)

A Multivocal Literature Review of Function-as-a-Service (FaaS) Infrastructures and Implications for Software Developers

Jake Grogan[1], Connor Mulready[1], James McDermott[1], Martynas Urbanavicius[1], Murat Yilmaz[1,2], Yalemisew Abgaz[1], Andrew McCarren[1], Silvana Togneri MacMahon[1,2], Vahid Garousi[3], Peter Elger[4], and Paul Clarke[1,2](✉)

[1] School of Computing, Dublin City University, Dublin, Ireland
{jake.grogan8, connor.mulready2, james.mcdermott7, martynas.urbanavicius2}@mail.dcu.ie, {murat.yilmaz, yalemisewm.abgaz, silvana.macmahon, paul.m.clarke}@dcu.ie, amccarren@computing.dcu.ie

[2] Lero – the Science Foundation Ireland Research Centre for Software, Limerick, Ireland

[3] School of Electronics, Electrical Engineering and Computer Science, Queens University, Belfast, UK
v.garousi@qub.ac.uk

[4] fourTheorem Limited, Cork, Ireland

Abstract. In this paper, we provide a multivocal literature review of Function as a Service (FaaS) infrastructures. FaaS is an important, emerging category of cloud computing, which requires that software applications are designed and deployed using distributed, highly-decoupled service-based architectures, one example of which is the microservices architecture paradigm. FaaS is associated with on-demand functionality and allows developers to build applications without the overhead associated with server management. As such, FaaS is a type of serverless provisioning model wherein a provider dynamically manages and allocates machine resources, with the developers deploying source code into a production environment. This research provides an analysis of scalability, cost, execution times, integration support, and the constraints associated with FaaS services provided by several vendors: AWS Lambda, Google Cloud Functions, and Azure Functions. We discuss the implications of the findings for software developers.

Keywords: Functions-as-a-Service · Infrastructures · Serverless · Cloud computing · Scalability · Constraints · AWS Lambda · Microsoft Azure · Google Cloud Functions

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 58–75, 2020.
https://doi.org/10.1007/978-3-030-56441-4_5

1 Introduction

Software engineering is a complex undertaking [44], the process for which must take account of various situational factors [45, 46]. One such factor is software architecture which comes in many forms, but which is presently challenged to develop software systems that can be deployed more quickly and with less disruption to existing operational code, such as is exemplified in continuous software engineering for microservices architectures [47]. A further benefit, and a key consideration in commercial software engineering settings, is the fact that microservices architectures can take advantage of emerging lower cost cloud hardware provisioning models, one example of which is Function-as-a-Service (FaaS), defined by IBM as "a type of cloud-computing service that allows you to execute code in response to events without the complex infrastructure typically associated with building and launching microservice applications." [29]. FaaS, perhaps the most central technology in the serverless model, is focused on the event-driven computing paradigm wherein application code, or containers, only run in response to events or requests [29]. Fundamentally, FaaS is about running backend code without the need for companies or individuals to manage their own server systems or long-lived server applications [18].

AWS Lambda, Google Cloud Functions, and Azure Functions are among the most commonly used FaaS services in industry today [30]. Each vendor offers a different set of capabilities with their FaaS infrastructure implementations, from language runtime support and memory usage to the ability to execute functions at regional edge cache locations in response to events generated by content delivery networks [31]. When considering FaaS as part of a systems architecture, it is vital to choose the solution that works best for the system under consideration. For this reason, it is vital that factors surrounding FaaS infrastructures which influence this decision are discussed and investigated.

A multivocal literature review (MLR) [49] approach was adopted to investigate this topic, with the goal of understanding the current FaaS landscape, identifying commonalities and differences in leading vendor offerings, and identifying implications for software developers. Both peer reviewed sources and non-peer reviewed, grey sources, were included. This is important as perspectives from both academia and industry professionals are considered. The area of "FaaS infrastructures" is broad and fragmented and for this reason it was of utmost importance to identify relevant search terms and to include a broad range of literature in this study.

The set of FaaS infrastructure factors under analysis in each of our primary sources varies from source to source, therefore, we suggest, that to analyse various FaaS infrastructures, as many of these factors as possible must be taken into account. The role of software engineer is one that is known to be changing [48], and we suggest that FaaS will contribute to further change for software engineers, we therefore outline some of the major implications for software developers as required by the FaaS innovation.

In this paper, factors which will be analysed include execution times, memory configurations, abilities to scale, pricing and cost of FaaS services, the constraints of FaaS infrastructures, and how well integrated vendor FaaS infrastructures are, not only with their own platforms, but how they can be integrated with third party services. This is an important factor as vendor lock-in is a major barrier to the adoption of cloud computing due to the lack of standardisation [32]. These factors were selected as they represent the major service offering characteristics as presented by the vendors under study.

2 Related Literature

2.1 Methodology

This research has been conducted in the context of a multivocal literature review [49, 51] using both grey and white literature. The primary research was conducted by a team of four students who were allocated the research paper title: “An Analysis of Functions-As-A-Service Infrastructures”. The topic was examined under four key sub-topics: How do FaaS providers compare in performance and scalability? How do constraints compare between AWS Lambda, GC Functions and Azure Functions? How do FaaS providers differ in pricing? and, How well integrated are FaaS infrastructures with vendor platforms?

The team identified both academic and non-academic sources from searches on google.com and scholar.google.com, as well as IEEE, ACM and Wiley, among others. From these sources, we identified keywords relevant to the research being conducted and snowballing was incorporated to identify important sections of current sources and to identify other possible sources for inclusion in this research. These identified sources were examined and marked according to an inclusion/exclusion criteria outlined in Sect. 2.2, leading to 43 sources being included in this analysis.

Given that the topic under study is broad, sub-topics were agreed upon by the four primary researchers as they emerged as the most pressing questions raised in other literature when comparing key elements of various FaaS providers. This research was conducted in the period February–April 2020, the primary focus of this research was to gain the best possible understanding of the topic area, dive deep on the identified research questions and present our findings in a coherent manner within this paper.

Each member of the team was allocated one of the four research questions identified in order to conduct a review of said research question. This led to each member gaining a deep understanding of their allocated topic. Each member spent four weeks researching their topic and combining their findings in a single, summarized and coherent document in the final two weeks. On a weekly basis throughout this period, the student team liaised with the research supervisor, for guidance on methodology and research direction.

2.2 Inclusion/Exclusion Criteria

Following the completion of the search and source selection phase, the identified sources underwent further evaluation in order to grade their relevance to our study. In our criteria, high impact factor journals were given preference over lower impact factor publications, and when evaluating white literature for inclusion in our sources, we graded these sources based on the number of citations they have received elsewhere and their relevance to the research questions. When including non peer-reviewed grey literature such as blogs or videos, the team approached with caution, only including those sources which are considered reliable and relevant based on the level of administration or moderation of the source and the level of bias introduced by the source, though clearly even well moderate online blogs fall some way short of the rigor, validity and generalization expected from academic peer reviewed research articles.

3 Research Questions and Findings

In this section, we will present the research questions and the associated findings from our literature review. The primary focus in this analysis is to examine three of the established major FaaS providers and to compare and contrast their various infrastructures: Microsoft, Google and AWS. Specific concerns for software developers targeting FaaS platforms are expressed in the form of research questions. For example, the ability of FaaS platforms to scale under higher system load is an important consideration, as it the basic runtime performance of the platform. The amount of memory available to runtime functions is also a key consideration in software design, as is the maximum allowable number of concurrent running functions. The FaaS pricing model is furthermore a key operational concern. These and other considerations are examined in each of the three providers under study.

3.1 RQ 1: How Do FaaS Providers Compare in Performance and Scalability?

In this section, we compare and contrast AWS Lambda, Google Cloud Functions and Azure Functions in terms of their relative performances and their abilities to scale. For the purposes of this research and due to limitations outlined in Sect. 5, we limit the definition of performance to hot-start execution times (Sect. 3.1.1) and cold-start execution times (Sect. 3.1.2).

From our primary studies [1–7, 22], hot-start execution times are dependent on a number of factors. In Sect. 3.1.1, we discuss the factors of language runtimes and memory configurations. Each vendor, AWS, Google and Microsoft Azure, offer different language runtimes which will be outlined in Sect. 3.2.1. Our primary studies focus on NodeJS, Python and C# .NET. Each vendor also offers different memory configurations, each of which will be outlined in Sect. 3.2.2, however, we will discuss the impact of varying memory configurations on performance in Sect. 3.1.2.

3.1.1 Hot-Start Execution Times

In this section we compare hot-start execution times between FaaS vendors. From our primary studies, hot-start times vary dramatically from runtime to runtime. In tests carried out by D. Jackson and G. Clynch [3] on AWS Lambda, the average execution times (ms) across all warm start tests show Python has the best execution time of 6.13 ms followed closely by C# .NET with an average execution time of 6.39 ms. Go had an average execution time of 19.21 ms making it the poorest performer [3].

It may be unexpected for C# .NET to come second in the test as the Just-In-Time nature of the compiler would be expected to be slower [3, 7].

Comparing to Azure, of the two runtimes (NodeJS and C# .NET) tested, C# far outperforms NodeJS, showing an average sub-millisecond performance of just 0.93 ms compared to 4.91 ms for NodeJS.

For hot-start tests, C# .NET significantly out-performed on Azure compared to AWS. This is expected however, as C# is a Microsoft technology and Azure Functions would be expected to have a solid support for .NET. This suggests that vendor specific implementations of their FaaS infrastructures impact execution time by language runtime as Azure Functions are run on windows containers compared to AWS Lambda which uses the open-source .NET Common Language Runtime on Linux containers for C# [3].

3.1.2 Cold-Start Execution Times

From our primary studies it is evident that cold-start times are of great importance when comparing execution times as very little literature focuses on hot-start executions. In this section we compare the cold-start times of AWS, Google and Azure using the NodeJS language runtime.

AWS Lambda with a memory configuration of 128 MB had a median cold-start execution time of 265.2 ms compared to Google which had a median execution-time of 493 ms with the same memory configuration [8]. Azure assigns their VM instances 1.5 GB of memory [5, 8]. The cold start execution time of Azure Functions was 3640 ms [8], a stark increase compared to AWS and Google. From the findings of the tests carried out by L. Wang *et al.* [8] an interesting result is that when the memory assigned to functions is increased (AWS at 1536 MB and Google at 2048 MB), the median cold-start execution of AWS is 250 ms, only a 15 ms decrease in execution time, however, the median cold start for Google Cloud Functions was 110 ms, a decrease of 140 ms [8]. This suggests that memory size has a far greater impact on cold-start execution times with Google Cloud Functions compared to AWS Lambda.

Further tests were carried out by L. Wang *et al.* [8], whereby cold-start tests were carried out over a 168 h period and the median execution times were calculated each hour across the three vendors. Results from these tests show AWS has the most stable cold-start execution times of $\sim$200 ms. Google Cloud Functions also had relatively stable cold-start times of $\sim$400 ms (except for a few spikes). Azure had the highest variation over time, ranging from 1500 ms to 16000 ms [8].

From other primary studies, it also suggests again that a vendor's infrastructure implementation can impact cold start times. One study carried out [3] compares cold-start times with AWS Lambda and Azure Functions on the C# .NET runtime. The primary study shows AWS had an average cold-start execution time of 2500 ms, compared to Azure which had an average cold-start time of 276.4 ms [3]. We refer our reasoning back to Sect. 3.1.1 where Azure Functions are run on windows containers compared to AWS Lambda functions which are run on Linux containers.

3.1.3 Scalability

Scalability is an extremely important factor for anyone considering building or moving parts of their infrastructure to a serverless infrastructure. In this section we analyse findings from our primary studies on the ability of FaaS infrastructures to scale.

In tests performed by G. McGrath *et al.* [2], a framework was developed to test the ability of serverless FaaS platforms to performantly invoke functions at scale. Starting with a single invocation call, every 10 s an additional concurrent call, up to a maximum of 15 was added. The findings of their tests show AWS Lambda is able to scale linearly and exhibits the highest throughput of the three platforms being examined in this review. Google Cloud Functions scales sub-linearly but begins to taper off as the number of concurrent requests approaches 15. Again, similar to the cold-start execution times discussed in Sect. 3.1.2, Azure experiences a high degree of variance in the number of concurrent requests it can handle. Initially it outperforms AWS and Google but that number drops as the number of concurrent requests increases, then decreases, and continues to fluctuate in this manner [2].

3.2 RQ 2: Constraints Comparison of FaaS Providers

In this section we compare FaaS vendors and highlight constraints associated with each. From our primary studies, the main constraints associated with FaaS vendors lie in the area of supported language runtimes, memory configurations and integration capabilities. On the topic of integration capabilities, we analyse the vendor lock-in problem and the extent to which each vendor suffers from this. The main constraints are outlined in Table 1.

Table 1. Constraints Comparison

	AWS	Azure	Google Cloud
Supported runtimes	Nodejs, Python, Ruby, Java, GO, C#, Powershell, additional languages via runtime API	Nodejs, Python, Java, C#, F#, Powershell	Nodejs, Python, GO
Maximum concurrent executions	1000 *upgradable (Sect. 3.2.1)	200	1000
Minimum function memory	128 MB	128 MB	128 MB
Maximum function memory	3008 MB	1536 MB	2048 MB
Maximum function timeout	900 s	600 s	540 s
Maximum deployment package size	50 MB (compressed) for sources 250 MB (uncompressed) for sources, modules	No limit	100 MB (compressed) for sources 500 MB (uncompressed) for sources, modules
Maximum HTTP request/response payload size	6 MB	100 MB	10 MB
Maximum number of functions	No limit	No limit	1000 functions per project

3.2.1 Performance and Scalability Constraints

AWS offers support for the greatest number of language runtimes. AWS supports NodeJS, Python, Ruby, Java, GO, C#, and Powershell language runtimes [3, 6] as well as support for third party runtimes via the AWS Lambda Runtime Interface [40]. Azure offers support for C#, F#, NodeJS, Java, Powershell, Python, and Typescript [3, 5]. Google Cloud has the least extensive list of supported language runtimes with support only for NodeJS, Python and GO [4].

AWS and Google Cloud have a default limit of 1000 simultaneous executions for a function, but AWS allows this limit to be increased on request. Azure in comparison has a limit of only 200 simultaneous executions [2, 4–6].

3.2.2 Memory Constraints

Each vendor has different limits on the amount of memory available to the function during execution. All three providers offer the same minimum memory configuration for functions, but have different maximum limits, Azure has the lowest maximum memory limit of 1536 MB [5] followed by Google Cloud with a limit of 2048 MB [1, 4] and AWS with the largest limit of 3008 MB [1, 6], therefore being able to support functions that require more memory resources than other vendor FaaS services can provide. For AWS lambda the allocated memory linearly translates to the CPU power available for the function [9].

FaaS providers limit the amount of time that a function is allowed to run before it times out. AWS has a timeout of 900 s, Azure 600 s and Google Cloud 540 s [4–6]. Functions are shipped as packages and some providers have certain package size limits. Out of the three providers Azure is the only FaaS vendor not to have any package size limit. AWS has a size limit of 50 MB for compressed sources and a 250 MB uncompressed limit for the whole package which includes dependencies [36]. Google Cloud doubles those limits at 100 MB and 500 MB respectively [4]. Functions can be triggered by HTTP request events which can contain data to be passed to the function for processing. AWS limits this data up to 6 MB, Google Cloud at 10 MB and Azure at 100 MB [4, 5, 36].

3.2.3 Vendor Lock-In

Vendor lock-in refers to a problem in cloud-computing wherein customers become dependent on a single cloud provider technology implementation. It may become difficult for a customer to move to another vendor without substantial costs or possible legal constraints [32].

In the case of FaaS, vendor migration can be very expensive if a language runtime does not have cross vendor support. As covered in Sect. 3.2.1 Google Cloud supports only a fraction of language runtimes in comparison to other providers. This can pose a migration challenge if the existing codebase languages are not supported by the platform.

As FaaS infrastructures are generally well integrated with the rest of a vendors serverless platform (as outlined in Sect. 3.4), they may not integrate as well into third party services or other vendor platforms. Interoperability and portability are essential qualities that affect FaaS infrastructures and pose as a major barrier to entry into cloud-computing [32].

As FaaS are event driven services, attempts have been made to standardise how event publishers describe events in order to support interoperability and portability. An example of one such attempt is CloudEvents [37]. Recently, Azure have announced first class support for CloudEvents [38]. AWS took a different direction when they launched EventBridge [39]. The format to describe events in AWS EventBridge differs from that of CloudEvents, which has not helped in solving the issue of cross-vendor interoperability [52, 53]. Therefore there is still a constraint of integrability between vendors and their FaaS infrastructures.

3.3 RQ 3: How Do FaaS Providers Differ in Pricing?

In this section, we conduct a cost analysis comparison of FaaS solutions such as AWS Lambda, Microsoft Azure, and Google Cloud Functions. This section looks at how the pricing of a function model is influenced by performance and invocation of that function. For the purposes of this research, we will look in particular at the factors that inform the difference in pricing by limiting the research to each FaaS vendor's free tier, the tier which is most widely accessible.

From our initial studies, we can discern that each vendor, AWS, Microsoft Azure and Google Cloud, offer broken down examples of how each billing is computed. There are three main factors in a function's execution cost - execution time, fixed invocation cost per individual function execution and memory allocated to the function [3]. Thus, we explore how request invocations, the duration of execution, memory provisioning and compute time influence the pricing model, while also taking time to delve into the double billing of functions, as it violates a principle of the serverless model [28]. This primary study does not focus on a particular set of languages or runtimes as pricing is language-agnostic and follows a flat charge rate for each factor.

3.3.1 Request Invocations

Our primary studies indicate the prioritisation of request invocations as an important contributing factor in the pricing of the FaaS solution. In this section we compare the number of free invocations granted per month as well as the charge rate after passing that threshold.

In accordance with the Google Cloud Functions free tier pricing [25] invocations are charged at a flat rate, independent of the source of the invocation, including functions invoked from a HTTP request (HTTP functions), background and invocations resulting from the call API. These invocations are charged at a per-unit rate of $0.0000004 per invocation, and this excludes the first 2 million free invocations per month, though these free invocations are charged regardless of the outcome of the function or its duration.

In the case of AWS Lambda pricing [23], we note that the first million requests per month are free. After this point, the invocation of a function is not charged at a per-unit rate but rather at a per-million rate, with the next one million requests charged at $0.20.

When looking at Microsoft's Azure Functions pricing [24], invocations are counted each time a function is executed in response to an event, triggered by a binding. The first million executions are included free each month. Further executions are charged like AWS Lambda, at a price-per-million of $0.20.

In regards to contrasting the factor of pricing request invocation, we can see from the primary studies that Google Cloud Functions grants the most free requests on a per monthly basis, boasting one million more free invocations per month than both Azure Functions and AWS Lambda. In terms of comparing the charge rate thereafter, we must look at price-per-million, as both Azure and AWS offer the price in this format already. Taking Google Cloud Functions price-per unit rate of $0.0000004 and multiplying by one million, we are left with the product of $0.40. We can now say that in regards to charges after passing the free request invocation threshold, Azure Functions and AWS Lambda both offer the cheapest access based on the number of requests invoked.

3.3.2 Duration

From our primary studies of all three FaaS vendors we can see the prioritisation of duration, or time until the completion of execution, as an important contributing factor in the pricing of the FaaS solution. In this section we compare the amount of compute time freely granted by each provider, along with how it was calculated, and the associated charging after passing this granted threshold. Compute time is measured from the time your function receives a request to the time it completes, either through your signaling of completion, or through a timeout, other failure or any other termination [25].

Taking a look at granted free tier resources we can note that Azure Functions pricing includes a monthly fee grant of 400,000 GB-seconds of resource consumption per month per subscription [24]. With AWS Lambda pricing, free tier provides 400,000 GB-seconds of resource consumption per month per subscription [23]. Finally looking at Google Cloud Functions pricing, the free tier also provides 400,000 GB-seconds [25].

Beyond this point we can see that AWS Lambda bills at $0.0000166667 per GB-second [3, 23, 25, 26], Azure Functions are billed at a flat rate of $0.000016 per GB-second [3, 24] and Google Cloud Functions at a rate of $0.0000025 per GB-second [25]. Comparing across vendor's we see each free tier capping the allowance of free resource consumption at the 400,000 GB-seconds mark.

3.3.3 Memory Provisioning

Looking beyond the granted allowance, in Azure Functions, memory used by a function is measured by rounding up to the nearest 128 MB, up to the max. memory size of 1,536 MB [24].

With AWS Lambda the memory allocated to a function can be between 128 MB and 3008 MB, in 64 MB increments [23, 26]. With Google Cloud Functions the memory allocated to a function can be between 128 MB and 2048 MB, in doubling increments [25]. With provisioning we see the widest range of options offered by AWS, offering up to 3008 MB, while Azure and Google Cloud respectively offer 1536 MB and 2048 MB max. [23–26].

3.3.4 Compute Time

The cost of serverless functions are directly related to their execution times. This is due to the prevalent billing model across the major serverless platforms of cost per milliseconds of execution [3]. Both AWS and Azure solutions round up to the nearest 1 ms [23, 24] while the Google Cloud solution rounds to the nearest 100 ms [25]. Across each provider except Azure, compute time is measured in 100 ms increments, with AWS Lambda offering a significantly cheaper solution (at \$0.0000002083 per 100 ms) than Google Cloud (\$0.0000025 per 100 ms) [3, 23–25]. The execution time of Azure Functions is calculated by rounding up to the nearest 1 ms, with a price of \$0.0000000034167 per 1 ms.

3.3.5 Double Billing

Double billing refers to charging for the same product twice. In the case of FaaS vendor's double billing refers to the synchronous invocation of other functions [27].

When a function makes a call to another service, you pay for the waiting time—even if the *code* is using async I/O. This means you wait for the result of each function called, and hence must pay for the wait time. This wait time is directly correlated to Sect. 3.3.2. This additional billing violates one of the serverless principles: pay for what you use, not idle time [28].

Without intrinsic support from the FaaS vendor, it is impossible to avoid double billing while still maintaining the core principle of the serverless computing model [28].

3.4 RQ 4: How Well Integrated Are FaaS Infrastructures with Other Services from the Same Vendor?

In this section we will cover what the serverless model is and the possibilities of integrating a pre-existing application and other services with it through the use of FaaS. In order to achieve this, it is necessary to cover the different architectural services attached to the serverless model for AWS, Google Cloud and Azure (Sect. 3.4.1). Depending on certain workflows different sections of the serverless model will be more prominent than others and understanding the entry criteria for these services for each provider will also be necessary (Sect. 3.4.2). By the end of this section the core differences between each provider's serverless model should be apparent, along with the integration capabilities of the provided services and FaaS.

3.4.1 What Is the Serverless Model?

Serverless models abstract the underlying computer infrastructure. It eliminates infrastructure management tasks such as server or cluster provisioning, patching, operating system maintenance, and capacity provisioning. It also works as a pay-what-you-use model [33–35].

In this section we will cover the different architectural related serverless services provided by each provider. AWS, Google Cloud and Azure offers traditional and unique services which fall under the compute, storage, data stores and integration categories. Noting that even though these categories are used in both serverless and traditional cloud computing, services retaining to traditional cloud computing are outside the scope of this paper.

3.4.2 What Architectural Services Does the Serverless Model Provide?

FaaS and serverless are generally synonymous with one another due to its ability to listen and act upon the events of other serverless services (Sect. 3.4.3). However FaaS is a subset of serverless [20], which resides within the compute category. By understanding the variety of categories and the services that lie within we can better ascertain the best provider for the users architectural needs. The core categories that relate to architecture that are consistent across all three providers are the following, compute, storage, data stores and integration. The main purpose of this research is to see how FaaS integrates with other serverless services. For the sake of brevity, other compute services are omitted as the core focus of this research is FaaS.

3.4.2.1 Storage

Storage provides consistent object-level data storage across all three providers [34, 35, 43]. However AWS also provides EFS; an NFS file system service that can be easily integrated with on-premises or cloud resources [35].

3.4.2.2 Data Store

In regard to data stores AWS provides the most unique services, those being DynamoDB, AWS DocumentDB, AWS MCS, and Amazon Aurora Serverless [35, 41, 42]. However, in regard to variety Azure Cosmos DB provides a multi-model with wire protocol compatible API endpoints for Cassandra, MongoDB, SQL, Etcd and Table [43]. Finally, Google Cloud provides their Cloud Firestore which is a NoSQL database [34].

3.4.2.3 Integration

As the topic of this section, integration provides the services required in order for pre-existing resources to interact with the cloud's serverless environment. Whilst integration has a plethora of services such as Simple Notification Service for AWS, Pub/Sub for Google Cloud and Messaging for Azure, the core services that will be of most use for integration in relation to FaaS are API related creation & management services such as API Gateway for AWS and API Management for Azure [35, 43]. Google Cloud has no direct service for this however their Cloud functions can be directly invoked by using traditional HTTP means (Sect. 3.4.3.1).

During our studies we also found that a large portion of the integration services based on serverless orchestration make heavy use of FaaS and act as a layer of abstraction to the service. An example of such a service would be AWS Step Functions [11, 21].

3.4.3 FaaS Entry Criteria

In this section, we will look at FaaS and its ability to execute functions based on events that occur within the serverless model. This functionality allows for the design of workflows and promotes automation within the serverless model. As services may differ for each provider, so do the events, triggers and bindings associated with the relative services as discussed in the sections below. The primary study here is not to list all unique triggers but establish the possibilities of executing FaaS using external resources and in turn, highlight its transitive nature by covering the possibilities of invoking additional calls to services.

3.4.3.1 On-Premises to FaaS

Executing lambda, cloud functions and Azure functions from outside their respective provider is possible, however the process of doing so differs from provider to provider but the underlying approach remains consistent.

Google Cloud Functions is the most straightforward with each function being accessible through a RESTful approach [4]. Accessing Lambda from outside the AWS platform requires the use of API Gateway. By using API gateway it is possible to create a public-facing API and attach a respective lambda function to an endpoint. With this additional step, the function can be triggered via a RESTful API call [9]. Azure Functions can be triggered over HTTP also without any additional set up. However, in order to receive a HTTP response from the function a corresponding binding (Sect. 3.4.3.2) must also be created for that particular function [10].

3.4.3.2 FaaS and Serverless Integration

As stated in Sect. 3.4.2, FaaS is a subset of the serverless architecture. As it lies within the domain of serverless, it has the ability to react to other serverless service actions. In this section we will cover the different criteria in which other services can execute serverless functions and the overhead involved in doing so.

AWS Lambda can be viewed as a reactionary service. Depending on specific events that occur within the cloud platforms environment, it may or may not trigger a Lambda function. Whilst the list of possible triggers is vast [9], the general consensus is that any event within the services defined in Sect. 3.4.2 contain events that can trigger a function. Additional services may also trigger Lambda for example, AWS Cloud Watch events or responses to HTTP as mentioned in Sect. 3.4.3.1 [9, 12, 13].

Google cloud defines events as “things that happen within your cloud environment that you might want to take action on” [14]. It also defines a trigger as “a declaration that you are interested in a certain event or set of events”. In regards to integration, the list of events provided by google may seem inherently limited [14], however a transitive approach may be taken with services to allow for additional integration with other resources within or outside of the platform [15].

Whilst Azure has the same functionality as the other platforms, it uses different terminology to a certain degree. Events exist on Azure as evidenced by the Event Hub service (a service which allows the user to orchestrate and consolidate particular events in a single easy to manage space [16]). Azure also provides the additional optionality of manually creating bindings between Azure functions and serverless resources [17]. Whilst Event Hub adds a layer of abstraction over the operation, bindings provide a hands on approach for the user thereby allowing them to further customise the level of granularity needed for integration. As the only difference between bindings and Event Hub is abstraction, any service available in Event Hub will also be available to listen to through bindings [17].

4 Limitations of Research

FaaS is a new and emerging technology. Whilst it first began to be discussed around 2010 [19], it is only in recent years that the paradigm itself has witnessed sustained interest and is now being offered by the largest cloud providers (AWS, Azure, Google Cloud) as a service. As the four primary researchers did not have previous experience with research, there was some difficulty in finding informative, trusted sources on the subject and a degree of uncertainty when judging the reliability of some grey sources included in this review. However, advice and training were provided by more senior researchers each week throughout the six-week literature review period on how to correctly conduct research, and in particular, multivocal literature reviews. Had more than six weeks been available to the primary research team, a broader and more comprehensive review would have been possible.

Although AWS, Google Cloud and Microsoft Azure provide documentation for their FaaS services, we found it to be largely oriented towards developers learning to use the service (API documentation) rather than the underlying infrastructure. Whilst more topics could have been included and certain topics which were included could have been expanded on, due to the six week time constraint given, the team had to limit themselves in terms of scope. Had more time been available to the researchers, the topics analysed in this paper could have undergone further testing and validation through broader and deeper research outlined in Sect. 2. Furthermore, that service offering information is from the vendors themselves represents a limitation in that it has not been subject to independent verification.

With these limitations having been considered, the research team feels the relevance of the work carried out has not been diminished and that this work is still a useful contribution to the study of serverless computing and, in particular, FaaS infrastructures.

5 Implications for Software Engineers

While FaaS infrastructures might not immediately appear to have a major impact on software developers and software development processes, upon deeper examination, we find that there are several important considerations that should not be overlooked. Apart from any specific changes related to FaaS adoption, it has previously been shown that changes in software contexts are a common phenomenon and that this has implications for the software process [54–57] and business performance [58]. FaaS, we argue, is a significant change in the background hardware provisioning context.

To take advantage of the potential cost savings and elastic scalability that FaaS infrastructure provides, it is necessary to design and architect a software system that is suited to FaaS, and as the analysis in Sect. 4 has shown, there are many details involved and they vary not insignificantly amongst the FaaS vendors. Traditional so-called monolith-based architectures will simply not be in a position to take advantage of the reduced operational risk and costs that FaaS can deliver, and so there is an emerging large job for the broader software industry to modernize their architectures to take advantage of FaaS. In the authors' opinion, this is an aspect of software development work which will be large and strategically important in many software companies in the coming years.

For individual software developers, the early parts of their careers may have been characterized by a largely distant relationship with operational infrastructure that was provisioned and managed by another department in their business or perhaps externally. However, with FaaS, individual developers may now be required to change their process of work towards actually issuing their own code in the form of functions to the FaaS infrastructure. The performance and cost of individual functions will be monitored and so the individual developer in a FaaS environment can expect to be drawn closer to the real operational system than might have been the case heretofore. Indeed, in some instances, function/service developer(s) might be entirely responsible for the creation, operation and maintenance of their code. Such a move would reduce costs related to communication and training in software companies, and could be disruptive to existing development processes, pushing DevOps to the limit and perhaps even into an entirely new space, where developers become responsible for vertical software services from conception through to operation and maintenance.

To adopt FaaS, software developers will need to ensure that their software follows a distributed high-decoupled service-based architecture such as microservices, and to take full advantage of the potentially significant cost savings FaaS can deliver, there are certain new software design issues that must be addressed. For example, the traditional thinking associated with chains of functions/methods executing in a stack and awaiting responses for potentially long periods will need to curtailed so as to avoid the double billing problem (where one function is executing in a waiting state, awaiting the output from a second with is running concurrently). These are not insignificant mindset changes for software developers, indeed the shift away from monolith architectures and towards distributed microservices itself poses a major challenge for many established successful firms.

There are many other changes that will result from the adoption of FaaS, the role and process adopted by software testers will change, in previous work we have already noted that this has changed significantly in recent years [50]. The space that testers have traditionally occupied between developers and operations becomes less distinct and may fundamentally reshape the role of software test. For example, the need to monitor and log FaaS systems is high and testers would at least have to re-train to develop skill in these areas. Various other changes will undoubtedly arise and will be the subject of much future research and practice.

6 Conclusion

Function-as-a-Service (FaaS) is an emerging serverless cloud computing hardware provisioning model that allows developers to essentially be completely abstracted from hardware concerns. Three separate FaaS vendors were analyzed (AWS Lambda, Google Cloud Functions and Microsoft Azure Functions) and we found that there is a complex cocktail of factors that developers should consider when selecting a FaaS provider. We recommend that careful consideration be given to the various factors, many of which are outlined in this paper. A conclusion from our research is that the trend towards increased adoption of FaaS is likely to be sustained or may even accelerate. Why? Because the speed of delivery of new software can be very rapid, concerns in relation to hardware scaling are removed, and the cost model is one where only the resources used are billed (e.g. software vendors pay for hardware only when their software is actually executing on it). These are large and disruptive forces, but caution needs to be adopted to avoid many pitfalls and FaaS is not a model that will work for all software developers.

There are various FaaS constraints to be considered, including the supported language runtimes, with significant differences to be observed in the vendors analyzed. The idea of custom runtimes means the scope of usage for AWS Lambdas is much wider than that of the other providers, and therefore is not a significant constraining factor. A further significant constraint discovered during our research was the idea of vendor lock in. Generally, migrating application functions from a non-cross vendor supported runtime can be very expensive due to requiring a complete rewrite of function logic to another language (Sect. 3.2.3). In our cost analysis our research found that the concept of double billing in the serverless model breaks a core principle of the serverless model and that without intrinsic support from the FaaS vendor, it may be impossible to avoid double billing [28]. We suggest that vendors consider the provision of a method to overcome this so they may truly achieve the serverless model. Beyond this, the cost analysis conducted by the research team concluded that Microsoft Azure and AWS Lambda offers the cheapest access to FaaS based on request invocations alone and also offers the best flat rate charge for duration of a function's execution - at \$0.000016 per GB-second (Sect. 3.3). The first real divergence outside of flat charges is found at memory provisioning with AWS Lambda giving the widest range of options [23]. With compute time, both AWS and Google Cloud round up to the nearest 100 ms, whilst Azure only rounds to the nearest 1 ms (Sect. 3.3.4).

In regard to the integration capabilities of each solution, Google Cloud provides a more streamlined approach due to the absence of integrating the function with an API, whereas both AWS and Azure require the user to either deploy or have a pre-existing API (Sect. 3.4.3.1). Generally across FaaS infrastructures, functions can be called via a HTTP request. This makes integrating functions into a pre-existing service a very simple process. These are all concerns that software developers need to be acquainted with.

Overall, our analysis of FaaS has yielded insights into the varying costs, performance, integration capabilities and constraints of each examined solution. Adopting FaaS requires significant changes to the way software is designed and deployed, and the process in these areas in particular will need to be adapted if FaaS is to be successfully and fruitfully implemented. Despite the difficult challenges to be managed in adopting FaaS, the economic imperatives driving its rise are compelling: delivery of software at a faster pace, with less impact on operational systems, and at a reduced hardware provisioning cost. FaaS in combination with well-disciplined microservices architectures may be the closest we have yet come to realizing Better and Faster and Cheaper software.

Acknowledgements. This work was supported, in part, by Science Foundation Ireland grant 13/RC/2094 and co-funded under the European Regional Development Fund through the Southern & Eastern Regional Operational Programme to Lero - the Irish Software Research Centre (www.lero.ie).

References

1. Pawlik, M., Figiela, K., Malawski, M.: Performance considerations on execution of large scale workflow applications on cloud functions. arXiv:1909.03555 [cs], September 2019
2. McGrath, G., Brenner, P.R.: Serverless computing: design, implementation, and performance. In: 2017 IEEE 37th International Conference on Distributed Computing Systems Workshops (ICDCSW), pp. 405–410 (2017). https://doi.org/10.1109/icdcsw.2017.36
3. Jackson, D., Clynch, G.: An investigation of the impact of language runtime on the performance and cost of serverless functions. In: 2018 IEEE/ACM International Conference on Utility and Cloud Computing Companion (UCC Companion), pp. 154–160 (2018). https://doi.org/10.1109/ucc-companion.2018.00050
4. Cloud Functions, Google Cloud. https://cloud.google.com/functions. Accessed 04 Mar 2020
5. Azure Functions Serverless Compute—Microsoft Azure. https://azure.microsoft.com/en-us/services/functions/. Accessed 04 Mar 2020
6. AWS Lambda – Serverless Compute - Amazon Web Services. Amazon Web Services, Inc. https://aws.amazon.com/lambda/. Accessed 04 Mar 2020
7. Hendrickson, S., et al.: Serverless computation with openLambda. In: Proceedings of the 8th USENIX Conference on Hot Topics in Cloud Computing, Denver, CO, pp. 33–39 (2016)
8. Wang, L., et al.: Peeking behind the curtains of serverless platforms. In: 2018 USENIX Annual Technical Conference (2018)
9. AWS Lambda - Developer Guide, p. 184. https://docs.aws.amazon.com/lambda/latest/dg/lambda-dg.pdf. Accessed 02 Mar 2020
10. Azure Functions HTTP triggers and bindings. https://docs.microsoft.com/en-us/azure/azure-functions/functions-bindings-http-webhook. Accessed 06 Mar 2020
11. AWS Step Functions - Developer Guide. https://docs.aws.amazon.com/step-functions/latest/dg/step-functions-dg.pdf. Accessed 01 Mar 2020
12. Using AWS Lambda with Other Services - AWS Lambda. https://docs.aws.amazon.com/lambda/latest/dg/lambda-services.html. Accessed 05 Mar 2020
13. AWS Lambda Event Source Mapping - AWS Lambda. https://docs.aws.amazon.com/lambda/latest/dg/invocation-eventsourcemapping.html. Accessed 05 Mar 2020
14. Events and Triggers—Cloud Functions Documentation. Google Cloud. https://cloud.google.com/functions/docs/concepts/events-triggers. Accessed 05 Mar 2020
15. Calling Cloud Functions—Cloud Functions Documentation—Google Cloud. https://cloud.google.com/functions/docs/calling. Accessed 28 Feb 2020
16. Event Hubs—Real-Time Data Ingestion—Microsoft Azure. https://azure.microsoft.com/en-us/services/event-hubs/. Accessed 02 Mar 2020
17. Triggers and bindings in Azure Functions. https://docs.microsoft.com/en-us/azure/azure-functions/functions-triggers-bindings. Accessed 03 Mar 2020
18. Martin Fowler: Serverless Architectures. https://martinfowler.com/articles/serverless.html. Accessed 03 Mar 2020
19. Fox, G., Ishakian, V., Muthusamy, V., Slominski, A.: Status of serverless computing and function-as-a-service (FaaS) in industry and research. arXiv:1708.08028 [cs] (2017). https://doi.org/10.13140/rg.2.2.15007.87206
20. Sewak, M., Singh, S.: Winning in the era of serverless computing and function as a service. In: 2018 3rd International Conference for Convergence in Technology (I2CT), pp. 1–5 (2018). https://doi.org/10.1109/i2ct.2018.8529465
21. García López, P., Sánchez-Artigas, M., París, G., Barcelona Pons, D., Ruiz Ollobarren, Á., Arroyo Pinto, D.: Comparison of FaaS orchestration systems. In: 2018 IEEE/ACM International Conference on Utility and Cloud Computing Companion (UCC Companion), Zurich, pp. 148–153 (2018). https://doi.org/10.1109/ucc-companion.2018.00049

22. Figiela, K., et al.: Performance evaluation of heterogeneous cloud functions. Concurr. Comput.: Pract. Exper. **30**(23), e4792 (2018). https://doi.org/10.1002/cpe.4792
23. AWS Lambda Pricing. https://aws.amazon.com/lambda/pricing/. Accessed 06 Mar 2020
24. Pricing – Functions—Microsoft Azure. https://azure.microsoft.com/en-us/pricing/details/functions/. Accessed 06 Mar 2020
25. Pricing—Cloud Functions Documentation—Google Cloud. https://cloud.google.com/functions/pricing. Accessed 06 Mar 2020
26. AWS Lambda Cost Guide. https://lumigo.io/aws-lambda-cost-guide/. Accessed 06 Mar 2020
27. The need for asynchronous FaaS call chains in serverless systems. https://read.acloud.guru/the-need-for-asynchronous-rpc-architecture-in-serverless-systems-ff168f1c8785. Accessed 06 Mar 2020
28. Composing Functions into Applications - Apache OpenWhisk - Medium. https://medium.com/openwhisk/composing-functions-into-applications-70d3200d0fac. Accessed 06 Mar 2020
29. An introduction to FaaS—a cloud computing service that makes it easier for cloud application developers to run and manage microservices applications (2020). https://www.ibm.com/cloud/learn/faas. Accessed 11 Mar 2020
30. Novkovic, N.: Top Function as a Service (Faas) Providers, Dashbird, 14 May 2018. https://dashbird.io/blog/top-function-as-a-service-faas-providers/. Accessed 11 Mar 2020
31. Edge Computing—CDN, Global Serverless Code, Distribution—AWS Lambda@Edge. Amazon Web Services, Inc. https://aws.amazon.com/lambda/edge/. Accessed 11 Mar 2020
32. Opara-Martins, J., Sahandi, R., Tian, F.: Critical analysis of vendor lock-in and its impact on cloud computing migration: a business perspective. J. Cloud Comput. **5**(1), 1–18 (2016). https://doi.org/10.1186/s13677-016-0054-z
33. Overview of serverless applications in Azure (2019). https://docs.microsoft.com/en-us/azure/architecture/serverless/. Accessed 25 Feb 2020
34. Serverless Architecture—Google Cloud. https://cloud.google.com/serverless/whitepaper/. Accessed 26 Feb 2020
35. Serverless Computing - Amazon Web Services. https://aws.amazon.com/serverless/. Accessed 28 Feb 2020
36. AWS Lambda Limits - AWS Lambda. https://docs.aws.amazon.com/lambda/latest/dg/gettingstarted-limits.html. Accessed 12 Mar 2020
37. CloudEvents. https://cloudevents.io/. Accessed 12 Mar 2020
38. Announcing first-class support for CloudEvents on Azure. https://azure.microsoft.com/es-es/blog/announcing-first-class-support-for-cloudevents-on-azure/. Accessed 12 Mar 2020
39. Amazon EventBridge - Amazon Web Services, Amazon Web Services, Inc. https://aws.amazon.com/eventbridge/. Accessed 12 Mar 2020
40. Custom AWS Lambda Runtimes - AWS Lambda. https://docs.aws.amazon.com/lambda/latest/dg/runtimes-custom.html. Accessed 12 Mar 2020
41. Amazon Managed Apache Cassandra Service - Developer Guide. https://docs.aws.amazon.com/mcs/latest/devguide/ManagedCassandraService.pdf#what-is-mcs. Accessed 03 Mar 2020
42. Amazon DocumentDB - Developer Guide. https://docs.aws.amazon.com/documentdb/latest/developerguide/developerguide.pdf#what-is. Accessed 04 Mar 2020
43. Azure Serverless—Microsoft Azure. https://azure.microsoft.com/en-us/solutions/serverless/#solutions. Accessed 28 Feb 2020
44. Clarke, P., O'Connor, R.V., Leavy, B.: A complexity theory viewpoint on the software development process and situational context. In: Proceedings of the International Conference on Software and Systems Process (ICSSP), co-located with the International Conference on Software Engineering (ICSE), pp. 86–90 (2016). https://doi.org/10.1145/2904354.2904369

45. Clarke, P., O'Connor, R.V.: The situational factors that affect the software development process: towards a comprehensive reference framework. Inf. Softw. Technol. **54**(5), 433–447 (2012)
46. Clarke, P.M., et al.: Exploring software process variation arising from differences in situational context. In: Stolfa, J., Stolfa, S., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2017. CCIS, vol. 748, pp. 29–42. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64218-5_3
47. O'Connor, R.V., Elger, P., Clarke, P.: Continuous software engineering - a microservices architecture perspective. J. Softw.: Evol. Process **29**(11), 1–12 (2017)
48. Meade, E., et al.: The changing role of the software engineer. In: Walker, A., O'Connor, R. V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 682–694. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_53
49. Garousi, V., Felderer, M., Mäntylä, M.V.: The need for multivocal literature reviews in software engineering: complementing systematic literature reviews with grey literature. In: Proceedings of the 20th International Conference on Evaluation and Assessment in Software Engineering. ACM (2016)
50. Cunningham, S., et al.: Software testing: a changing career. In: Walker, A., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 731–742. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_57
51. Garousi, V., Felderer, M., Mäntylä, M.V.: Guidelines for including grey literature and conducting multivocal literature reviews in software engineering. J. Inf. Softw. Technol. **106**, 101–121 (2019)
52. Nogueira, E., Moreira, A., Lucrédio, D., Garcia, V., Fortes, R.: Issues on developing interoperable cloud applications: definitions, concepts, approaches, requirements, characteristics and evaluation models. J. Softw. Eng. Res. Dev. **4**(1), 1–23 (2016). https://doi.org/10.1186/s40411-016-0033-6
53. Ali, H., Moawad, R., Hosni, A.A.F.: A cloud interoperability broker (CIB) for data migration in SaaS. In: 2016 IEEE International Conference on Cloud Computing and Big Data Analysis (ICCCBDA), Chengdu, pp. 250–256 (2016)
54. Clarke, P., O'Connor, R.V.: Changing situational contexts present a constant challenge to software developers. In: O'Connor, R., Umay Akkaya, M., Kemaneci, K., Yilmaz, M., Poth, A., Messnarz, R. (eds.) EuroSPI 2015. CCIS, vol. 543, pp. 100–111. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24647-5_9
55. O'Connor, R.V., Elger, P., Clarke, P.: Exploring the impact of situational context: a case study of a software development process for a microservices architecture. In: Proceedings of the International Conference on Software and Systems Process (ICSSP), co-located with the International Conference on Software Engineering (ICSE), pp. 6–10 (2016). https://doi.org/10.1145/2904354.2904368
56. Giray, G., Yilmaz, M., O'Connor, R.V., Clarke, P.M.: The impact of situational context on software process: a case study of a very small-sized company in the online advertising domain. In: Larrucea, X., Santamaria, I., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2018. CCIS, vol. 896, pp. 28–39. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-97925-0_3
57. Marks, G., O'Connor, R.V., Clarke, P.M.: The impact of situational context on the software development process – a case study of a highly innovative start-up organization. In: Mas, A., Mesquida, A., O'Connor, R.V., Rout, T., Dorling, A. (eds.) SPICE 2017. CCIS, vol. 770, pp. 455–466. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-67383-7_33
58. Clarke, P., O'Connor, R.V., Leavy, B., Yilmaz, M.: Exploring the relationship between software process adaptive capability and organisational performance. IEEE Trans. Softw. Eng. **41**(12), 1169–1183 (2015). https://doi.org/10.1109/tse.2015.2467388

Adaptive Predictive Energy Management Strategy Example for Electric Vehicle Long Distance Trip

Nikolay Pavlov[1(✉)], Boyko Gigov[1], Maria Stefanova-Pavlova[1], and Zlatina Dimitrova[2]

[1] Technical University of Sofia, Sofia, Bulgaria
{npavlov, bgigov}@tu-sofia.bg,
stefanova-pavlova@citt-global.net
[2] PSA Groupe, Vélizy-Villacoublay, France
zlatina.dimitrova@mpsa.com

Abstract. In this paper the factors that influence the energy consumption of electric vehicles are examined. The main factors affecting the driving resistance such as load, grades, vehicle speed, and additional factors are considered. For example the climate control system and the influence of ambient temperature on the electric vehicle range. The impact of the electric drive efficiency map is also taken into account. The impact of each of the factors was evaluated through a numerical study. Recommendations are given for the strategy of an adaptive predictive model for the energy management of an electric vehicle. To be planned the point of next recharge for a long distance trip, the travel conditions must be taken into account. This is done by measuring some parameters before and during the trip. The information from GPS navigation for the intended trip must also be taken into account. It could give information for road inclines and the location of the charging stations.

Keywords: Electric vehicles · Driving resistance · Efficiency map · Energy consumption · Prediction · Charging

1 Introduction

The rapid development of the automotive industry in recent years has led to environmental problems associated with the release of harmful substances contained in the exhaust gases of Internal Combustion Engines (ICE). Passenger cars and vans (‘light commercial vehicles’) are responsible for around 12% and 2.5%, respectively, of total EU emissions of carbon dioxide (CO_2), the main greenhouse gas. On 17 April 2019, the European Parliament and the Council adopted Regulation (EU) 2019/631 setting CO_2 emission performance standards for new passenger cars and for new vans in the EU [7]. This Regulation started applying on 1 January 2020, replacing the previous Regulations for cars and vans. The new Regulation will contribute to the achievement of the EU’s commitments under the Paris Agreement. It should lead to a 15% reduction of CO_2 from 2025 on and 37.5% reduction from 2030 on for new cars. It also includes

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 76–91, 2020.
https://doi.org/10.1007/978-3-030-56441-4_6

incentive mechanism for zero- and low-emission vehicles with CO_2 emissions between 0 and 50 g/km.

European carmakers are focusing on building their models on modular platforms, which are able to integrate any of the powertrains in its line-up, whether internal combustion, electric or hybrid. These modular, multi-energy and global platforms provide technological flexibility to meet customers' various expectations. They allow the industrial plants to produce internal combustion, hybrid or electric vehicles on a single production line, depending on demand.

One of the European car makers, Group PSA says that will offer clean mobility solutions in the form of all-electric zero-emissions vehicles or plug-in hybrids emitting less than 49 g of CO_2 per kilometer. Starting in 2019, all new petrol and diesel models will also systematically come in a hybrid or all-electric version [13].

Research and technology efforts for the near future:

Electric Powertrains. New technologies are geared towards customer needs and increasing autonomy, with work on:

- Future Li-ion battery technologies
- Hydrogen fuel cell technologies
- Research for the best efficiency from the entire powertrain
- Optimization of energy use, with energy optimization algorithms.

The decrease of the recharging time in line with the future recharging powers available on the fats charging terminals (150 kW, 350 kW) and the comfort of connection to recharging points (induction recharging technologies, for example) in the homes or the urban car parks.

The Rechargeable Gasoline Hybrid Electric Vehicles. For hybrid powertrains and in particular on rechargeable petrol/electric powertrains, the main working axes are:

- Reducing emissions by optimizing the engine/electrification system
- The search for "mainstream" rechargeable hybrid solutions, complementary to the premium offer arriving in 2019
- Improvement of thermal comfort in the passenger compartment (heating and cooling if the cabin).

The Recharging Ecosystem. In support of the deployment of rechargeable vehicles as Battery Electric Vehicle (BEV) and Plug in Hybrid Electric Vehicle (PHEV) - a lot of work is being done to ensure that the deployment of the recharging infrastructure is consistent, especially on:

- Smart charging solutions or smart grids to absorb the peaks of consumption generated by a growing fleet of rechargeable vehicles
- Future standards allowing the interoperability of charging solutions
- Services improving the user experience of charging.

The electrification is clear trend in the European automotive market and employs considerable efforts of the European automotive industry and academia in research and development of the main technologies enabling sustainable mobility.

Second important innovation axe is the automated driving and the connectivity of the vehicles. Connected automated driving is the opportunity to address several important societal challenges of road transport: safety, energy efficiency, congestion, urban accessibility and social inclusion [6].

2 E-Mobility Key Areas

The electrification of the vehicles connects the vehicles to the electricity grid. They have to be recharged. The vehicle energy system is considered as grid connected system. The electrified vehicles need to access to the service of recharging such as availability of the recharging points, protocol compatibility, data exchanges, on demand booking systems for the recharging slots.

The key areas for the E-Mobility and the infrastructure are identified and presented in Fig. 1.

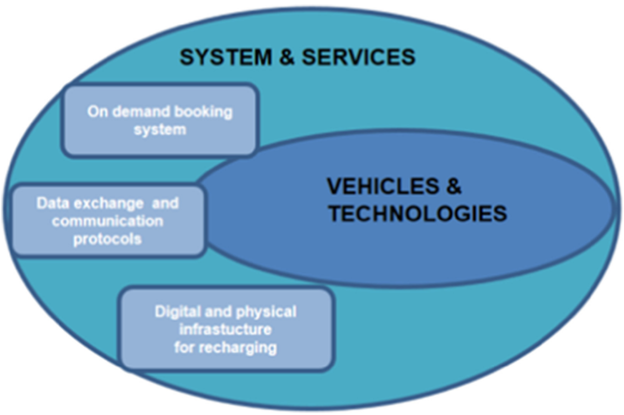


Fig. 1. Key areas E-mobility and infrastructure

Table 1 presents the key areas of the E-Mobility, applied to the recharging infrastructure and the recharging services.

Table 1. Vehicles technologies and key areas for services

Vehicles technologies	System & service
Plug-In hybrid electric vehicles	Recharging stations with connectivity layer
Electric vehicles	Data exchanges, booking of recharging slots to the stations
Fuel cell electric vehicles	Hydrogen infrastructure planning management
V2X connectivity: V2V and V2X	Connectivity to the recharging stations & Data exchange of their status

Other question is the E-Mobility for long ranges, for long distances trips. Special model needs to be developed for the long-range mobility [4]. The recharging infrastructure has an important investment cost. There is a techno-economic trade-off between the density of location of the charging stations and their economic sustainability. This research question has to be modelled and studied to find optimal location and usage of the recharging stations.

3 Recharging Technologies and Infrastructure

The recharging systems are classified according to their power. Figure 2 defines the charging systems to slow and fast charging systems. The power determines the time to charging of the battery. The time and the occupation rates of the chargers are as well evaluated. The optimality is researched for long way drives on electric mode. The results illustrate a method to rethink the usage and the energy management on the vehicle during the driving. The target is to propose a high quality service for the electric mobility for long way drives.

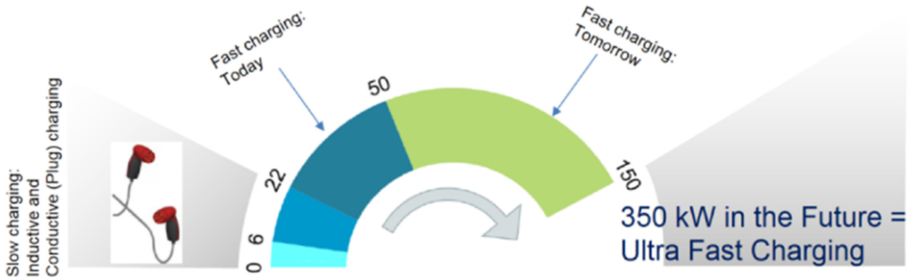


Fig. 2. Definition of the recharging systems

In an urban environment, electric cars have indisputable advantages. One is the lack of local pollution in the city. The other is related to the ability of electric vehicles to regenerate kinetic energy during braking. This leads to significant energy savings. For long journeys, the car moves at a steady speed and cannot take advantage of regenerative braking. The speed is usually high. Therefore, the air resistance is considerable and hence the energy consumption. The driving modes, as well as other factors, influence the current energy consumption. In a poorly developed network of charging stations, the battery of an electric vehicle may be exhausted before reaching the next charging station. Therefore, it is necessary to take into account the route and the operational factors in advance.

4 Factors Affecting Energy Consumption and Range of Electric Vehicles

There are a number of factors affecting the energy consumption of an electric vehicle. When driving an electric vehicle, resistance forces occur on the road: rolling resistance, grade resistance, aerodynamic drag, and inertial resistance. To overcome these resistances, the stored energy in the battery is used up. Furthermore, electric drive is characterized by different efficiency for different operating modes, which is illustrated by the so-called efficiency map. These maps are obtained by measurement on test bench and usually represent the conversion efficiency as a function of the load and the speed [3]. Some of the factors affecting the resistance are conditionally determined, such as the frontal area, aerodynamic drag, constructive weight, etc.

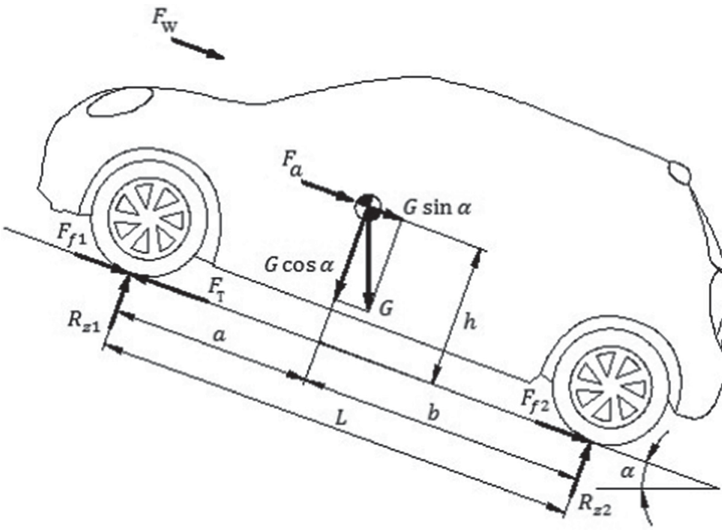


Fig. 3. Forces acting on a vehicle travelling uphill

Other factors are operational - payload, road gradients, speed, and more. Also, the correct value of the tire internal pressure affects the rolling resistance [12]. Figure 3 shows the forces acting on a vehicle in the general case. The influence of various operating factors on energy consumption and electric vehicle range are analysed below.

4.1 Payload

The total mass of the electric vehicle, and therefore the payload, influence the inertial resistance F_a during acceleration. In the long range aspect, the influence of inertial resistance can be neglected, since for most of the time the electric vehicle is moving at a constant steady speed. The effect of mass on its steady state motion can be analysed by the following equation:

$$F_\psi = F_f + F_i = \psi \cdot G \quad (1)$$

Where F_ψ is a road resistance, F_f is a rolling resistance, F_i is a gradient resistance (downhill force), $G = m \cdot g$ is the vehicle weight, $\psi = f + i \cdot 10^{-2}$ is the road resistance coefficient (for small grades cosα ≈ 1 and sinα ≈ tgα = i/100), f is the rolling resistance coefficient, i is a road grade in %.

It is used to determine the total resistance in a steady motion of an electric vehicle:

$$F_r = F_\psi + F_w \quad (2)$$

Where $F_w = 0.5 \cdot C_x \cdot \rho_в \cdot A \cdot v^2$ is an aerodynamic drag, C_x is the drag coefficient, $\rho_в$ is the air density, A is the frontal area, v is the vehicle speed.

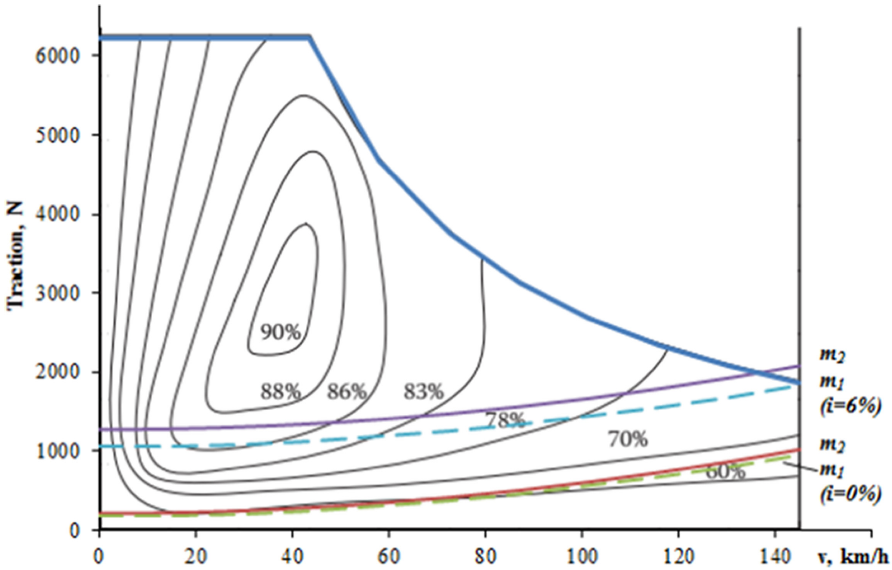


Fig. 4. Effect of the payload on the required tractive effort

Figure 4 shows the effect of payload on the required tractive effort F_T of an electric vehicle. For this purpose, calculations were carried out on an electric vehicle with driver-only mass m_1 = 1500 kg and full-load mass (including driver, four passengers and their luggage) – m_2 = 1800 kg. As can be seen the influence of the mass in a steady motion and a horizontal road (i = 0%) is not great. However, as the slope increases (i = 6%), the influence increases and it is not negligible.

4.2 Grading Resistance and Influence of Speed

For the most part of the length of the routes, the roads are not horizontal and occur positive or negative grades. There is a requirement when designing long-distance roads for a maximum grade of up to 9% [10]. However, such sections are rarely encountered.

Maximum gradients often are up to 6% for long-distance roads and up to 4% for highways. Using Formula (2), the calculations of the required tractive effort at different grades for climbing and descending in the range of 0 to the maximum speed of the electric vehicle have been carried out. The calculations were performed in full mass m_2 = 1800 kg. The results are presented on the traction diagram of an electric vehicle, which also include an efficiency map typical of a permanent-magnet electric drive electric motor (Fig. 5).

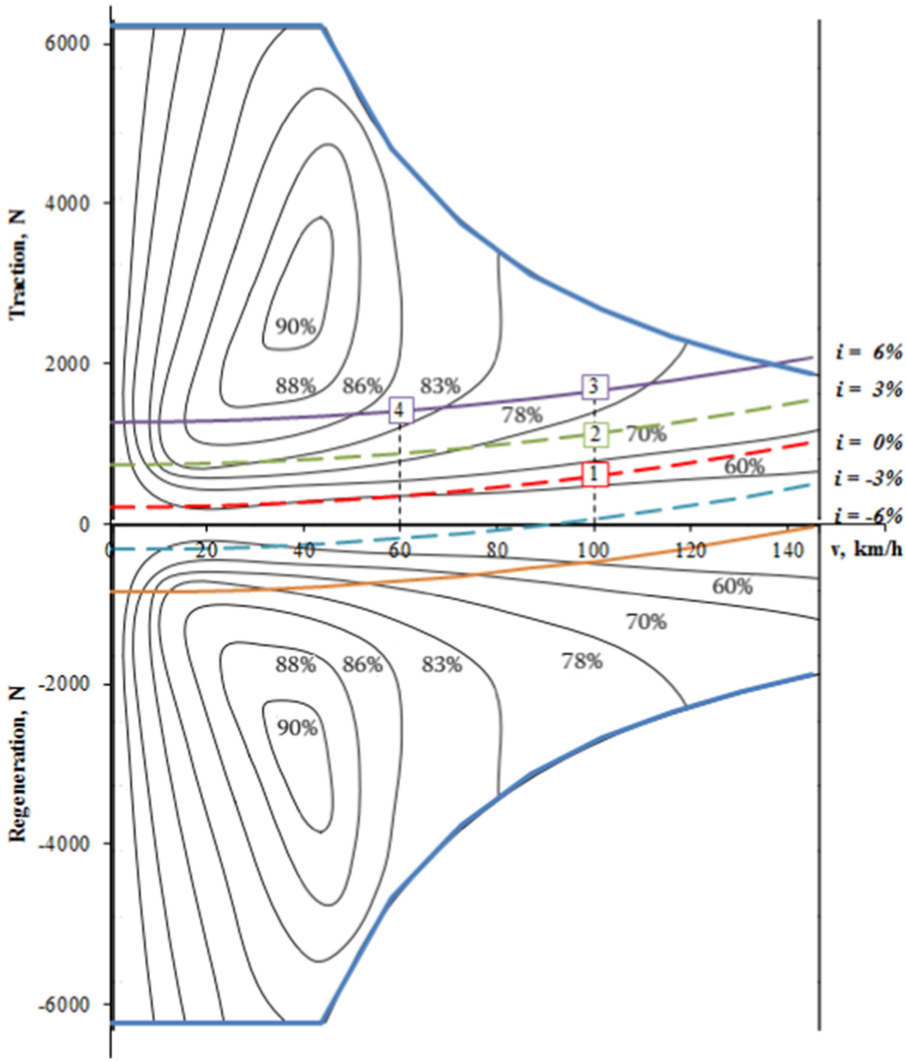


Fig. 5. Effect of the grades and speed on the required tractive effort

In this way, both the required tractive effort and the energy efficiency of the electric drive can be taken into account. The traction mode is shown above the horizontal axis and the energy regeneration mode below it. In steady state motion, this mode can be achieved when driving downhill. As the speed of the electric vehicle increases, the aerodynamic drag (with the square of the speed) increases and the regenerated energy decreases. In the Fig. 5 is shown an example of changing the efficiency of the electric drive when driving at a constant speed of 100 km/h and overcoming different grades - operating points 1, 2 and 3. It is seen that with increasing the slope i from 0% to 3% and then to 6% the required tractive effort increases. But each successive point falls into an area of higher efficiency than the previous one in order of 10%. Thus, the need for more electric energy to overcome the grade is to some extent offset by the increasing efficiency of the electric drive. Working point 4 shows how, as speed decreases, the required tractive effort also decreases, but at the same time the efficiency of the electric drive is again increased and the effect of the reduction of speed when overcoming large grades is enhanced. Each electric drive has its own specific efficiency map, which must be taken into account when calculating and predicting the energy consumption of the electric vehicle during its operation. The example shows that this electric drive is more suitable for urban traffic. With higher accelerations and regenerative braking, the operating points will fit into the higher efficiency zones.

4.3 Cabin Microclimate Energy Consumption

The continental climate, which is found in much of Europe, Asia and North America, is characterized by a significant annual temperature range - hot summers and cold winters. In order to maintain the necessary microclimate for the driver and passengers, it is necessary to heat the cabin during the cold part and cool it during the warm part of the year. Different amounts of electricity are used depending on the ambient temperature. At ambient temperatures around 20 °C, only ventilation is required. As the temperature rises, the energy required for cooling increases. At low temperatures, a large amount of electric energy is consumed for both the interior heating and the defrosting of the windows and rear-view mirrors. Unlike conventional cars, where the heat generated by the ICE can be used to heat the cabin, in electric cars the energy must come from the battery. At extremely low ambient temperatures, the required energy can reach up to 3.5 kW and this seriously affects the electric vehicle range [8, 9]. An example of the energy, consumed for heating and ventilation depending on external temperature is shown in Fig. 6.

4.4 Effect of Temperature on Battery Capacity

The temperature at which the battery is discharged has a pronounced effect on its capacity. This is due to the decrease in chemical activity and the increase in internal resistance of the battery at lower temperatures [14]. Figure 7 illustrates this variation for a single lithium-ion cell [1].

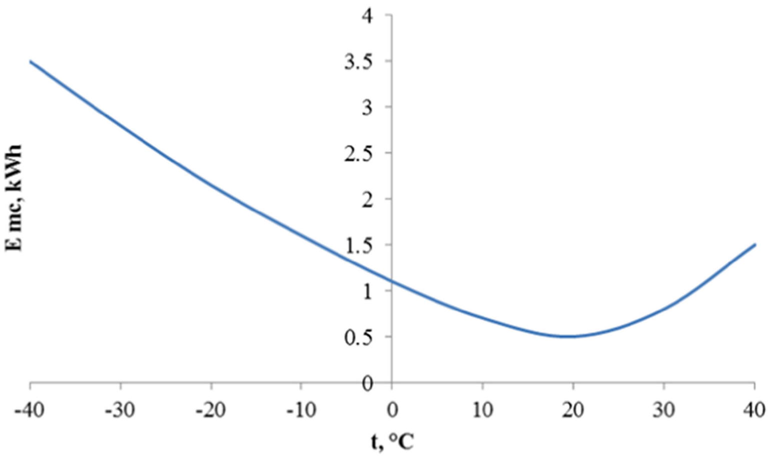


Fig. 6. Example for cooling/heating energy consumption depending on temperature

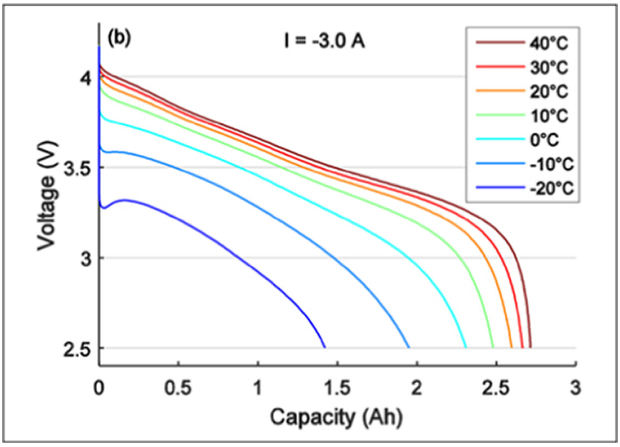


Fig. 7. Discharge voltage of a Li-ion cell at 3 A and various temperatures [1]

5 Evaluation of the Impact of Individual Factors Affecting Electric Vehicle Long Distance Trip Range

To illustrate the degree of influence of each of the above factors on the energy balance of an electric vehicle and its range with a single charge of the battery, it is possible to calculate the total power required to overcome the various resistances (P_{Σ}) and the power for the additional consumers (P_0). Based on the calculated power and the battery energy, it is possible to be estimated the range of the electric vehicle.

As shown above, the most factors are interdependent, and changing one of them leads to a change in the other factors, which can have the opposite effect on power. To assess the degree of influence of only one factor, all others are at the baseline conditionally fixed. The following baseline values have been adopted (Table 2):

Table 2. Initial values of the factors

f_0	C_X	Air dens., kg/m^3	G, N	Road grade, %	Motor eff., %	Ambient t, °C	v, km/h	E_{batt}, kWh	a, m/s^2	P_0, kW
0.018	0.31	1.226	10465	0	59	20	70	39.6	0	1

In the calculations a possible change of each factor by 10% in the direction of increase or decrease according to reality is set. For the acceleration and incline of the road is assumed to be an increase of about 10%, but based on the most probable real average values, since the initial values are zero (valid for established moving mode and horizontal road). In addition, the following specificities for each of the factors are considered:

- The rolling resistance coefficient is calculated using the empirical formula: $f = f_0 \cdot (1 + k_v \cdot v^2)$, to take into account the influence of speed over 60 km/h;
- For the drag coefficient (C_x) are accepted the minimum values valid for motion in a stationary air environment when the aerodynamic drag is directed exactly in the direction of travel (special case);
- The initial value for air density is assumed to be at sea level, with the possibility of decreasing with increasing altitude and temperature;
- The initial weight value is formed at an optimistic chassis weight, the battery weight and the driver with the possibility to change the number of passengers;
- The efficiency of the electric motor is accounted from the efficiency map for the respective operating mode with the possibility of increasing when the speed and/or the load increases;
- The power electronics and transmission efficiency coefficients change much less than the electric motor efficiency and they are therefore constant considered [2].
- It is accepted that the ambient temperature most likely to change in the range −20 to 40 °C. The dependence of the energy, required to heat or cool the passenger compartment, from the ambient temperature is approximated and analytically presented with a third degree polynomial. Changing the temperature by only 10% would practically not have a significant impact, so changes in both directions up to the limit values were adopted;
- For the speed, average operating value with the possibility of increasing is accepted. Note that its increase influences in the opposite direction the power consumed by the battery and the range. With the same available energy, an increase in power leads to a reduction in range by reducing travel time. But the increase in speed compensated to some extent this reduction;
- The initially battery energy is calculated for operating voltage of 360 V and capacity of 110 Ah.

- This work assumes a motorway travel with a relatively constant speed and limited road gradients. In the calculation of the inertial force, a coefficient of conditional mass increase δ = 1.01 was adopted, which takes into account the influence of the rotating parts. When the load on the motor increases, whether due to acceleration or other factors, the increase in efficiency is not accounted for due to its entering the optimum zone, which would have a compensating effect on the power;
- It assumed that the battery, motor and control electronics operate in the optimum temperature area by cooling with appropriate systems. The power required for this, as well as for other additional consumers, taken in the order of 1 kW.

The results of this numerical study are illustrated in Fig. 8 and in Table 3.

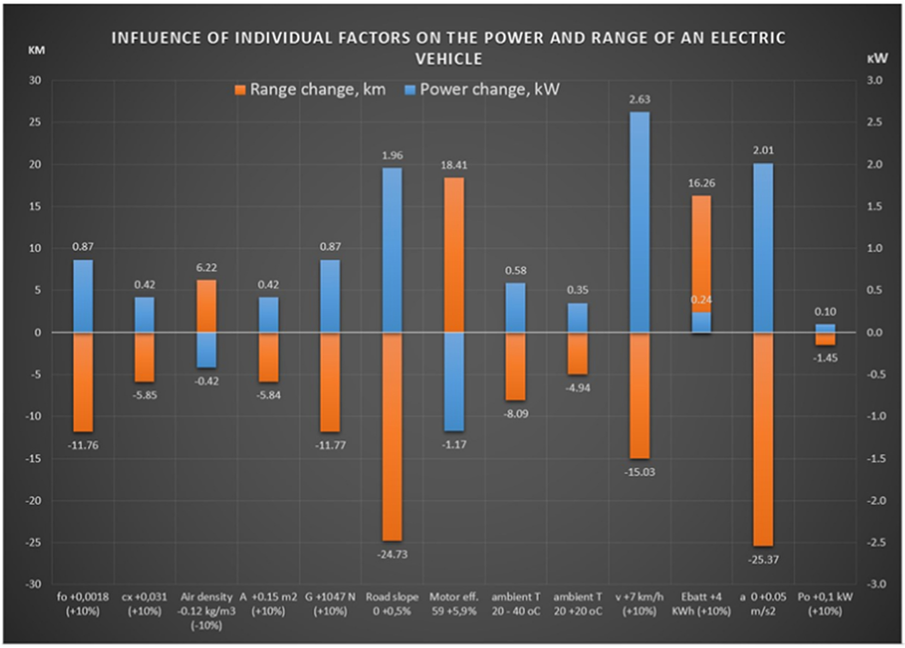


Fig. 8. Numerical study of the factors, affecting the power consumption and travel range of an electric vehicle

Table 3. Change in factors, power and range of an electric vehicle in absolute and relative units

	Δ	Δ, %	Δ*P*, kW	Δ*P*, %	Δ*R*, km	Δ*R*, %
f_0	0.0018	10.00	0.87	6.3	−11.76	−5.9
C_X	0.031	10.00	0.42	3.0	−5.85	−2.9
Air density	−0.1226	−10.00	−0.42	−3.0	6.22	3.1
A	0.1556	10.00	0.42	3.0	−5.84	−2.9
G	1047	10.00	0.87	6.3	−11.77	−5.9
Road grade	0.5		1.96	14.1	−24.73	−12.4

(*continued*)

Table 3. (*continued*)

	Δ	Δ, %	ΔP, kW	ΔP, %	ΔR, km	ΔR, %
Motor eff	5.9	10.00	−1.17	−8.4	18.41	9.2
Amb. t	−40.0	−200.00	0.58	4.2	−8.09	−4.0
Amb. t	20.0	100.00	0.35	2.5	−4.94	−2.5
v	7.0	10.00	2.63	18.9	−15.03	−7.5
E_{batt}	4.0	10.00	0.24	1.7	16.26	8.1
a	0.05		2.01	14.5	−25.37	−12.7
P_0	0.1	10.00	0.10	0.7	−1.45	−0.7

The analysis of the results of the numerical study shows that in most cases a change in a factor causes a change in range because of a change in power consumption. The increase in power causes a proportional reduction in range and vice versa in almost the same percentage.

The effect of the speed is an exception, as its increase of 10% causes an increase of power of about 19% and a decrease in range by only 7.5%, which is due to the speed influence on the range, explained above. Another exception is the impact of the increase in battery capacity, which should not effect on the power consumed, but only on the range. Since the increase in the mass of the battery is also into account taken in this case, is obtained a relatively small increase in power consumption (by 1.7%), which reduces the increase in range.

The change in the coefficient f_0 causes practically the same effect on the power and range as the change in weight, since it evaluated, when traveling on a horizontal road without slope and at a relatively low speed. However, if the slope increased by only 1%, the impact of the change in weight increases, with the power consumed by the increase in weight increasing by 1.4 kW and the range decreasing by about 3 km compared to the changes caused by the increase of the coefficient f_0. This follows from the equation, by which the resistance force from the road calculated. It can expected, that the increase in speed will increase the influence of the coefficient f.

The influence of the ambient temperature is significantly less and incomparable with other influences. It can cause variations in power and range up to about 4% at the temperature change of 200% relative to the accepted initial value of 20 °C. The power of the additional consumers can also be related to the slightly influencing factors on the range.

The change in the C_x coefficient has the same effect on the parameters studied, as the frontal area (A) and the air density. Under these initial conditions, this influence is relatively small (in the order of 3%) when the factors change by 10%. At higher speeds, this influence can expected to increase significantly. In addition, in the presence of wind direction in a general position - different influence of the three factors, that can hardly be estimated without experiment [5].

The most significant is the influence on the power and range of the road slope and the acceleration. The conditional increase of 10% relatively the most likely average values leads to change of the power and of the range in the range of 12 to 14% and an increase in the efficiency of the electric motor by 10% leads to an increase in range of about 9%, which is also not to be ignored.

6 Adaptive Predictive Energy Management Model

Based on the analysis of the factors affecting energy efficiency, a model can be created that predicted energy consumption for the planned trip (Fig. 9). For high-accuracy predictions, it is necessary to determine the load on the electric vehicle at the beginning of the trip (point A). This can be done by taking information from the sensors for the presence of passengers, and if they determine the weight, the accuracy will be improved. Such sensors can also be used in the car trunk to determine the weight of the luggage. Satellite navigation information can be used to determine the road profile and expected resistances. However, the end point of the trip (B) must be entered for this purpose. Information about the current driving mode and its efficiency is got from the electric vehicle speed sensor, the accelerator pedal sensor, the current sensor and the efficiency map entered into the software.

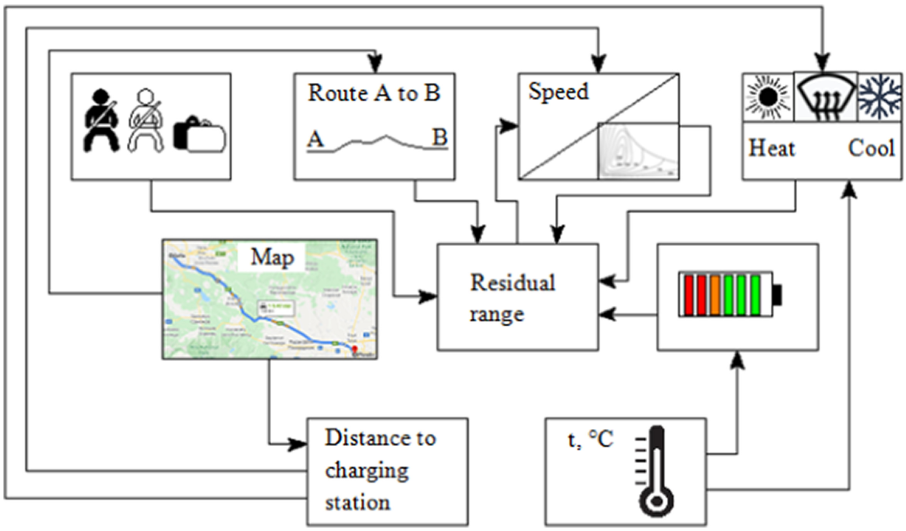


Fig. 9. Adaptive predictive energy management model

Depending on the ambient temperature, it can be predicted how much energy will be required to heat or cool the passenger compartment for the entire trip. Based on the ambient temperature and taking into account the temperature characteristic of the battery (Fig. 7), the expected range of the electric vehicle can be predicted. In order to determine the current expected range, it is also necessary to take into account the current amount of energy in the battery during the trip. To determine on which charging station it is best to charge the battery, information is needed to their location. Such information may come also from navigation or via connected vehicle to infrastructure [11]. In case the current mode of motion consumes more energy than is allowed for the respective section and the model recognizes that there is a danger of exhausting the battery energy before reaching the charging station, an automatic correction in the driving mode should be made. It results in a reduction in speed so as to reduce the tractive effort and the required power. In this case, the operating point may fall into more or less efficient electric driving mode. This should be accounted by the software and to correct the calculations for the residual range. The heating or cooling power of the passenger compartment can be further reduced in order to save energy.

7 Conclusion

This paper has presented the influence of different factors that affect the electric vehicle long distance trip range. This is important because there is still no well-developed network of charging stations. There is a risk of running out of battery before reaching a charging station. The risk can be minimized by taking into account the impact of different operating factors and road conditions. By taking into account some factors, such as vehicle mass, road profile, additional energy consumption, etc., it is possible to predict what the residual range will be. The GPS system may receive information on the location of the charging stations. Depending on the received information, driving modes can be modified by the software to extend the residual range and to reach the charging station. Also, the proposed model can be used as a resource tool in the training of electric powertrain engineers and they will be able to estimate the impact of various factors on the range of electric vehicles.

8 Relevance to SPI Manifesto

The most recent interpretation of the SPI term is System, Software, Services, Safety, and Security Process and Product Improvement, Innovation, and Infrastructure [15]. EuroAsiaSPI has developed as experience and knowledge exchange platform for Europe and worldwide, where Software Process Improvement (SPI) practices and knowledge can be gathered and shared. The connected SPI Manifesto [16] defines the required values and principles for most efficient SPI research and industry projects and is a Strategy Paper for the further development on the base of those principles. One main goal is to support changes by innovation and include all stakeholders. There are two important principles, which supports the whole development process namely “SPI is inherently linked with change” means that change is part of the normal development process which leads to new business opportunities [12]. A current major change in the automotive industry affects the mobility concepts and the presented paper gives new insights into the energy management for electric vehicles. “Create a learning organization” is important principle for exchange of best practices on the researched topics [17]. A further principle is proposed “Use dynamic and adaptable models as needed”, which supports also the effectivity of the development process.

Acknowledgements. This work is supported by the ECEPE project. The ECQA Certified Electric Powertrain Engineer project (ECEPE) is co-funded by the Erasmus+ Call 2019 Round 1 KA203 Programme of the European Union under the agreement 2019-1-CZ01-KA203-061430.

References

1. Battery University, BU-502: Discharging at High and Low Temperatures. https://batteryuniversity.com/learn/article/discharging_at_high_and_low_temperatures. Accessed 06 Apr 2020
2. Břoušek, J., Zvolský, T.: Experimental study of electric vehicle gearbox efficiency. In: Gigov, B., Nikolov, N., Stoilova, S., Kralov, I., Todorov, M., Stoilov, V. (eds.) BulTrans-2018, MATEC Web of Conferences, vol. 234, p. 02004 (2018). https://doi.org/10.1051/matecconf/2018234004
3. Dimitrova, Z.: Vehicle propulsion systems design methods. In: Gigov, B., Nikolov, N., Stoilov, V., Todorov, M. (eds.) BulTrans-2017, MATEC Web of Conferences, vol. 133, p. 02001 (2017). https://doi.org/10.1051/matecconf/201713302001
4. Dimitrova, Z.: Optimal designs of electric vehicles for long-range mobility. In: Gigov, B., Nikolov, N., Stoilova, S., Kralov, I., Todorov, M., Stoilov, V. (eds.) BulTrans-2018, MATEC Web of Conferences, vol. 234, p. 02001 (2018). https://doi.org/10.1051/matecconf/201823402001
5. Dobrev, I., Massouh, F., Danlos, A., Todorov, M., Punov, P.: Experimental and numerical study of the flow field around a small car. In: Gigov, B., Nikolov, N., Stoilov, V., Todorov, M. (eds.) BulTrans-2017, MATEC Web of Conferences, vol. 133, p. 02004 (2017). https://doi.org/10.1051/matecconf/201713302004
6. ERTRAC, European Road Map, Electrification of Road Transport, 3rd edn. (2017). https://www.ertrac.org/uploads/documentsearch/id50/ERTRAC_ElectrificationRoadmap2017.pdf. Accessed 10 Apr 2020
7. European Commission. CO2 emission performance standards for cars and vans (2020 onwards). https://ec.europa.eu/clima/policies/transport/vehicles/regulation_en. Accessed 10 Apr 2020
8. Evtimov, I., Ivanov, R., Stanchev, H., Kadikyanov, G., Staneva, G., Sapundzhiev, M.: Energy efficiency and ecological impact of the vehicles. In: Sładkowski, A. (ed.) Ecology in Transport: Problems and Solutions. LNNS, vol. 124, pp. 169–250. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-42323-0_4
9. Farrington, R., Rugh, J.: Impact of Vehicle Air Conditioning on Fuel Economy, Tailpipe Emissions, and Electric Vehicle Range. Earth Technologies Forum Washington, D.C, 31 October 2000
10. https://dv.parliament.bg/DVWeb/showMaterialDV.jsp?idMat=129951. Accessed 08 Apr 2020
11. Macher, G., Armengaud, E., Schneider, D., Brenner, E., Kreiner, C.: Towards dependability engineering of cooperative automotive cyber-physical systems. In: Stolfa, J., Stolfa, S., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2017. CCIS, vol. 748, pp. 205–215. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64218-5_16
12. Messnarz, R., Ekert, D., Grunert, F., Blume, A.: Cross-cutting approach to integrate functional and material design in a system architectural design – example of an electric powertrain. In: Walker, A., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 322–338. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_25
13. PSA GROUPE. https://www.groupe-psa.com/en/automotive-group/innovation. Accessed 10 Apr 2020
14. Reddy, T.: Linden's Handbook of Batteries, 4th edn. McGraw-Hill Education, New York (2010)
15. https://2020.eurospi.net/index.php/manifesto. Accessed 02 Apr 2020

16. The SPI Manifesto: EuroSPI 2009, Alcala (2009). https://2019.eurospi.net/images/eurospi/spi_manifesto.pdf
17. Korsaa, M., et al.: The people aspects in modern process improvement management approaches. J. Softw.: Evol. Process **25**(4), 381–391 (2013)

SPI and Standards and Safety and Security Norms

Integrating Approaches in Software Development: A Case Analysis in a Small Software Company

Mary Sánchez-Gordón[1](✉), Ricardo Colomo-Palacios[1], Alex Sánchez[2], and Sandra Sanchez-Gordon[3]

[1] Østfold University College, Halden, Norway
{mary.sanchez-gordon, ricardo.colomo-palacios}@hiof.no
[2] LogicStudio Panamá, Panama City, Panamá
alex.sanchez@logicstudio.net
[3] Escuela Politécnica Nacional Quito, Quito, Ecuador
sandra.sanchez@epn.edu.ec

Abstract. There are a myriad of software development methods, methodologies, frameworks, techniques and practices in both traditional and agile software development. Seeking synergy between these approaches has become necessary for the evolution of a software development process. Software companies deal with that challenge by combining well-structured comprehensive methods and flexible agile practices. In fact, some studies have revealed that mixed approaches in software industry are not uncommon. This paper analyzes a case study of the evolution of a software development process in a small company, which is based on Design Thinking, PMBOK and SCRUM. Results show the natural evolution of the software process over 15 years and lessons learned from a pragmatic process selection. The evolution in the company depends on its adaptability which captures the capacity of the company to learn, combine experience and knowledge, and adjust its software process to a changing context. Despite the results are promising further studies should be done.

Keywords: Agile software development · Case study · Lessons learned

1 Introduction

Software practitioners, especially in small companies because of their nature [1], face a major challenge in shaping the many available methods, methodologies, frameworks and techniques [2, 3]. Even in the software process improvement field, there is a lot of diversity [1]. In fact, there is no silver-bullet or one-size-fits-all solution to all software development settings [3, 4]. Although, the literature has reported Waterfall as one of the most popular traditional (non-agile) approaches [4, 5], there are a large number of them, e.g. prototyping, spiral model and unified process. Additionally, Scrum is a well-known and popular agile approach [4, 5] but there are also a large number of them, e.g. Scrum, eXtreme Programming, Feature Driven Development and Crystal. Each approach, whether traditional or agile, is characterized by an individual specific structure that reflects the particular point of view and experiences of who created it [2].

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 95–106, 2020.
https://doi.org/10.1007/978-3-030-56441-4_7

Given that each approach also has its own style and terminology to describe its selected practices, sometimes it is hard to distinguish common practices.

In this scenario, integrating agile and traditional approaches has become necessary for the evolution of a software development process [6]. It is revealed in the growing popularity of the mixed (hybrid) approaches in software industry [4, 7] while it is gaining increasing attention from the software engineering research community [2, 5, 8, 9]. Hybrid development is a fruitful research field [2, 8] in which there is a lack of evidence concerning combination patterns and contextual factors that drive the creation of hybrid approaches [9]. By conducting a case analysis of a software development process in a small software company, this exploratory study aims to better understand the historical evolution. The findings confirm that many variations of software development approaches could occur, even within a small company.

2 Study Context

2.1 Company Background

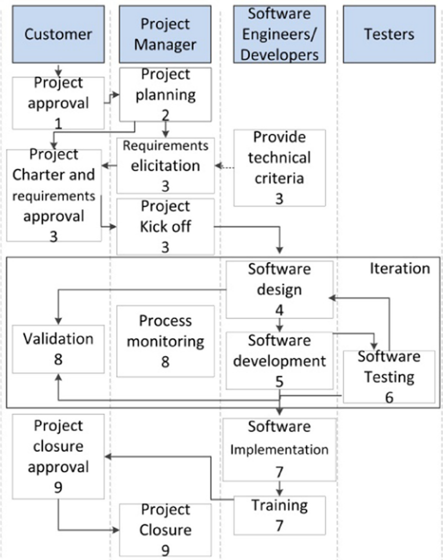
Logic Studio is a recognized Latin American outsourcing and software company that provides solutions to corporations, banks, and the public sector taking advantage of web and mobile technologies in order to create innovative services. Founded in 2003, Logic Studio has more than 100 successful projects in 7 countries. Logic Studio, with 90 employees, has branches in 2 countries, although its headquarters is located in Panama. The headquarters made around 1.8 million dollars in 2018 and 1.9 million dollars in 2019. Moreover, Logic Studio has commercial representatives in Lima-Peru and Florida-USA. Logic Studio is also a Microsoft Gold Certified Partner, whose experts have been awarded with the Microsoft Most Valuable Professional (MVP) and Microsoft Regional Director (RD) awards.

2.2 High-Level Software Development Lifecycle and Product Perspective

Logic Studio aims to empower their customers in order to significantly increase the success of their software development projects, while complying with the best practices on project management, software engineering, and agile development. To do so, Logic Studio manages the software development life cycle (SDLC) using an approach based on 1) the understanding of customer needs through Design Thinking (DT) [10]; 2) the best practices of the Project Management Body of Knowledge (PMBOK) [11]; and 3) SCRUM [12] for the agile development of software. The main goals are: 1) Delivery of products that meet company regulations and customer requirements on time and within budget; 2) Alignment of the vision of both customer and development team; 3) Motivated and involved team members to achieve better results; and 4) High productivity and quality.

Figure 1(a) shows a high-level view of the SDLC in the company. The DT approach disclosed by d.school was chosen to address a better understanding of the customer need through a systematic exploration at the initiation phase before project approval. The aim is to understand the right product to develop. This approach encompasses five phases: Empathize, Define, Ideate, Prototype and Test.

(a) High-level Software Development Lifecycle



(b) Product Perspective

Products: Phase	Products: Name	Product Owner	Project Mngmnt	Scrum Master	Developer	Tester
1. Initiation	Sales	A	R	S		
2. Design	Project Charter	R	S	R	S	S
	Work Breakdown Structure	C	S	R	S	I
	Schedule			R		
	Risk matrix	C	S	R	S	I
3. Requirements	Software Requirements Specification	R	S	R	S	S
4. Design	Design	C	S	R	S	I
	Data dictionary			R		
	Architecture	C	S	R	S	I
5. Development	Source code			R	R	
6. Test	Test	C/I	C/I	C/I	C/I	R
7. Delivery	Installation manual*	A		R	R	
	Certificate of installation*	A	R	R	A	
	User manual	A		R		S
8. Monitoring	Meeting Record	I	R			
	Progress Status Record	I	R	S	S	S
	Change Request	A	R	C		
9. Close	Acceptance Record	A	R			
	Lessons learned	C	R	C/I	C/I	C/I

Fig. 1. High-level SDLC (a); product perspective on the SDLC (b).

The project management process is based on five main categories of PMBOK: Initiating, Planning, Executing, Monitoring and Controlling, and Closing. The processes are often iterated prior to completing the project and can have interactions within a Process Group and among Process Groups. Using the context of the software project, i.e. the specific project characteristics and environment, the project team seeks to understand the constraints in which the team should focus their efforts. The relationship among factors such as scope, quality, schedule, budget, resources and risks is such that if any factor changes, at least one of the other factors might be affected. Therefore, the project team needs to be able to assess every situation, balance demands and maintain proactive communication with stakeholders in order to deliver a successful project. Such project should be what stakeholders are expecting so they can find the project acceptable.

On the other hand, Scrum includes an iterative, development approach with early deliverables and well-defined responsibilities that promotes transparency, inspection and adaptation. A sprint is a 2–4 weeks period of development time. Moreover, the responsibilities of the traditional role of project manager are divided and complemented between three roles: Product Owner, Scrum Master and development team. In the last one, analysts, designers and testers are involved. The “look and feel” design is done by a graphic designer, if required, and validated by the customer in each delivery. The main focus is to deliver value continuously in a “time boxed” manner while the team members work together and not individually to build the product. The processes are ISO-9001 certified and the engineering activities are operating at CMMi Level 3.

From the product perspective, there is a set of expected work products related to each SDLC phase. Figure 1(b) shows the relationships between products and the actions of each role—Responsible (R), Approve (A), Read (C), Informed (I), and Support (S). The specific project management approach is selected according to the nature and duration of the project as well as the amount of the project budget.

3 Research Method

A single-case study was carried out to explore the natural evolution of software process in a small software company. Hence, we used qualitative data regarding the (general) process use and experiences in order to gain an understanding of the underlying motivation, opinions, and practice of software engineers when tailoring the software process. In what follows a brief description of the procedure is described.

Documentation Analysis. Two authors were granted access to the documentation, related to both the software development processes and/or work products (December 2018). The preliminary documentation analysis was used as a basis to define the study context. Moreover, the documentation analysis was important when validating research and analyzing the content or reasoning contained within a document.

Survey. Two authors conducted the survey (January 2019) with project team members identified by the Chief Executive Officer (CEO). The questionnaire is based on the instrument developed by HELENA study [5]. The survey was anonymous and participation was voluntary. However one reminder was sent to each potential participant. The data was gathered from 17 projects in order to better understand the software development process in the company. Only one of the responses was incomplete—as respondents were given the option to skip questions. Specifically, three questions were not answered. The HELENA is an international exploratory multistage survey-based study on the use of "**H**ybrid d**E**ve**L**opm**EN**t **A**pproaches in software systems development" [13] that was launched in 2016. Two of three stages have been conducted globally in more than 25 countries and involved about 75 researchers [5]. The questionnaire comprises 38 questions aimed at collecting data on general process use, process use in the context of norms and standards, process improvement, and experiences. Results revealed that hybrid development approaches in software system development are a reality that affects companies regardless of size and industry sector.

Interview. After a first analysis of the survey results, two authors conducted the interview with the CEO/Funder (beginning February 2019) in order to understand better the results of the survey and the evolution of the software process in the company. The interview allowed more in-depth discussions and targeted questions about the previous results. The interview (semi-structured and in-depth) lasted approximately 45 min and it was voice recorded and transcribed verbatim. Finally, the transcript was sent to the interviewee in order to correct misunderstandings.

Data Analysis. Two authors constantly analyzed, coded and reviewed the transcript in the light of data collected from the previous survey and documentation analysis. To do so, the constant comparative method used within the Grounded Theory (GT) method was employed to evaluate new data. Thus, the iterative analysis of the data from the interview was augmented by the documentation analysis. It was achieved by constantly referring to the information provided in the intranet for checks and validation. The findings were also consolidated by using constant comparative method. The author not involved in the data analysis was tasked to provide the quality assurance.

4 Results and Lessons Learned

4.1 Status Report

Respondents informed that, they work in all size projects, but the majority of them are medium (2–6 person months). The projects are running in some industry sectors: "Financial Services", "Cloud Applications and Service", "Web Applications and Services", "Mobile Applications" and "Public Sector/Public Contracting". Almost all respondents recognized that a software failure conceivably can impact the company's reputation and company's business. About half of the respondents also believe that such a failure can lead to system (service) degradation. Similarly, legal consequences (civil law) and financial loss were identified in 3 and 2 projects, respectively. It is not surprising that respondents informed all projects of the company are operated according to the same (potentially customized) standard process. Although the project-specific development approach was mostly defined by a project manager who tailors the process in the beginning of a project, respondents recognized others ways as well. Sometimes, the process follows defined rules or specific practices and methods that are selected in the project according to customer demands. However, two less experienced (1–2 years) respondents pointed out that specific practices and methods are selected in the project on demand and the process is not tailored at all.

An overview of the use of software development approaches is shown in Fig. 2(a). It is based on the respondents' rating in four categories: we rarely use it, we sometimes use it, we often use it, and we always use it. Other categories such as "do not know a specific framework or method" and "do not use it" are not depicted. As it was expected Scrum is on the top, followed by Kanban, Classic Waterfall, Iterative Development and DevOps while HELENA study [5] for the combined process use in the entire (non-filtered) dataset shows Scrum followed by Iterative Development, Kanban, Classic Waterfall and DevOps. Therefore, there is no surprise in top-5. A likely explanation for the use of Waterfall in LogicStudio is in the outsourcing requirements. Regarding PMP, it was not mentioned at all. Although PMP is part of the homegrown process, the respondents would probably be not aware of it. Finally, DT was incorporated as part of the homegrown process and its use is limited at the time.

Despite that, previous studies [5, 8, 14] have revealed the diversity of combinations of individual practices and methods from both traditional and agile approaches, it was unexpected the variety of them within a small company (see Fig. 2(a) and (b)). The results of the analyses of the entire (non-filtered) dataset in HELENA study [5, 15]

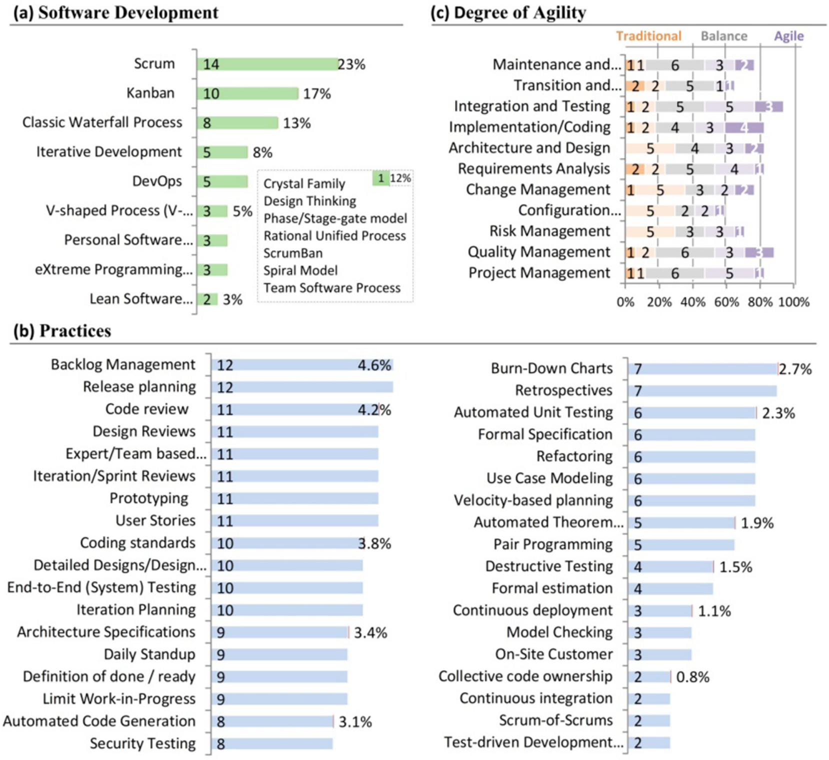


Fig. 2. a) Overview of the use of software development (a); practices (b); degree of agility (c).

identified three practices—Code Review, Coding Standards, Release Planning—as the most commonly used practices. In LogicStudio, Backlog Management and Release planning were two of the most used practices along with other six practices—Code Review, Design Reviews, Expert/Team based estimation, Iteration/Sprint Reviews, Prototyping and User Stories.

In spite of the fact that 71% of the respondents pointed out that implementing agile is easy in their context project, the remaining 29% perceive challenges. In HELENA study [5], half of the participants state that implementing agile is easy in their context (59%). The literature, e.g. [8, 14], identifies as a major challenge how to combine mixed approaches. However, in this study the identified challenges were documentation requirements to compliance with external standards (50%), resistance to change (25%) and lack of knowledge about the agile methods and practices (25%). By analyzing the data gathered through the inspection to the project repositories we did not find the description of all the used methods and practices. Therefore, it seems that some respondents rather use their experience to define their software processes. A likely explanation is that an explicit construction process is missing. In the literature, we found two relevant proposals that highlight its importance and aim to address this issue at both organizational and project levels [2, 8]. With regard to the degree of agility, respondents state that the approach tends to be agile as Fig. 2(c) shown. However, it is based on self-perception of the terms “agility” and “traditional”, and classified according to SWEBOK categories. When compared to the results of the HELENA study [5], there were less “don’t know” and “not answered” answers than the results in Fig. 2(c). However, 7 of the 11 SWEBOK categories seem to be aligned in both studies. In particular, “Change Management” is perceived more “agile” in the HELENA study [5] while “Maintenance and Evolution”, “Quality Management” and “Project Management” are perceived more “traditional” than the results in Fig. 2(c).

Similarly, according to the respondents’ perception, the current development approach is helpful. On one hand, about half of the respondents state that they are generally satisfied with the level of agility and they believe it is sufficient but 80% of them would change or improve it if they could. Those results seem to be aligned with the results of HELENA study [5]. On the other hand, 59% of the respondents reported that they (intentionally) combine different development approaches while 78% of participants positively answered this question in the HELENA study [5]. When comparing the reasons for external standards reported by this study with the reasons reported by HELENA study [5], most respondents agreed that, external models and standards such as CMMI and ISO 9001 are due to mainly the internal policies (65%, 60%), follows by external triggers (35%, 58%) and requirements of the company or the project business (12%, 52%). Therefore, despite that the reasons were the same their distribution was different. Regarding the assessment of compliance of both studies, in most of the projects, the compliance is firstly assessed by internal assessments (82%, 74%). However, respondents also point out that they have applied constructive measures (35%, 43%) and analytical measures (29%, 40%). Additionally, HELENA study [5] reported external project assessments (63%).

4.2 Lessons Learned

According to Jacobson and Stimson [2], it is evident that “we need better ways of working that put us on the road to real software engineering”. In fact, literature has underlined that the mixed approaches are a reality in the software industry and the situation reported in this case study provides empirical evidence to better understand the historical evolution of the software process in a small software company. Since the beginning the CEO/Funder realized that having a good software method provides a competitive advantage, but he knew by experience that one size does not fit all projects. Obviously, programming was the major concern in the early years so that the other important things were ad hoc. Over the period from 2003 to 2007, Logic Studio grew in terms of sales, customers and employees. However, the individual perspectives, prejudices and experiences of project managers were reflected, and not what the company as a whole had collectively learned during those years. In other words, every “method” was controlled by a “warden” as explained in [2]. By 2007, the corporate customers such as HSBC, Banesco, ASSA, Copa Airlines, Adidas and Sage USA expected that the projects were managed according PMBOK then Logic Studio adopted it. It was agreed that traditional processes provide predictability, stability and high quality assurance [6]. In fact, the CEO got a Project Management Professional (PMP) Certification that year. Then, the first approach reused what the company considered the best practices for their particular challenges and purposes. This finding is supported by Kuhrmann et al. [14], who concluded that experience, learning and pragmatism driven a natural process evolution.

Eventually, it was necessary that such an approach was explained and explicit. Hence, a digital repository was implemented. However, the schedules were inflexible and predetermined so that it was frequently necessary to adjust the estimated time as the company is mainly focused on innovation. In that context, agile offered the flexibility to more easily adjust to changes in project requirements [6]. Therefore, a small team was immediately given Scrum training. As a result, the CEO was the first person in becoming a Certified Scrum Master (CSM, 2008), Certified Scrum Product Owner (CSPO, 2010) and Agile Certified Professional (PMI-ACP, 2014). Nevertheless, adopting an agile approach was challenging not only for the practices, but also because, customers approved a fixed budget. Indeed, contract negotiation is an important aspect that remains a fundamental business instrument in many engagements [1]. Despite that fact, the company decided to take that risk in 2013 by developing a mixed approach that combined the unique strengths, and lessened the weaknesses of both approaches Scrum and PMBOK. Thus, project-specific software processes were built by choosing practices from PMBOK to give a "safe" environment for managers and practices from Scrum to achieve freedom for developers. In this sense, the systematic literature review carried out by Theocharis et al. [9] in 2015 reported that there is a clear trend toward adopting Scrum and some indication that Scrum is often used in combination with other software development approaches. Once the mixed approach was defined, the company developed a training plan, including conferences and workshops about Scrum, PMP and the mixed approach. Project managers were the first to attend the workshops followed by all development team members. Moreover, new employee orientation and induction processes were defined.

The mixed approach was defined as follows: starting with the most critical requirements for the specific project in order to deliver maximum value to the customer in the first iterations while the remaining requirements are left for next versions and additional planning cycles. In this way, according to the company, customers could see something tangible and feel pleased. These results are consistent with the benefit reported in some studies, e.g. [16, 17]. The company also acknowledged that the specific momentum was crucial in the adoption. It coincided with the banking innovation in Panama so that such a sector really appreciated the workshops and training courses about Scrum and PMP offered by the company. In 2014, the results provided some indication not only of the benefits but also of the need to strengthen the process. In this context, an increasing interest in software process improvement (SPI) enriched by continuous learning practices [18] emerged. Thus, a training course focused on "CMMI for Development v1.3" was undertaken and SPI became a high priority. In 2015, the software process were ISO 9001:2008 certified and after three annual audits the company has achieved an ISO 9001:2015 certification. In 2018, after 2 years of efforts, the company accomplished the goal of receiving a CMMI Level 3 certification. As a result, the defects have decreased although previous studies, e.g. [16, 17], have reported such an effect in CMMI Level 5 companies. Scrum and CMMI together bring a combination of adaptability and predictability [16]. The company relies on the current processes and the development team's freedom to pick between a full agile or full predictive project management and mix those practices.

Emerging challenges and new business opportunities, especially in USA, required more flexibility. In this context, DevOps captured the attention of the company due to its strong focus on rapid and continuous delivery. Then, a subset of DevOps practices was chosen in 2017 as a feasible approach to minimize risks. Apart from that, the SPI initiatives helped the company to realize that innovation projects should be a majority in the future. Therefore, after attending training on DT, it has been adopted since 2018. The initial results are in line with a previous study [19] that point out the extra effort spent on it helps the development teams to have a deeper understanding of the problem to be solved.

Regarding the effects of the whole mixed approach, two kinds of perceptions were identified. They are well illustrated by the following two quotes: "for our customers, more innovative projects which generate/add value from the beginning that are easier to justify and win approval of the [customer] management" and "for us, a more efficient planning and the satisfaction of keeping our customers satisfied". However, it is worth noting that the approach has a homegrown presentation that makes hard to compare it with others [2]. So far, the company cannot see value on use a formal specification to define it. Finally, it is worth to note that Logic Studio plans to adopt the indicators proposed in the CMMI Level 4 in the near future.

4.3 Threats to Validity

There are some threats to the validity of this study. The major threat is related to the sample size since this study is focused on only one small software company. The case company is not representative for all software small companies therefore the results must be interpreted with some caution when moving away from the features of the studied company. Moreover, surveys and interviews may be subject to post hoc rationalization and recall biases. However, the research approach allowed us to perform an in-depth qualitative analysis that used multiple sources of evidence. Such a triangulation process allowed us to validate insights by accessing different perspectives.

Another potential threat is related to conducting interviews due to the fact that it always includes some possible bias from the interviewer. To mitigate that threat two authors were involved and an interview guideline, which was previously reviewed by the other authors, was developed. The transcript was also sent to the interviewee to correct misunderstandings. The number of participants—convenience sampling strategy—and the self-reporting structure of the survey is another limitation of this study. To handle the risk related to convenience sampling strategy, this study followed the approach used in the HELENA study [15]. It means before analyzing the data, data preprocessing including a consistency check of the data was implemented. Regarding the HELENA survey instrument, it was developed and refined in several iterations by a team of researchers who built the questionnaire, tested and revised it. The threats to validity reported in [15] regarding the use of this instrument are briefly discussed below.

The online questionnaire might lead to incomplete or wrong answers since multiple-choice questions might have been incomplete and respondents may have misunderstood questions or answer options. To mitigate these threats, multiple-choice questions were complemented with a free-text option and a qualitative analysis of the free-text answers were carried out by the authors. The questionnaire was administered in Spanish to mitigate the risk of misunderstandings due to language issues. Finally, the results and conclusions were validated by the company.

5 Conclusions

In this paper, the historical evolution of a software development process in a small company was analyzed. In this case, such an approach was motivated by improving the quality of the services and products, not only for providing compliance with the international standards, but also for improving the flexibility and speed of response to customer needs and innovation challenges. The company was founded in 2003 but an incremental harmonization of different software development approaches started in 2007 and took around a decade. The first approach adopted by the company was PMBOK (2007). After five years (2013), the natural evolution resulted in a mixed approach based on PMBOK and SCRUM. That triggered a gradual improvement of company's capabilities for compliance with ISO 9001 (2015) and CMMI Level 3 (2018). However, emerging challenges and new business opportunities, especially in USA, required more flexibility, and rapid and continuous delivery then DevOps was integrated (2017). Lately, global competition has forced the company to adopt an innovation process so that DT (2018) has been incorporated. In this way, the company aims to make business successful while faces the need for change which is in line with *Business* and *Change* values of SPI Manifesto [20]. Agile approaches offer an appealing combination of economy and simplicity that allows small software company to increase the degree of success when they undertake software projects. In this sense, the CEO pointed out "we tried several other approaches to improve our software development life cycle, most of them with the precept of *'no pain, no gain'*. Once we found agility our customer easily come on board with the focus on early delivery of results and continuous value generation, now our motto is *'no gain, no way'*".

The findings from this study about the pragmatic process selection and its evolution over time are aligned with the results of HELENA study [5], but this study obtains further in-depth information. The sense of urgency and vision, called "common sense" by the CEO, along with continuous learning practices have addressed the evolution of the software development process. But the actual practice is what makes the company agile and enables it to adapt to changing situations without sacrificing formality. In fact, the actual practice reflects the mixed approach developed. One interesting point is the unexpected diversity of software development approaches and practices founded within a small company. Therefore, further research is needed to understand this fact.

Despite the effort made, this mixed approach also requires a construction process in order to enable a systematic design of the project-specific development approach as suggested by [2, 8]. The evolution of software process in the company depends on its adaptability which captures its capacity to learn, combine experience and knowledge, and tailor its software process to changing environment. The main limitation of this study is the sample, which is one small company as well as the collected and analyzed data is from (17) respondents mostly located in Panama. Therefore, further research is also needed to increase coverage. Although the results of the mixed approach are promising, additional studies are necessary to know how agile and traditional approaches are combined and how they relate to the particular company context [9]. Moreover, it is worth noting that the DT approach integrated with SCRUM has been little studied in the scientific literature [19].

Finally, this study aims to encourage empirical research and documentation of the lessons learned from companies that seek synergy between traditional and agile software development approaches.

References

1. Sánchez-Gordón, M.-L., Colomo-Palacios, R., de Amescua Seco, A., O'Connor, R.V.: The route to software process improvement in small- and medium-sized enterprises. In: Kuhrmann, M., Münch, J., Richardson, I., Rausch, A., Zhang, H. (eds.) Managing Software Process Evolution, pp. 109–136. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-31545-4_7
2. Jacobson, I., Stimson, R.: Escaping method prison – on the road to real software engineering. In: Gruhn, V., Striemer, R. (eds.) The Essence of Software Engineering, pp. 37–58. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-73897-0_3
3. Clarke, P., O'Connor, R.V., Yilmaz, M.: In search of the origins and enduring impact of agile software development. In: Proceedings of the 2018 International Conference on Software and System Process, pp. 142–146. ACM Press, Gothenburg (2018). https://doi.org/10.1145/3202710.3203162
4. Vijayasarathy, L.R., Butler, C.W.: Choice of software development methodologies: do organizational, project, and team characteristics matter? IEEE Softw. **33**, 86–94 (2016)
5. Kuhrmann, M., Tell, P., Klünder, J., Hebig, R., Licorish, S., MacDonell, S.: Complementing Materials for the HELENA Study (Stage 2). https://doi.org/10.13140/rg.2.2.11032.65288
6. Boehm, B., Turner, R.: Balancing Agility and Discipline: A Guide for the Perplexed. Addison-Wesley, Boston (2003)
7. Kuhrmann, M., et al.: Hybrid software development approaches in practice: a european perspective. IEEE Softw. **36**, 20–31 (2018)
8. Küpper, S., Rausch, A., Andelfinger, U.: Towards the systematic development of hybrid software development processes. In: Proceedings of the 2018 International Conference on Software and System Process, pp. 157–161. ACM, New York (2018). https://doi.org/10.1145/3202710.3203158
9. Theocharis, G., Kuhrmann, M., Münch, J., Diebold, P.: Is *Water-Scrum-Fall* reality? On the use of agile and traditional development practices. In: Abrahamsson, P., Corral, L., Oivo, M., Russo, B. (eds.) PROFES 2015. LNCS, vol. 9459, pp. 149–166. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-26844-6_11
10. The Hasso Plattner Institute of Design: Get Started with Design Thinking. https://dschool.stanford.edu/resources/getting-started-with-design-thinking
11. Project Management Institute: A Guide to the Project Management Body of Knowledge (PMBOK Guide). Project Management Institute (2013)
12. Schwaber, K., Sutherland, J.: Home—Scrum Guides. https://www.scrumguides.org/

13. Tell, P., MacDonell, S., Licorish, S.A.: 3rd workshop on hybrid development approaches in software system development. In: Kuhrmann, M., et al. (eds.) PROFES 2018. LNCS, vol. 11271, pp. 433–440. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-03673-7_34
14. Kuhrmann, M., et al.: Hybrid software and system development in practice: waterfall, scrum, and beyond. In: Proceedings of the 2017 International Conference on Software and System Process, pp. 30–39. ACM, New York (2017). https://doi.org/10.1145/3084100.3084104
15. Tell, P., et al.: What are hybrid development methods made of? An evidence-based characterization. Presented at the International Conference on Software and Systems Process ICSSP, Montreal, QC, Canada, Canada (2019). https://doi.org/10.1109/ICSSP.2019.00022
16. Sutherland, J., Jakobsen, C.R., Johnson, K.: Scrum and CMMI level 5: the magic potion for code warriors. In: Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008), p. 466 (2008). https://doi.org/10.1109/HICSS.2008.384
17. Ortiz, G.A., Trujillo, M.E.M., Oktaba, H., Hernandez, E.R.: Integrating agile methods into a level 5 CMMI-DEV organization: a case study. IEEE Lat. Am. Trans. **14**, 1440–1446 (2016). https://doi.org/10.1109/TLA.2016.7459632
18. Kuhrmann, M., Müench, J.: SPI is dead, isn't it? Clear the stage for continuous learning! In: 2019 IEEE/ACM International Conference on Software and System Processes (ICSSP), pp. 9–13 (2019). https://doi.org/10.1109/ICSSP.2019.00012
19. Pereira, J.C., de F.S.M. Russo, R.: Design thinking integrated in agile software development: a systematic literature review. Procedia Comput. Sci. **138**, 775–782 (2018). https://doi.org/10.1016/j.procs.2018.10.101
20. Pries-Heje, J., Johansen, J.: SPI Manifesto. http://www.iscn.com/Images/SPI_Manifesto_A.1.2.2010.pdf

A Developer Driven Framework for Security and Privacy in the Internet of Medical Things

Ceara Treacy(✉), John Loane, and Fergal McCaffery

Regulated Software Research Centre & Lero,
Dundalk Institute of Technology, Dundalk, Ireland
{ceara.treacy, john.loane, fergal.mccaffery}@dkit.ie

Abstract. The General Data Protection Regulation has propelled privacy to an equal requirement with security for data protection. A key regulation requirement is data protection by design and a corresponding data protection impact assessment that details security and privacy by design for a project processing data. Security and privacy of data in the Internet of Medical Things (IoMT) handles sensitive health data and accordingly is bound by additional regulatory safety, security and privacy requirements. Health data in the IoMT can potentially flow through a diversity of apps, systems, devices and technologies, open networks. This exposes the transmitted data in the IoMT to additional attack surfaces and consequently increases on the need for security and privacy. Applying security and privacy regulatory requirements is a struggle for developers in small to medium enterprises due to a lack of knowledge, understanding, training, experience and financial constraints. This paper proposes a framework aimed at developers in small to medium enterprises, to assist in meeting regulatory requirements for security and privacy of data in flow in the IoMT. This framework expands on the basic established threat modeling steps to apply both security and privacy properties to protect data in flow in the IoMT. The framework includes a set of categorised technical security and privacy controls developed through medical device security standards to mitigate identified security and privacy threats. This framework also provides a foundation for the administration of a data protection impact assessment.

Keywords: Security · Privacy · Threat modeling · IoMT · Security and privacy controls

1 Introduction

The IoMT, is a growing domain as the benefits of moving healthcare into the community and centering health with the patient becomes common. However, as the IoMT grows, cybersecurity and privacy risks have risen [1, 2]. Reports [3, 4], have shown that security and privacy in the healthcare domain is lacking in maturity and that small to medium enterprises (SMEs) and developers struggle with understanding and fulfilling security and privacy requirements in the domain. The lack of maturity, understanding and experience is due to a variety of historic and new issues related to security and privacy in development in this domain. Historic issues for privacy and security include [5]: budget constraints, deficiency in knowledge and lack of trained personnel

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 107–119, 2020.
https://doi.org/10.1007/978-3-030-56441-4_8

[3, 4, 6]. In the same way, the historic 'bolt on' approach, where security is bolted on at a later phase in development or after a system is designed, has been shown to be problematic, expensive, and failure prone [7]. Likewise, the consideration of privacy throughout the system development lifecycle known as a Privacy by Design approach [8], is characterised by anticipating and preventing privacy-invasive events before they happen [9]. Security and privacy is often treated as a feature that can be "bolted on" late in the development lifecycle [10–12].

The newer issues include the increased complexity and compatibility issues in terms of the variety of technologies in use in the IoMT, which have an impact on the rise in security and privacy risks [13]. Data in flow in the IoMT can be through various apps, systems, devices, technologies and open networks, which are inherently insecure such as wireless sensor networks and the cloud [3]. This is vastly different to the closed networks the healthcare domain has traditionally been based in. An equally problematic issue, is the rush into the lucrative healthcare domain and the speed the healthcare domain is embracing the IoMT, without a profound understanding of the security and privacy risks [14–16]. Added to these issues, are the complexities for SMEs and developers in understanding the security and privacy regulatory requirements [17]. The EU General Data Protection Regulation (GDPR) regulatory requirement of *'data protection by design and by default'* [18] and expert recommendations [7, 9, 19–21], direct that security and privacy are designed into a project from the beginning. This includes considering security and privacy in the devices, the communication protocols and the services [22]. The GDPR requirement of Article 24(1) states an organization is to *'implement appropriate technical and organizational measures to ensure and demonstrate compliance with the Regulation'* and document a data protection impact assessment (DPIA) [18]. A DPIA is an effective way to assess and demonstrate a project's compliance with the data protection principles and obligations [23]. There are guidance documents provided on what a DPIA should include but there are no templates as data protection is unique to each project, which makes this requirement challenging for small and medium enterprises. The framework fosters documentation of the process, which contributes to evidence that the project has fulfilled GDPR requirements. For an inclusive DPIA the current framework would require further extension to include data storage, as this research looks explicitly at the security and privacy of data in flow in the IoMT. There is also a lack of knowledge and understanding of the standards and limited European or international standards designed to assist SMEs towards ensuring appropriate protection of data [24].

Given the summarized complexities, this paper presents a framework aimed at developers in SMEs, to assist in meeting regulatory compliance for security and privacy of data in flow in the IoMT. This paper is structured as follows. Section 2 looks at related work and positions the framework with respect to these. Section 3 provides a detailed description of the framework. Finally, Sect. 4 details our conclusions and future work.

2 Related Work

Threat Modeling (TM) is generally used for assessing security threats in software. It was developed by Microsoft and is a part of its security development lifecycle [25]. It is widely used and there are many TM frameworks and methods for assessing security threats [6, 21]. Some of the most common attack models are; STRIDE [26], Attack Trees [27], Persona non Grata [28] and Developer Driven Threat Model Process (DDTM) [6]. Common TM frameworks include; Quantitative Threat Modeling Method [29], Trike [30], PASTA [31], OCTAVE [32] and Hybrid Threat Modeling Method (hTMM) [29]. STRIDE is the most widely used TM tool for security threats [33] but it is important to recognize that it *does not consider privacy threats* [34]. A TM tool that addresses privacy threats is LINDDUN [34], which is a systemic approach to assist with the elicitation and mitigation of privacy threats in software systems. There are tools emerging mainly for security threats such as the Microsoft Tool for TM [35], and the OWASP project, the Threat Dragon tool [36]. There is a tool in development for security and privacy threat analysis called SPARTA [37]. Other research in this area comprises of privacy extensions for DFDs that extend from LINDDUN [38] to include new DFD elements and labels to address privacy in the TM process. Other DFD extensions research concerns already known security decisions and constraints before a project begins [39]. Similarly research in this specific area [40], aims to provide DFDs extensions in a structured way for pre-determined security decisions and constraints.

This framework expands on the established TM steps in the DDTM created by Dhillon [6] within his professional security development experience with Microsoft. The DDTM already features STRIDE, the framework extends this to include the LINDDUN privacy threats to provide both security and privacy threat elcitation in the process. The framework also provides standardized security and privacy properties mapped to these two threat categories and the threat categories are mapped to the OWASP Top 10 [41]. The framework provides a set of Data Flow Security and Privacy Controls (DFSPCs), based on the medical standard IEC/TR 80001-2-8:2016 [42], to mitigate the identified threats. The DFSCs are categorized to the standardized security and privacy properties, the common goals the framework wants to protect.

3 Framework

The framework has six steps: **Step 1:** Contextual knowledge - to link security and privacy properties; **Step 2:** System decomposition; **Step 3:** Threat identification; **Step 4:** Threat Analysis; **Step 5:** Identify security and privacy properties against threats; **Step 6:** Selection of controls to mitigate threats.

3.1 Step 1 Contextual Knowledge

Framework step 1 aims to provide contextual knowledge to assist new or inexperienced developers to understand the security and privacy context of the framework. There are three parts to step 1 i.e. part 1, part 2 and part 3.

Part 1 Security and Privacy Objectives. Personal data security and privacy objectives are the foundation of the framework and should map to the organisations whilst balancing the regulatory data privacy and security obligations. The framework uses the security objectives of the Information Security Management System standard ISO/IEC 27001 [43] and the privacy objectives of the Privacy Information Management Systems standard ISO/IEC 27701 [44]. Security and privacy objectives should secure and keep data private that an organisation considers confidential. This is explicitly personal data in the IoMT and the regional categorization of personal data. Regional categorization of personal data is referenced in ISO/IEC 27701 and varies between different regulatory regimes. The framework outlines a classification scheme for personal data in accordance to ISO/IEC 27701 and to include any regional categorization. The classification includes definitions and explanations of: the nature of the data - e.g. personal health information; personal data principals concerned - e.g. personal data relating to children, the regulatory requirements differ in relation to the age of consent and how the information is processed and used; changing or extending the purposes for the processing of personal data, which will require updating and/or revision of the legal basis and additional consent for use from the data subject.

Part 2 Security and Privacy Properties. Security and privacy properties are the common goals the framework wants to protect. Traditional data security properties observe ensuring the confidentiality, integrity and availability (CIA) of information. However, Whitman and Mattord [45] note that the CIA triangle model no longer adequately addresses the constantly changing environment. Consequently, the framework uses the security properties in Part 3 of the standard ISO/IEC 27033 for network security [46]. An added security property Authorization referenced in STRIDE is included, it is not an individual security property in ISO/IEC 27033. The security properties of the framework are: Authentication, Integrity, Non-repudiation, Confidentiality, Availability, Authorization, Access Control, Communication or Transport Security, and Privacy/Opacity. The framework employs the privacy properties used to develop the LINDDUN framework [47]. The acronym LINDDUN refers to the privacy threat categories: Linkability, Identifiability, Non-repudiation, Detectability, Disclosure of information, Unawareness of content, and Non-compliance of policy, which map to the privacy properties Unlinkability, Anonymity & Pseudonymity, Plausible deniability, Undetectability & Unobservability, Confidentiality, Content Awareness and Policy and Consent Compliance.

Part 3 Administer Two LINDDUN Privacy Threat Categories. Part 3 manages the two LINDDUN soft privacy properties of Content Awareness and Policy and Consent and their corresponding privacy threat categories of Unawareness and Non-compliance. The Policy and Consent property is violated by the non-compliance threat category. Non-compliance means not following the (data protection) legislation, the advertised policies or the existing user consents of the regulatory jurisdiction [48]. The framework outlines GDPR privacy policy and consent requirements. Additionally, developers need to know that any process being developed is within the privacy policy consent acceptance. Recognition that consent has been received from the user for collection of the data and that the data is used according to the consent and not used for any other purposes. Developers need to adhere to the specifics of the privacy policy to know where in development; when to get consent for collecting data, understanding who the product is demographically targeting including special requirements, the reason and use of the collected data. Any data collected or used outside the privacy policy would be a breach of privacy regulatory requirements.

3.2 Step 2 System Decomposition

Step 2 of the framework is system decomposition based on Dhillon's DDTM [6], which features guidelines for creating DFDs for developers with or without security expertise. The framework uses DFDs because they are an established tool used by TMs, are easy to use [21, 49] and are used by both STRIDE and LINDDUN [50]. DFDs support following the flow of data through a system. Problems tend to follow the data flow and DFDs facilitate discovery of problems, which can be highlighted and addressed [21]. Additionally, ISO/IEC 27701 advises the use of DFDs as helpful tools to inform a DPIA and risk assessment transfer, which assists with regulatory requirements. STRIDE and LINDDUN also provide a set of threat types in relation to the elements of DFDs, which makes the framework easier to use. Recommended personnel present for system decomposition through DFDs include: software architect, lead developers, product Owner or representative, project manager or lead, security champion and business owner to represent the objectives of the organisation. In SMEs personnel may fulfil more than one role. The framework defines three key features for creating DFDs to assist inexperienced developers and SMEs: 1. DFDs Elements and Symbols 2. Decomposition Levels 3. Annotations.

DFDs Elements and Symbols. The framework uses the set of DFD symbols devised by Gane and Sarson [51] data stores, processes, data flows, and external entities, with the adapted process symbols presented by Ibrahim and Yen [52, 53]. The framework also includes the diversity of boundaries, Machine Boundary and Process Boundary, presented by Osterman [49] and Shostack [54], and used by Dhillon [6]. While the use of DFDs for TM security in systems is well established, their use for privacy is an emerging and developing area. Antignac et al. [38], completed research in the area of privacy and DFDs outlining that GDPR meaning that personal data processing involves: collection, disclosure, usage, record, retrieval, and erasure. They applied the definition of these concepts to establish a link with the privacy requirements of personal data as specified in standards and regulatory texts. They presented a usable DFD extension to the LINDDUN TM process called Privacy-Aware Data Flow Diagrams (PA-DFDs). The researchers state that the PA-DFDs privacy extensions reflect the personal data processing concepts from the GDPR and references the privacy principles of the ISO/IEC 29100:2011+A1:2018 [55]. PA-DFDs introduced three different types of privacy external entities; data subjects, data controllers, and data processors, established from the three main types of natural and legal persons discussed in regulatory texts. They also added two new process elements, *usage* and *complex usage* processes and introduced a new element called Erasure, which pertains to the purpose of ensuring that erased data comply with the GDPR *'right to be forgotten'* principle. The framework incorporates the principles from these DFD models to provide for both security and privacy.

Decomposition Levels. The purpose of decomposition is to break down the system to create a security and privacy profile based on areas of vulnerability for security and privacy. The framework uses the DFDs decomposition levels presented in research by [56] and [6]. The DFDs hierarchical levels are: Level 0 - context diagram, system as a single entity; Level 1 - more detailed DFD through refinement of the system; Level 2 - more complex systems requiring decomposition of some components with annotations; Level 3 – further decomposition of complex systems. Dhillon [6] specified that in Microsoft they have not needed to decompose to Level 3. Levels 1 and 2 of DFD abstraction would include the annotations outlined in the next section. The framework fosters documentation for both security and privacy beginning at decomposition Level 0 that include the security and privacy properties and personal data classification of the project. The security and privacy documentation will change with the level of decomposition but, uniformity in information and personal data is required through the levels of decomposition. The framework applies the eleven security privacy principles detailed in the privacy framework standard ISO/IEC 29100 [55], to provide a knowledge basis for developers in privacy.

Annotations. The framework added annotations for both security and privacy. The security annotations are based in Dhillon's [6] DDTM process. He recommends the use of annotations on the DFDs to capture additional information to include interactions within the system. Dhillon found that DFD elements alone do not capture all necessary details to perform effective threat modelling. The addition of annotations that consider the interactions of the system to the DFD, provides additional information to assist and quicken the TM process by bringing focus on the typical areas attackers are interested in and which are common sources of vulnerabilities [6, 57]. The vulnerabilities can be used to identify interactions that could introduce weaknesses making the identification process faster [57]. Security annotations examples include critical security functions such as; authentication, password management, and cryptographic operations. Network and local dataflow; HTTP, API call and file I/O.

Antignac et al. [38], introduced privacy annotations for the elements and interactions between the DFD elements. They added the annotation of purpose, *purp*, to the usage and complex usage process elements, to track the purposes to which the data subject has consented. Antignac et al. [38], note regulations expect all kinds of personal data processing to be associated with a purpose. They labelled the data flow element carrying personal data with *pdata* and provided a link through a dotted line to the corresponding data subject, which the personal data refers to [38]. Labelling the data flow *pdata* provides records of where personal data is in the system and supports developers in identifying where to provide the added controls to keep the data secure and private. Annotations can be added at any level of abstraction but, as stated previously should evolve in line with the previous abstractions of the system.

3.3 Step 3 Threat Identification

Threat identification is central to the TM process but, is also one of the most difficult aspects of the process for developers with little or no experience [6]. Categorizing threats makes it easier to understand what the threat allows an attacker to do and supports in assigning priority and mitigation [33]. The framework uses the STRIDE threat categories for security and the LINDDUN threat categories for privacy. The recommendation from the LINDDUN creators is to implement LINDDUN alongside STRIDE [57]. Both STRIDE and LINDDUN complete threat identification using element-based and per interaction-based methods centered upon the building blocks of DFDs [40]. The framework utilizes both per-element and per interaction-based methods.

Threat identification per-element is where each element in the DFD is analysed through the threat categories. Not all threat categories from both STRIDE and LINDDUN apply to all DFD element types. The overlapped STRIDE and LINDDUN (Bold) threat categories to DFD elements mapping are presented in Table 1. The LINDDUN threat categories unawareness and non-compliance refers to regulatory compliance and are addressed in step 1. However, it is vital that developers understand the regulatory compliance requirements for personal data including the personal data classification, consent and the privacy policy addressed in step 1 of the framework.

Table 1. STRIDE and LINDDUN threats categories mapped per DFD element.

Threat Category STRIDE & **LINDDUN**	Entity (External)	Data Flow	Data Storage	Process
Spoofing	X			X
Tampering		X	X	X
Repudiation **Non-repudiation**	X	**X**	X	X
Information disclosure **Disclosure of information**		X	X	X
Denial of service		X	X	X
Elevation of privilege				X
Linkability	**X**	**X**	**X**	**X**
Identifiability	**X**	**X**	**X**	**X**
Detectability		**X**	**X**	**X**
Unawareness	**X**			
Non-compliance		**X**	**X**	**X**

Threat identification per-interaction considers all interactions taking place in the system. It considers the origin, destination and interaction in the system and identifies threats against them [21]. Entry and exit points are the places where data enters or exits the system [58] and identifying these throughout the system will be aided through establishing the interactions. The key categories for entry points to a system that offer a way-in for attackers and generally overlap with trust boundaries include; communication (especially wireless), software and physical (hardware).

3.4 Step 4 Threat Analysis

Threat analysis is a challenging aspect of TM and there is a degree of understanding and knowledge required to map to tangible attacks [6]. Within security TMs, it is recognized that without a degree of understanding and knowledge, STRIDE threats can't be used effectively. In addition, both the STRIDE and LINDDUN categories are abstract, which means that attacks could apply to one or more of the threat categories, which requires understanding and knowledge. Given the level of knowledge and experience required to complete threat analysis, the framework maps the STRIDE and LINDDUN threat categories to the OWASP Top 10 [41]. The purpose of this mapping is to provide initial guidance for developers with no experience, to assist in developing knowledge and grow confidence with threat elicitation and analysis. Future mappings for the framework could include other threat and vulnerability resources. Stage 3 adopts the threat analysis approach from NIST SP 800-30 and the formula Risk = Likelihood $\times$ Impact [59]. The framework uses as a starting point with appropriate tailoring, Appendices G and H of NIST SP 800-30. The appendices, provide sets of exemplary tables for use in determining adverse likelihood and impacts quantitatively. Using NIST SP 800-30, provides standardized guidance for SMEs and developers with little or no knowledge and experience. Threat prioritization is guided by the quantitative outcomes of the formula and managed by the sensitivity of the personal data and safety of the patient associated with the threat and vulnerabilities.

3.5 Step 5 Map Threats to Security and Privacy Properties

The use of the security and privacy properties to categorize the DFSCs was completed because of their relationship to the standards. Each STRIDE and LINDDUN threat category maps to the security and privacy property it violates. The threat categories are affiliated to the security and privacy properties, not to the standards. step 5 maps the threats identified in step 4 to the frameworks security and privacy properties. The framework has extracted controls from the technical standards, defined in step 6, according to the security and privacy properties. The purpose of this mapping is to simplify identification of appropriate security and privacy controls to mitigate the identified threats and subsequent vulnerabilities. Similar to the threat elicitation and analysis stages, there are commonalities and overlapping controls in the property categories. Developers find the associated security or privacy category to through the mapped threat category. They then consult the identified security or privacy category's controls and find the control suitable to mitigate the identified threat and vulnerability. An example of this process would be the identification of the cross-site request forgery attack vulnerability in step 4, which is A7 OWASP Top 10 threat. This vulnerability corresponds to the STRIDE threat category spoofing, which violates the security property category authentication. The developer examines the security controls for authentication to find a control to address the cross-site vulnerability.

3.6 Step 6 Data Flow Security and Privacy Controls

Step 6 of the framework is the identification of countermeasures to protect the security and privacy properties through the application of appropriate security and privacy controls. The lack of standards and guidance specifically in this domain that straddle development for both security and privacy, has created a vacuum for inexperienced SMEs and developers. This means existing standards from security and privacy, must be applied in a patchwork method to address security and privacy [5], which requires extensive knowledge and understanding. A key part of the framework is a set of technical controls applicable for the security and privacy of data in flow, the DFSPCs. The DFSPCs are a set of technical controls based in standards to assist with security and privacy regulatory requirements and close the gap in knowledge and understanding in this area. The DFSPCs fill the vacuum of specific technical controls for the security and privacy of data in flow to assist developers to comply with the regulatory requirements. The DFSPCs originated in the international standard IEC/TR 80001-2-8 [42], which should be considered a basic foundation in security [60]. The IEC/TR 80001-2-8 controls are to manage risks to confidentiality, integrity, availability and accountability of data and systems and do not consider privacy of data or the span of security and privacy properties identified for the framework. Therefore, the development of the DFSPCs was expanded to include technical standards used for IEC/TR 80001-2-8. There were four standards used for technical controls: ISO/IEC 15408-2 [61], ISO/IEC 15408-3 [62], NIST 800-53 Rev. 4 [63] and IEC 62443-3-3 [64]. ISO/IEC 15408-3 defines the assurance requirements of the evaluation criteria, which was deemed out of scope for the DFSPCs and was excluded. The remaining three standards were examined to establish security and privacy technical controls for the DFSPCs. There are currently 63 controls validated for security. The privacy controls are still in development and will be validated by two privacy experts. The technical controls for security have been categorised into the security properties. The validation of the DFSPCs is currently being finalized and on completion they will be implemented in the framework.

4 Conclusions

In this paper, we presented a developer driven framework for security and privacy of data in flow in the IoMT. This framework aims to support and assist inexperienced developers and SMEs elicit threats and apply security and privacy controls, to meet regulatory requirements. The framework focuses on providing a structured process for SMEs and developers to address both security and privacy and develop confidence in TM to ensure that the data in flow of their systems is secure and private.

Future work on this framework will entail a comprehensive validation of the framework. The security controls have been validated by experts from the security development and the medical standards domains. Experts from the privacy standard domain and industry have been identified to validate the privacy controls. Three experts in the TM domain have been identified to complete and expert review of the framework. On completion of the expert validation, the framework will be implemented within two companies from the medical domain.

Acknowledgements. This work was supported with the financial support of the Science Foundation Ireland grant 13/RC/2094.

References

1. Brien, G., et al.: SP 1800-8 Securing wireless infusion pumps. In: Healthcare Delivery Organizations. National Institute of Standards and Technology (2018)
2. Papageorgiou, A., et al.: Security and privacy analysis of mobile health applications: the alarming state of practice. IEEE Access **3536**(c), 1–13 (2018)
3. Ponemon Institute: The State of Cybersecurity in Healthcare Organizations in 2018. https://ponemonsullivanreport.com/2018/03/the-state-of-cybersecurity-in-healthcare-organizations-in-2018/. Accessed 07 Apr 2020
4. Cisco 2018 Annual Cybersecurity Report. https://www.cisco.com/c/dam/m/hu_hu/campaigns/security-hub/pdf/acr-2018.pdf. Accessed 07 Apr 2020
5. Treacy, C., McCaffery, F.: Data security overview for medical mobile apps assuring the confidentiality, integrity and availability of data in transmission. Int. J. Adv. Secur. **9**(3&4), 146–157 (2016)
6. Dhillon, D.: Developer-driven threat modeling: lessons learned in the trenches. IEEE Secur. Priv. **9**, 41–47 (2011)
7. Schneier, B., Shostack, A.: Breaking Up Is Hard to Do: Modeling Security Threats for Smart Cards, Chicago (1999)
8. Cavoukian, A.: Privacy by Design…take the challenge. Information & Privacy Commissioner, Ontario (2009)
9. De Francesco, G.P.: The general data protection regulation's practical impact on software architecture. Computer (Long. Beach. Calif.) **52**(4), 32–39 (2019)
10. McGraw, G.: Software Security Building Security. Addison-Wesley, Reading (2006)
11. Hall, J.L., McGraw, D.: For telehealth to succeed, privacy and security risks must be identified and addressed. Health Aff. **33**(2), 216–221 (2014)
12. Dhanvijay, M.M., Patil, S.C.: Internet of Things: a survey of enabling technologies in healthcare and its applications. Comput. Netw. **153**, 113–131 (2019)
13. Alsubaei, F., Abuhussein, A., Shandilya, V., Shiva, S.: IoMT-SAF: Internet of medical things security assessment framework. Internet Things **8**, 100123 (2019)
14. Dilawar, N., Rizwan, M., Ahmad, F., Akram, S.: Blockchain: securing internet of medical things (IoMT). Int. J. Adv. Comput. Sci. Appl. **10**(1), 82–89 (2019)
15. Hatzivasilis, G.O., et al.: Review of security and privacy for the internet of medical things (IoMT): resolving the protection concerns for the novel circular economy bioinformatics. In: Proceedings of the 15th Annual International Conference on Distributed Computing in Sensor Systems, Greece, pp. 457–464 (2019)
16. Sun, W., Cai, Z., Li, Y., Liu, F., Fang, S., Wang, G.: Security and privacy in the medical Internet of Things: a review. Secur. Commun. Netw. **2018**, 1–9 (2018). https://www.hindawi.com/journals/scn/2018/5978636/. Accessed 22 Jul 2020
17. Parker, L., et al.: A health app developer's guide to law and policy: a multi-sector policy analysis. BMC Med. Inform. Decis. Mak. **17**(1), 1–13 (2017)
18. GDPR: (EU) 2016/679 of the European Parliament and of the council on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC General Data Protection Dir. pp. 1–88, EU (2016)

19. Tondel, I.A., Jaatun, M., Meland, P.: Security requirements for the rest of us: a survey. IEEE Softw. **25**(1), 20–27 (2008)
20. Danezis, G., et al.: Privacy and Data Protection by Design - from policy to engineering. ENISA (2014)
21. Shostack, A.: Threat Modeling Designing for Security. Wiley, Indiana (2014)
22. McManus, J.: Security by design: teaching secure software design and development techniques. J. Comput. Sci. Coll. **33**(3), 75–82 (2018)
23. ICO: What is a DPIA? Data Protection Impact Assessments (DPIAs). https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/data-protection-impact-assessments-dpias/what-is-a-dpia/. Accessed 03 Jan 2020
24. Manso, C., Rekleitis, E., Papazafeiropoulos, F., Maritsas, V.: Information security and privacy standards for SMEs. Recommendations to improve the adoption of information security and privacy standards in small medium enterprises. ENISA, Heraklion (2015)
25. Howard, M., Lipner, S.: The Security Development Lifecycle. Micorsoft Press, Redmond (2006)
26. Kohnfelder, L., Garg, P.: The threats to our products. Microsoft Press (1999)
27. Schneier, B.: Academic: attack trees. Dr. Dobbs J. **24**(12), 21–29 (1999)
28. Cleland-Huang, J.: Meet elaine: a persona-driven approach to exploring architecturally significant requirements. IEEE Softw. **30**(4), 18–21 (2013)
29. Mead, N.R., Shull, F., Vemuru, K., Villadsen, O.: A hybrid threat modeling method. Carnegie Mellon University, Virginia (2018)
30. Saitta, P., Larcom, B., Eddington, M.: Trike v. 1 Methodology Document. https://www.octotrike.org/papers/Trike_v1_Methodology_Document-draft.pdf. Accessed 07 Apr 2020
31. UcedaVelez, T., Morana, M.M.: Risk Centric Threat Modeling: Process for Attack Simulation and Threat Analysis. Wiley, Hoboken (2015)
32. Alberts, C., Dorofee, A., Stevens, J., Woody, C.: Introduction to the OCTAVE Approach. Carnegie Mellon University, Pittsburgh (2003)
33. Hussain, S., Kamal, A., Ahmad, S., Rasool, G., Iqbal, S.: Threat modelling methodologies: a survey. Sci. Int. (Lahore) **26**(4), 1607–1609 (2014)
34. Deng, M., et al.: A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements. Requirements Eng. **1**(16), 3–32 (2011). https://doi.org/10.1007/s00766-010-0115-7
35. Microsoft: Microsoft Threat Modeling Tool - Azure | Microsoft Docs. https://docs.microsoft.com/en-us/azure/security/develop/threat-modeling-tool. Accessed 06 Feb 2020
36. OWASP Threat Dragon 2020. https://owasp.org/www-project-threat-dragon/. Accessed 03 Mar 2020
37. Sion, L., Van Landuyt, D., Yskout, K., Joosen, W.: SPARTA: security & privacy architecture through risk-driven threat assessment. In: IEEE 15th International Conference Software Architecture Companion (ICSA-C), Seattle, pp. 89–99. IEEE (2018)
38. Antignac, T., Scandariato, R., Schneider, G.: A privacy-aware conceptual model for handling personal data. In: Margaria, T., Steffen, B. (eds.) ISoLA 2016. LNCS, vol. 9952, pp. 942–957. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-47166-2_65
39. Yampolskiy, M., et al.: Systematic analysis of cyber-attacks on CPS-evaluating applicability of DFD-based approach. In: Proceedings of the 5th International Symposium on Resilient Control Systems, ISRCS, Salt Lake City, pp. 55–62. IEEE (2012)
40. Sion, L., Yskout, K., Van Landuyt, D., Joosen, W.: Solution-aware data flow diagrams for security threat modeling. In: Proceedings of the ACM Symposium on Applied Computing, France, pp. 1425–1432. ACM (2018)
41. OWASP Top Ten. https://owasp.org/www-project-top-ten/. Accessed 05 Mar 2020

42. IEC/TR: 80001-2-8:2016 Application of risk management for IT-networks incorporating medical devices — Application guidance Part 2–8 : Guidance on standards for establishing the security capabilities identified in IEC TR 80001-2-2. British Standards Publication (2016)
43. ISO/IEC: 27001:2017 Information technology - Security techniques - Information security management systems - Requirements (ISO/IEC 27001:2013). British Standards Publication (2017)
44. ISO/IEC: 27701:2019 Security techniques — Extension to ISO/IEC 27001 and ISO/IEC 27002 for privacy information management — Requirements and guidelines. British Standards Publication (2019)
45. Whitman, M.E., Mattord, H.J.: Management of Information Security, 3rd edn. Cengage Learning, Boston (2010)
46. ISO/IEC: 27033-3 -Information technology — Security techniques — Network security Part 3: Reference networking scenarios — Threats, design techniques and control issues. BSI Standards Publication (2010)
47. Wuyts, K., Scandariato, R., Joosen, W.: Empirical evaluation of a privacy-focused threat modeling methodology. J. Syst. Softw. **96**, 122–138 (2014)
48. Wuyts, K., Joosen, W.: LINDDUN Privacy Threat Modeling: A Tutorial. Department of Computer Science, Belgium (2015)
49. Osterman, L.: Threat modeling, once again – Larry Osterman's WebLog, Microsoft | Developer (2007). https://blogs.msdn.microsoft.com/larryosterman/2007/08/30/threat-modeling-once-again/. Accessed 12 Jan 2020
50. Sion, L., Wuyts, K., Yskout, K., Van Landuyt, D., Joosen, W.: Interaction-based privacy threat elicitation. In: Proceedings of the 3rd IEEE European Symposium on Security and Privacy Workshops. EURO S&P, London, pp. 79–86. IEEE (2018)
51. Gane, C., Sarson, T.: Structured Systems Analysis: Tools and Techniques, 1st edn. Prentice Hall Professional Technical Reference, Englewood Cliffs (1979)
52. Ibrahim, R., Yen, S.Y.: An automatic tool for checking consistency between Data Flow Diagrams (DFDs). World Acad. Sci. Eng. Technol. **70**, 615–619 (2010)
53. Ibrahim, R., Yen, S.Y.: A formal model for data flow diagram rules. ARPN J. Syst. Softw. **1** (2), 60–69 (2011)
54. Shostack, A.: Experiences threat modeling at Microsoft. In: CEUR Workshop Proceedings vol. 413, pp. 1–11 (2008)
55. ISO/IEC: 29100:2011 + A1:2018 Information technology — Security techniques — Privacy framework. British Standards Publication (2018)
56. Ibrahim, R., Yen, S.Y.: Formalization of the data flow diagram rules for consistency check. Int. J. Softw. Eng. Appl. **1**(4), 95–111 (2010)
57. Scandariato, R., Wuyts, K., Joosen, W.: A descriptive study of Microsoft's threat modeling technique. Requirments Eng. **20**(2), 163–180 (2015). https://doi.org/10.1007/s00766-013-0195-2
58. Burns, S.: Threat Modeling: A Process To Ensure Application Security. SANS (2005). https://www.sans.org/reading-room/whitepapers/securecode/paper/1646
59. NIST SP 800-30 Guide for Conducting Risk Assessments https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-30r1.pdf. Accessed 07 Apr 2020
60. Jump, M., Finnegan, A.: Using standards to establish foundational security requirements for medical devices. Biomed. Instrum. Technol. **51**(s6), 33–38 (2017)

61. ISO/IEC: 15408-2:2008 Information technology - Security techniques - Evaluation criteria for IT security - Part 2: Security functional components. British Standards Publication (2008)
62. ISO/IEC: 15408-3:2008 Information technology - Security techniques - Evaluation Criteria for IT Security - Part 3: Security assurance components. British Standards Publication (2008)
63. NIST SP 800-53 Revision 4 Security and Privacy Controls for Federal Information Systems and Organizations Security and Privacy Controls for Federal Information Systems and Organizations. Joint task Force Transformation Initiative. U.S. Department of Commerce (2014)
64. IEC: 64223-3-3 Industrial Communication Networks - Network and System Security – Part 3-3: System security requirements and security levels. British Standards Publication (2013)

Assessing Risk Estimations for Cyber-Security Using Expert Judgment

Michael Krisper(✉), Jürgen Dobaj, and Georg Macher

Graz University of Technology, 8010 Graz, Austria
{michael.krisper,juergen.dobaj,georg.macher}@tugraz.at

Abstract. In this paper, we show a use-case of structured expert judgment to assess the risk of a cyber-security attack. We showcase the process of elicitating unknown and uncertain values using multiple experts and combining these judgments by weighing the experts based on their performance. The performance of an expert is assessed using the information and calibration score calculated from the judgments of calibration questions. The judgments are stated with three-points estimates of minimum, most likely, and maximum value, which serve as input for the PERT probability distribution. For the use-case, the input values frequency, vulnerability, and impact were asked. The combined results are propagated along an attack path to calculate the risk of a cyber-security attack. This was done using RISKEE, a tool for assessing risk in cyber-security and implementing the combination of expert judgments and propagation of the values in an attack-tree. It uses an attack graph to model the attack paths and applies probability distributions for the input values to consider the uncertainty of predictions and expert judgments. We also describe experiences and lessons-learned for conducting an expert elicitation to acquire input values for estimating risks in cyber-security.

Keywords: Risk assessment · Expert elicitation · Cyber-security · Expert judgment · Probabilistic methods

1 Introduction

Assessing the risk of cyber-security becomes more and more important nowadays. With the rise of IoT (Internet of Things) and the evergrowing number of devices connected to the internet, the possibility of hacker attacks is ever increasing. Smart housing, connected vehicles, and the general trend of all devices being able to communicate via wireless technologies (like WiFi, Bluetooth, Zigbee), increase the complexity tremendously. Besides being convenient and useful to the user and operators, this high connectivity at the same time widens the attack surface and invites hackers to use exploits in order to gain control over the devices. With such control, hackers could steal data, money, and identities, install ransomware,

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 120–134, 2020.
https://doi.org/10.1007/978-3-030-56441-4_9

deploy botnets, or even destroy hardware. This goes as far as committing terrorist attacks if they target critical infrastructures like the power grid, medical infrastructures, public transportation, or governmental institutions.

The problem is that risk for cyber-security is immensely difficult to estimate and even harder to predict for the future [25,42]. In safety, experts can extrapolate from past events and apply statistical methods to calculate future risk, cyber-security risk is behaving more chaotic and, therefore, can not be easily modeled [12]. The last resort, after not being able to rely on historical or reliability data, is to fall back to expert judgment. However, this also has its flaws. Humans have cognitive biases [18,33,34] and especially regarding probabilities we are really bad in estimating [23,24]. It was even shown that the majority of experts have terrible skills in giving quantitative judgments [3]. The topic of expert elicitation and how to get the best out of the combination of several expert judgments in the form of opinion pools is discussed for many years now [7,10]. We tackle this topic by combining several techniques from different domains and applying it to the cyber-security domain. We can calculate the risk of attacks by using attack trees and combining them with probability distributions and uncertainty propagation for the properties of attack frequency, attack probability, and impact. Furthermore, we apply structured expert judgment using the IDEA protocol [19,20] having two rounds of expert elicitation and doing performance-based weighting based on previously asked calibration questions. By doing so, we get a realistic image of risk, based on the answers of the best experts and can make better-informed decisions based on that.

2 Background and Related Work

Since the upcoming of the first computers, cyber-security played an important role [47]. Bruce Schneier was one of the cyber-security pioneers coining the term attack trees [46]. From this idea on the method of Attack Tree Analysis (ATA) [2,48] was developed and from that several other ways to describe multi-step cyber attacks were invented, like the Kill Chain [26] or Bayesian Attack Graphs [43]. In our work, we build upon this technique of modeling attacks in an attack graph by describing every step an attacker has to undertake to reach a goal. Furthermore, we combine that with quantifiable risk values like the factor analysis of information risk (FAIR) [17] does. For the evaluation of the survey in this paper, we developed and used the tool RISKEE [36] (coined from the fundamental structure of our approach: the risk tree). This is an ongoing approach to create a method for integrating safety and security risk in a holistic way [13].

Still open was the question of how to get the actual input data. Although there are many existing databases about security vulnerabilities like the Common Vulnerability Scoring System (CVSS) [15,39] or Common Criteria, and there are approaches on extracting the attack probabilities from them [21], it is still difficult to apply them to whole systems and calculate the risk for multistep attacks [25]. There are many standards for risk assessment, which we looked into: NIST SP-800 [32], NIST CSF [41], ISO/IEC 27000 [29,31], ISO 31000 [30],

EVITA [44], STRIDE [22,35], COBIT [28], OCTAVE [1], SAHARA [38], and FMVEA [45]. They model risk mainly in a qualitative way using ordinal or semi-quantitative scales, which was not sufficient for our purposes. Cox et al. proved that using such scales can lead to severe problems and has many pitfalls [9,10]. That is why we fell back to the last resort: Expert judgment. We found a robust method in Structured Expert Judgment (SEJ) [7,8] by Roger Cooke from TU Delft. It works by measuring the judgment performance of experts by asking them calibration questions and calculating a calibration score (how well did the expert catch the correct value) and information score (how big or small is the uncertainty in the judgments). These scores define the weight an expert should get for the combination into a linear opinion pool.

3 Method

In this section, we describe our applied method for expert elicitation. It is based on structured expert judgment (SEJ) [7] and the IDEA elicitation procedure [19,20]. IDEA stands for Investigate, Discuss, Estimate, and Aggregate and uses the basic principles of the Delphi method [37]. In this method, probability-based estimations and calibration questions are used to calculate performance-based weights for the individual experts. Based on these weights, their actual judgments are combined accordingly to get consolidated results. These results are discussed, and a second elicitation round is undertaken to get refined results. This method combines the best from both worlds: mathematical performance-based combinations and behavioral group-think mechanisms.

In September 2019, we tested this method during a conference workshop to judge the risk of a specific cyber-security incident. This workshop was held at the EuroSPI 2019 conference in Edinburgh [51]. RISKEE, the tool used for this workshop, was also presented at this conference [36]). In total, 21 participants with mixed backgrounds and domains participated in this workshop (experts and laypeople from many domains, e.g., cyber-security, safety, automotive, medical, industry, and academia). Of those 21 participants, only 17 were eligible for being accepted, because four either did not do the calibration questions or did not finish the elicitation (due to either arriving after the calibration phase or leaving the room earlier).

4 Information and Calibration Score

In SEJ, two scores are used for judging the performance of an expert: information score and calibration score [5]. The information score assesses the uncertainty of a given expert (his "precision"), and the calibration score, how well the true value was hit (the "accuracy"). These are calculated based on e.g., three-point estimates stating the 90% confidence interval via the minimum and maximum plausible value, and the most likely value. These three estimates define a probability distribution with five areas for the true value to fall into: with 5% probability it should be lower than the minimum; with 45% probability, it should be between the minimum and the most likely, and equally with 45% between the most likely and the maximum. The remaining 5% of cases it should be larger than the maximum value. Inside these areas, SEJ assumes that the probability is uniformly distributed. This would be the ideal "reference expert". A log-likelihood χ^2-test (also known as G-test) is used to compare this to the actual realizations of an expert. The result is the calibration value and signifies how well the expert fits the assumed reference expert. For our method, however, we had to make some changes, summarized as follows: Instead of uniform areas, we decided to take the PERT probability distribution for the three-point estimate. The distribution is used heavily in project management to predict uncertain time schedules [50]. Albeit the choice for PERT, our method is not limited to this and works with other distributions also. We use the distribution to get the probability with which the true value was predicted. Combining all realizations of the calibration questions, the resulting histogram H should correlate to a normal distribution $\mathcal{N}$. This is again tested via the log-likelihood χ^2-test to calculate the calibration score, similar to how it was done in SEJ. Ultimately this results in calculating the Kullback-Leibler distance and applying this to the χ^2 distribution to get the model fit.

$$Calibration = 1 - \chi^2 \left(2 \sum H_i \log \frac{H_i}{\mathcal{N}_i} \right) \tag{1}$$

Moreover, the information score is calculated by calculating the entropy of given estimation compared to a uniform background distribution over a reasonably large range.

$$Information = \sum_{Questions} \frac{-\left(\sum \mathcal{P}_i \log \frac{\mathcal{P}_i}{\mathcal{U}_q} \right)}{\log \left(\mathcal{U}_{max} - \mathcal{U}_{min} \right)} \tag{2}$$

Calibration and information scores are then multiplied to get the final combined score. The resulting weight is then the normalized value over the sum of combined scores for all experts.

$$Combined\ Score = Calibration \cdot Information \tag{3}$$

This combination now allows us to compare the expert to each other. Experts with higher calibration and information scores get higher weights, which is what we wanted.

4.1 Procedure

The whole workshop took approximately one hour and was split up in the following steps:

1. **Introduction** (10 min) We described the general process of expert judgment using distributions ("thinking in ranges" instead of single-point estimates) and explained some basic information about risk estimation. Every participant got a unique anonymous identifier to trace the different responses for each one of them.

2. **Calibration questions** (10 min) The calibration was done with a prepared questionnaire that should be filled out by the participants.
3. **Scenario explanation** (10 min) We explained the scenario of an imaginary car rental company, which could be the victim of a key-fob-hack [52]. The actual scenario was: Imagine being the CISO (chief information security officer) of a car rental company, having a fleet of 100 Tesla Model S cars in the beautiful city of Leuven (the city of the COSIC research group). What are the risks regarding the previously mentioned key-fob hack?
4. **First Round of expert elicitation** (5 min) Participants had to fill in the attack frequency, the attack success probability, and the resulting impact of a successful attack in the form of [minimum-mode-maximum]. This was done individually and without prior discussion.
5. **Discussion** (10 min) and **interpreting the results** (5 min)
During the discussion, we clarified the scenario even more and let the participants exchange opinions and arguments. In the meantime, we analyzed the first round and showed the results to the participants.
6. **Second round of expert elicitation** (5 min) Participants were asked to judge the same scenario as in the first round, but now with their revised judgments.

4.2 Visualisation of Results: Guppy Plots

A *guppy plot*[1] is a specialized violin plot, which shows the responses of all experts in the form of probability distributions drawn horizontally along the x-axis. Figure 1 shows an example for this. The experts are listed along the y-axis, and the whole value range is displayed along the y-axis, which is scaled either linearly or logarithmically, depending on how different the magnitudes of the responses are. If the guppy plot is used to show calibration questions and answers, the true value is also indicated using a vertical line (or area) as an overlay. More examples for this in a simpler form (as vertical lines plots) can be found in [6,20].

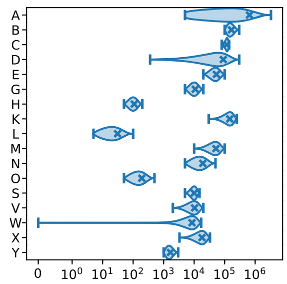


Fig. 1. An example guppy plot on a logarithmic scale.

4.3 Calibration Questions

The calibration questions have the purpose of examining the judgment quality of the experts. Based on the judgment quality, the assessor can calculate weights for the individual experts to combine the answers for the real questions in a

[1] The term "guppy plot" was coined by the wife of one of the authors, since they look like a group of guppy fish swimming in a fish tank. This is especially recognizable on logarithmic scales.

reasonable way. The calibration questions must fulfill three criteria to get a reasonable weighting of experts:

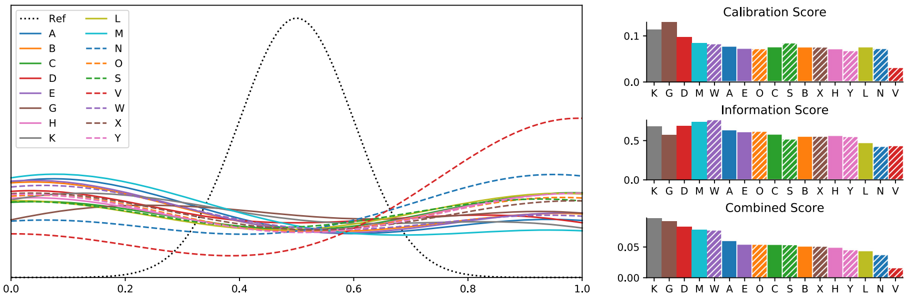
Firstly, they have to be on the same topic and ask for approximately the same facts as the real survey questions, so that the expert needs to apply the same know-how and skills to answer the real questions. This is important to assure that the performance of the calibration questions is transferable to the performance of the real survey questions.

Secondly, the answer must be unknown to the experts at the time of the elicitation. This is important because if the expert already knows the answers, the calibration does not assess the appropriate skills, and the weights are not applicable. If an expert already would know the answers, the elicitation would not be needed.

Thirdly, the calibrations questions should involve some uncertainty in order to assess their ability to judge it. Since the real survey questions inherently also involve uncertainty, it is essential to assess the skills of estimations and dealing with uncertain predictions.

5 Results of the Calibration

In this section, we visualize and discuss the results of the calibration questions and the actual survey results. Figure 2 shows the scores of calibration and information as well as the combined score for the participants. This combined score defines the weights to be applied for the actual survey results. The final weights are shown in Table 1.



(a) The calibration distributions for the participants in this survey. The dotted reference curve shows the expected results of a good expert.

(b) The resulting scores for all participants.

Fig. 2. The graph shows the performance of the participants during the calibration.

The results showed that the participants were uncertain and somewhat inadequate at estimating. This is of no surprise because Colson et al. [3,4] already showed that most experts are wrong at estimating. Furthermore, the survey participants had many different backgrounds and knowledge; therefore, it is quite natural to have vastly different opinions and responses. The reference curve in Fig. 2a depicts how the responses by a good participant are expected to look like. A good expert is one who hits the most-likely value most of the time. It can be seen that the participants in our survey did not match this expectation. The calibration distribution also shows which participants tend to over- or underestimate the true values.

Table 1. The weights for the participants in decreasing order. These weights were calculated based on the performance on the calibration questions.

K	G	D	M	W	A	E	O	C	S	B	X	H	Y	L	N	V
0.10	0.09	0.08	0.08	0.08	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.04	0.04	0.02

Calibration Results for Attack Frequency. The results for the calibration on attack frequency are visualized in Fig. 4. They show that the surveyed participants greatly disagreed and are very uncertain about the actual values, which resulted in quite a disperse distribution in the end. Participants V, S, and O have a zero calibration score since they never hit the correct value. Therefore their combined score also results in zero. The other participants all had approximately the same calibration and information scores (as can be seen in Fig. 3). The questions were the following:

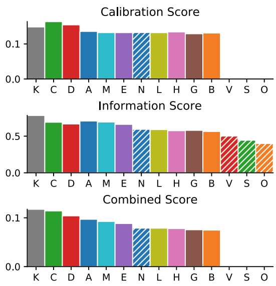
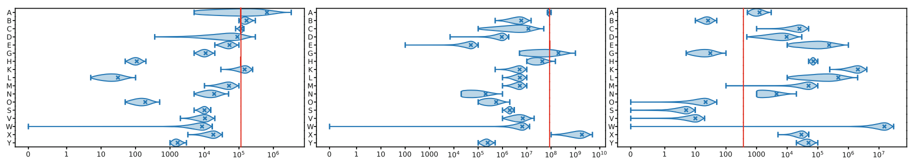


Fig. 3. The scores for the attack frequency.

Q1 How many cars are stolen per year in the UK? (3.3 Million registered cars) → 114 660 (Source: Statista [11])
Q2 How many people were affected by the Cambridge Analytica Facebook data leak in 2018? → 86 600 000 (Source: Facebook [40])
Q3 How many bigger healthcare-specific data breaches happened in 2018 in the US? → 371 (Source: U.S. Department of Health and Human Services [49])



(a) Question 1 (b) Question 2 (c) Question 3

Fig. 4. The responses to the first three calibration questions, which assess the attack frequency showing rather small uncertainties but great disagreement amongst the participants.

Calibration Results for the Vulnerability. The results for the vulnerability calibration questions again show that participants are quite uncertain about the values. Furthermore, the resulting calibration scores in Fig. 5 show that no participant was significantly better than the other, as all calibration scores are about the same at the value range of 0.1. However, regarding the information score, they differed. Participant W and K had slightly better results than the others. The resulting combined score is still a quite linear slope from the best participants W and K to the worst participants N and H. Due to space limitations, we left out the visualization of the calibration questions. In summary, they showed huge uncertainty amongst the participants.

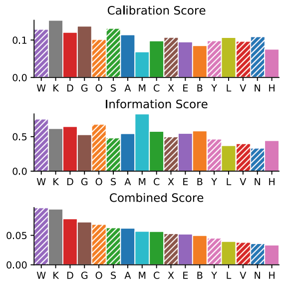


Fig. 5. The scores for the vulnerability questions.

Calibration Results for the Impact. Figure 7 shows some of the responses for the calibrations questions regarding the impact. It can again be seen that the participants are quite uncertain about the judgments (especially participant W). Three participants (W, V, S) did not have calibration scores since they never hit the actual value. For the others, the scores were quite evenly distributed (see Fig. 6). Interestingly, when looking at the combined score, the participants G, K, and M had slightly better scores than the others. The visualization for the impact was often heavily distorted by the difference in magnitudes of values (even when using logarithmic scales). Here are some questions which have been asked during the calibration:

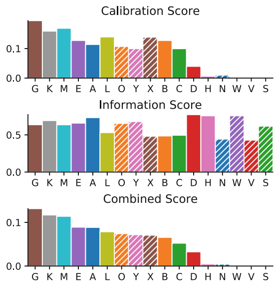
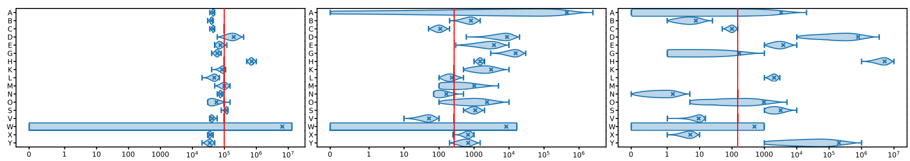


Fig. 6. The scores for the impact calibration.

- Q12 What is the price for a new Tesla Model S car? → 99000 (Source: Energysage [14])
- Q13 What is the cost of replacing a compromised electronic door lock in an office building? → € 270 (Source: Fixr [16])
- Q14 What is the estimated cost per record of a leaked user-sensitive data entry? → €135 (Source: IBM/Ponemon [27])



(a) Question 12 (b) Question 13 (c) Question 14

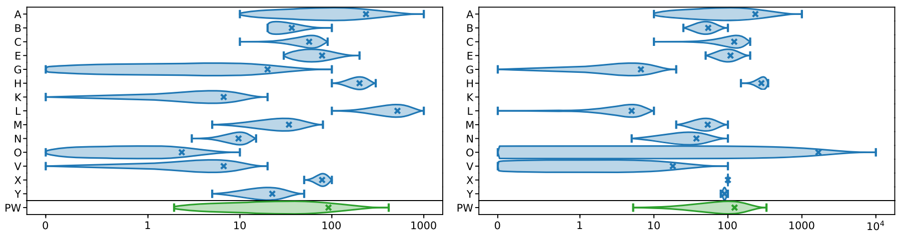
Fig. 7. The first three responses to the calibration questions about impact, showing a high variety of outcomes.

6 Results of the Risk Survey

Attack Frequency. The attack frequency defines how many attack attempts are made per timeframe (per year in our case). On comparing the results of the first elicitation round to the second, it can be concluded that the participants increased their minimum slightly and most-likely greatly but also significantly lowered the overall maximum of attacks per year after the discussions. This leads to a much higher concentration in the upper quantiles of the attack frequency distribution. Table 2 shows the actual values.

Table 2. The expected attack frequency for the entry node.

Round	Attack Frequency
1	2 – 34 – 415
2	5 – 100 – 332



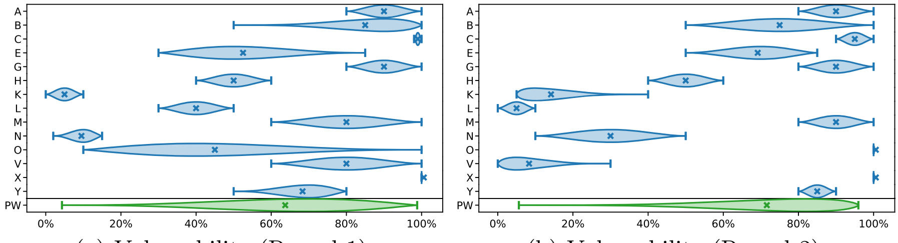
(a) Attack Frequency (Round 1) (b) Attack Frequency (Round 2)

Fig. 8. The estimations of attack frequency for the two elicitation rounds. The PW row on the bottom is the performance-weighted combined score.

Vulnerability. The vulnerability was assessed for four nodes along the attack path. The resulting values are shown in Table 3. In the second round, the minimum and most-like values were increased, and the maximum estimation decreased slightly. This concentrated the density more in the higher quantiles of the distribution.

Table 3. The estimated vulnerabilities in percent for all four nodes along the attack path.

Node	First round	Second round
1	4.3 – 69.8 – 98.8	5.6 – 82.0 – 95.9
2	2.6 – 48.2 – 98.8	4.6 – 55.9 – 92.1
3	2.9 – 40.7 – 98.8	4.4 – 46.7 – 98.9
4	1.3 – 15.7 – 96.7	3.1 – 14.4 – 96.9



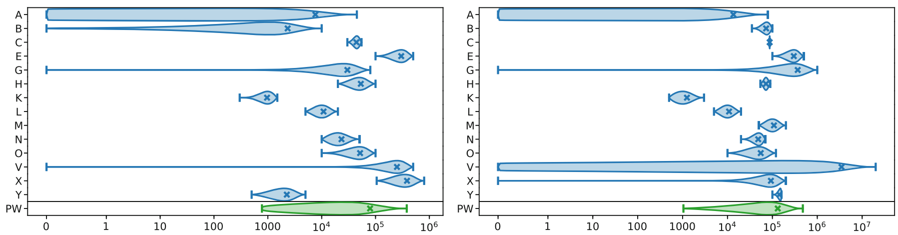
(a) Vulnerability (Round 1) (b) Vulnerability (Round 2)

Fig. 9. The estimations of the vulnerability for the two elicitation rounds. The PW row on the bottom is the performance-weighted combined score. For the sake of conciseness and space only one of the four responses is shown.

Impact. The impact defines the financial loss in the case the whole attack is executed successfully, which in our case meant that a car gets stolen. In the second round, the participants increased their estimations, which led to an increase in the minimum value, most-like value (more than double), and maximum value (Table 4).

Table 4. The expected impact in Euro if the whole attack is successful. The values are rounded for simplification sake.

Round	Impact
1	€ 780 – 23 500 – 380 000
2	€ 1000 – 78 000 – 477 000



(a) Impact (Round 1) (b) Impact (Round 2)

Fig. 10. The estimation of the impact for the two elicitation rounds. The PW row on the bottom is the performance-weighted combined score.

6.1 Final Risk Results

After weighing the participants by their performance during the calibration and using these values to calculate the risk along the risk tree with RISKEE, the final results depict the total risk of this cyber-security use-case. The results are shown in Fig. 11 and Fig. 12: The expected attack frequency will around 123 attacks per year (node 1), and the loss impact of a single attack being successful is around €131 800 (node 4).

Regarding the expected risk, there is a 95% chance that the loss will exceed approx. €24 000, the average expected loss will be around €868 800 (having a chance of 30%), and there is a 5% chance that the loss will even exceed around €3 Million. The value of having a 50% chance of exceeding is around €420 500, which is less than half of the average value. Furthermore, of the four attack steps, the first one is the most vulnerable (with 72% probability), while the last one is assumed to be the most difficult (with just 26% probability). This can be explained by the fact that in the first step, the attacker had only to be in the vicinity of the car, while in the last step the attacker has to unlock the car, enter it, start it, and drive away, which would be much more noticeable than to pass by.

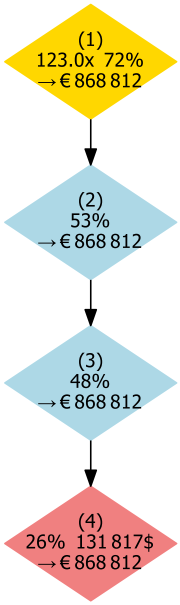


Fig. 11. The attack graph.

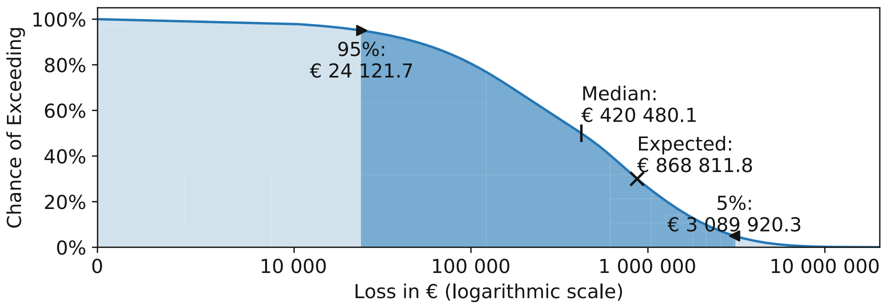
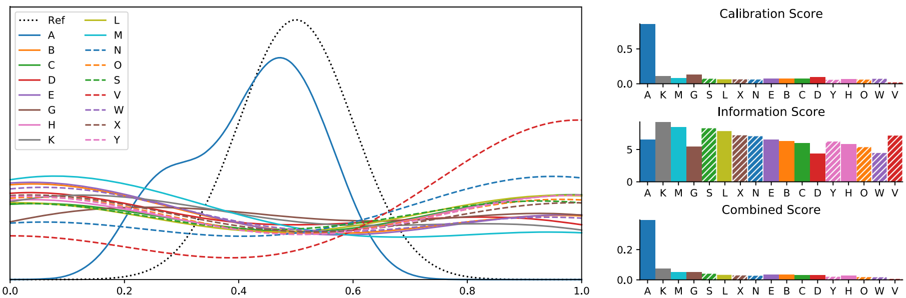


Fig. 12. The loss exceedance curve shows the probabilities of exceeding a certain amount of money along the whole range of possible outcomes. The values for the 95%, the median, the 5%, and the expected average are shown.

7 Discussion

In this survey, the scores are evenly distributed, which means that the participants performed equally on the calibration questions. The only problem here is that the calibration scores were very low. Figure 13 shows how this survey would have looked like if at least one participant had excellent responses. It can be seen that the weight for this participant would have been much higher than the other participants, and in such a way, this one would have overruled the other participants, which is desired behavior. A good expert should get more weight than a bad expert.



(a) The calibration distribution for the responses. A good participant would approximate the reference curve.

(b) The calibration score and combined score for the good participant dwarfs the scores of the others participants.

Fig. 13. The responses of participant A with artificial data representing the expected responses. It shows that a good expert would significantly overrule the other participants.

We assessed the participants' general estimation capability by combining all 14 calibration questions to get calibration and information scores. During analyzing the data, we realized that it would also have been interesting to assess the performance of the individual areas (attack frequency, vulnerability, and impact). For future surveys, we plan to assess the individual areas in more detail.

8 Conclusion

In this paper, we described a use-case of an expert elicitation for the assessment of cyber-security risks. We applied the method of structured expert judgment to get the input values, and used the tool RISKEE for uncertainty and risk propagation on risk trees to calculate the resulting loss expectancy. This served as a feasibility study to see if the method applies to real-world use-cases for estimation of risk in cyber-security. Our results support the argument of existing literature that people are bad in estimating uncertain values. Furthermore, they show that a risk assessment approach is feasible and can be done quite easily. The resulting loss exceedance curve proved to be an appropriate visualization of risk, because it showed a fully informative view over the whole risk space, while still being easy to comprehend.

In future work, we intend to improve the tools for the elicitation to streamline input and usage. Furthermore, we want to elaborate on different aspects of the risk analysis, e.g., modeling risk over time, or improving the results even more by de-biasing the experts based on their calibration which could ultimately pave the way for a self-calibrating expert judgment method that corrects the biases and errors of experts as good as possible to get even better prediction results.

References

1. Alberts, C., Dorofee, A., Stevens, J., Woody, C.: Introduction to the octave approach: Technical report, Defense Technical Information Center, Fort Belvoir, VA, August 2003. https://doi.org/10.21236/ADA634134. http://www.dtic.mil/docs/citations/ADA634134
2. Ammann, P., Wijesekera, D., Kaushik, S.: Scalable, graph-based network vulnerability analysis. In: Proceedings of the 9th ACM Conference on Computer and Communications Security - CCS 2002, p. 217. ACM Press, Washington (2002). https://doi.org/10.1145/586110.586140
3. Colson, A.R., Cooke, R.M.: Cross validation for the classical model of structured expert judgment. Reliab. Eng. Syst. Saf. **163**, 109–120 (2017). https://doi.org/10.1016/j.ress.2017.02.003
4. Colson, A.R., Cooke, R.M.: Expert elicitation: using the classical model to validate experts' judgments. Rev. Environ. Econ. Policy **12**(1), 113–132 (2018). https://doi.org/10.1093/reep/rex022. https://academic.oup.com/reep/article/12/1/113/4835830
5. Colson, A.R., et al.: Quantifying uncertainty about future antimicrobial resistance: Comparing structured expert judgment and statistical forecasting methods. Plos One **14**(7), e0219190 (2019). https://doi.org/10.1371/journal.pone.0219190

6. Cooke, R.M.: Quantifying uncertainty on thin ice: expert judgement assessment. Nat. Clim. Change **3**(4), 311–312 (2013). https://doi.org/10.1038/nclimate1860. http://www.nature.com/articles/nclimate1860
7. Cooke, R.M.: Experts in Uncertainty: Opinion and Subjective Probability in Science. Environmental Ethics and Science Policy Series. Oxford University Press, New York (1991)
8. Cooke, R.M., Goossens, L.L.: TU Delft expert judgment data base. Reliab. Eng. Syst. Saf. **93**(5), 657–674 (2008). https://doi.org/10.1016/j.ress.2007.03.005
9. Cox, A.L.: What's wrong with risk matrices? Risk Anal. **28**(2), 497–512 (2008). https://doi.org/10.1111/j.1539-6924.2008.01030.x
10. Cox, L.A.: Risk Analysis of Complex and Uncertain Systems. International Series in Operations Research & Management Science, vol. 129. Springer, New York (2009). https://doi.org/10.1007/978-0-387-89014-2
11. Clark, D.: Number of motor vehicle theft offences recorded in England and Wales from 2002/03 to 2018/19, July 2019. https://www.statista.com/statistics/303551/motor-vehicle-theft-in-england-and-wales/
12. Deb, A., Lerman, K., Ferrara, E.: Predicting cyber events by leveraging hacker sentiment. Information **9**(11), 280 arXiv: 1804.05276 (2018). https://doi.org/10.3390/info9110280
13. Dobaj, J., Schmittner, C., Krisper, M., Macher, G.: Towards integrated quantitative security and safety risk assessment. In: Romanovsky, A., Troubitsyna, E., Gashi, I., Schoitsch, E., Bitsch, F. (eds.) SAFECOMP 2019. LNCS, vol. 11699, pp. 102–116. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-26250-1_8
14. EnergySage: What are typical Tesla car prices? Model S, Model X crossover and Model 3 costs explained. https://news.energysage.com/how-much-does-a-tesla-cost/
15. FIRST: Common Vulnerability Scoring System version 3.1 Revision 1 (2019). https://www.first.org/cvss/v3-1/cvss-v31-specification_r1.pdf
16. FixR: Smart Lock Installation Cost. https://www.fixr.com/costs/smart-lock-installation
17. Freund, J.: Measuring and Managing Information Risk: A FAIR Approach, p. 00000. Butterworth-Heinemann, Amsterdam (2015)
18. Gilovich, T., Griffin, D.W., Kahneman, D. (eds.): Heuristics and Biases: The Psychology of Intuitive Judgment. Cambridge University Press, Cambridge, New York (2002)
19. Hemming, V., Burgman, M.A., Hanea, A.M., McBride, M.F., Wintle, B.C.: A practical guide to structured expert elicitation using the IDEA protocol. Methods Ecol. Evol. **9**(1), 169–180 (2018). https://doi.org/10.1111/2041-210X.12857
20. Hemming, V., Walshe, T.V., Hanea, A.M., Fidler, F., Burgman, M.A.: Eliciting improved quantitative judgements using the IDEA protocol: a case study in natural resource management. Plos One **13**(6), e0198468 (2018). https://doi.org/10.1371/journal.pone.0198468
21. Houmb, S.H., Franqueira, V.N., Engum, E.A.: Quantifying security risk level from CVSS estimates of frequency and impact. J. Syst. Softw. **83**(9), 1622–1634 (2010). https://doi.org/10.1016/j.jss.2009.08.023
22. Howard, M., LeBlanc, D.: Writing Secure Code, 2nd edn. Microsoft Press, Redmond (2003)
23. Hubbard, D.W.: The Failure of Risk Management: Why It's Broken and How to Fix it. Wiley, Hoboken (2009). oCLC: ocn268790760
24. Hubbard, D.W., Seiersen, R.: How to Measure Anything in Cybersecurity Risk, p. 00000. Wiley, Hoboken (2016)

25. Husak, M., Komarkova, J., Bou-Harb, E., Celeda, P.: Survey of attack projection, prediction, and forecasting in cyber security. IEEE Commun. Surv. Tutor. **21**(1), 640–660 (2019). https://doi.org/10.1109/COMST.2018.2871866. https://ieeexplore.ieee.org/document/8470942/
26. Hutchins, E.M., Cloppert, M.J., Amin, R.M.: Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains, p. 14, January 2011
27. IBM: How much would a data breach cost your business? (2019). https://www.ibm.com/security/data-breach
28. ISACA: COBIT — Control Objectives for Information Technologies — ISACA (2019). https://www.isaca.org/resources/cobit
29. ISO: ISO 27000 - ISO 27001 and ISO 27002 Standards (2019). http://www.27000.org/
30. ISO/IEC: ISO 31000:2018 Risk management - Guidelines (2018). https://www.iso.org/iso-31000-risk-management.html, 00000
31. ISO/IEC: ISO/IEC 27000:2018 (2018)
32. Joint Task Force Transformation Initiative: Guide for conducting risk assessments. Technical report. NIST SP 800–30r1, National Institute of Standards and Technology, Gaithersburg, MD (2012). https://doi.org/10.6028/NIST.SP.800-30r1, https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-30r1.pdf
33. Kahneman, D., Frederick, S.: Representativeness revisited: attribute substitution in intuitive judgment. In: Gilovich, T., Griffin, D., Kahneman, D. (eds.) Heuristics and Biases, pp. 49–81. Cambridge University Press, 1 edn. (2002). https://doi.org/10.1017/CBO9780511808098.004
34. Kahneman, D., Tversky, A.: Subjective probability: a judgement of representativeness. Cogn. Psychol. **3**(3), 430–454 (1972)
35. Kohnfelder, L., Garg, P.: The threats to our products, p. 8 (1999)
36. Krisper, M., Dobaj, J., Macher, G., Schmittner, C.: RISKEE: a risk-tree based method for assessing risk in cyber security. In: Walker, A., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 45–56. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_4
37. Linstone, H.A., Turoff, M. (eds.): The Delphi Method: Techniques and Applications. Addison-Wesley, Reading [usw.] (1975). oCLC: 251991541
38. Macher, G., Sporer, H., Berlach, R., Armengaud, E., Kreiner, C.: SAHARA: a security-aware hazard and risk analysis method, pp. 621–624. In: IEEE Conference Publications (2015). https://doi.org/10.7873/DATE.2015.0622
39. Mell, P.M., Scarfone, K., Romanosky, S.: A Complete Guide to the Common Vulnerability Scoring System Version 2 (2007)
40. Schroepfer, M.: An Update on Our Plans to Restrict Data Access on Facebook, April 2018. https://about.fb.com/news/2018/04/restricting-data-access/
41. National Institute of Standards and Technology: Framework for Improving Critical Infrastructure Cybersecurity, p. 41 (2014)
42. Oliver Wyman Forum: Cybersecurity Why Is It So Hard And Getting Harder? (2019). https://www.oliverwymanforum.com/cyber-risk/2019/sep/why-is-cybersecurity-so-hard-and-getting-harder-what-can-be-done.html
43. Poolsappasit, N., Dewri, R., Ray, I.: Dynamic security risk management using Bayesian attack graphs. IEEE Trans. Dependable Secure Comput. **9**(1), 61–74 (2012). https://doi.org/10.1109/TDSC.2011.34
44. Ruddle, A., et al.: EVITA D2.3 v1.1 (2009)

45. Schmittner, C., Gruber, T., Puschner, P., Schoitsch, E.: Security application of failure mode and effect analysis (FMEA). In: Bondavalli, A., Di Giandomenico, F. (eds.) SAFECOMP 2014. LNCS, vol. 8666, pp. 310–325. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10506-2_21
46. Schneier, B.: Attack trees - modeling security threats. Dr. Dobb's J. **24**(12), 21–29 (1999)
47. Yost, J.R.: The origin and early history of the computer security software products industry. IEEE Ann. Hist. Comput. **37**(2), 46–58 (2015). https://doi.org/10.1109/MAHC.2015.21
48. Tidwell, T., Larson, R., Fitch, K., Hale, J.: Modeling Internet attacks. In: Proceedings of the 2001 IEEE Workshop on Information Assurance and Security, p. 7 (2001)
49. U.S. Department of Health and Human Services Office for Civil Rights: Breach Portal: Notice to the Secretary of HHS Breach of Unsecured Protected Health Information. https://ocrportal.hhs.gov/ocr/breach/breach_report.jsf
50. Vose, D.: Risk Analysis: A Quantitative Guide, 3rd edn. Wiley, Chichester, Hoboken (2008). oCLC: ocn174112755
51. Walker, A., O'Connor, R.V., Messnarz, R. (eds.): EuroSPI 2019. CCIS, vol. 1060. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5
52. Wouters, L., Marin, E., Ashur, T., Gierlichs, B., Preneel, B.: Fast, furious and insecure: passive keyless entry and start systems in modern supercars. IACR Trans. Crypt. Hardw. Embed. Syst. **2019**(3), 66–85 (2019). https://doi.org/10.13154/tches.v2019.i3.66-85. https://tches.iacr.org/index.php/TCHES/article/view/8289

Separation of Concerns in Process Compliance Checking: Divide-and-Conquer

Julieth Patricia Castellanos Ardila(✉) and Barbara Gallina

IDT, Mälardalen University, Västerås, Sweden
{julieth.castellanos,barbara.gallina}@mdh.se

Abstract. Compliance with multiple standard's reference models has the potential to improve process quality but is a challenging task faced by manufacturers in the safety-critical context. To facilitate this task, we propose a method for automated process compliance checking that can be used as a basis for decision making. Our method requires users to create a knowledge base of formalized requirements and processes checkable for compliance. In this paper, we exploit the natural separation of concerns in the state of practice to offer adequate means to facilitate the creation of the required concepts by using a divide and conquer strategy. For this, we discuss the impact of process factors in compliance assessment and provide separation of concerns based on SPEM 2.0 (Systems and Software Process Engineering Metamodel). Then, we illustrate the defined concerns and discuss our findings.

Keywords: Software process compliance checking · Safety standards · SPI · SPEM 2.0 · Separation of concerns · Divide and conquer

1 Introduction

In the safety-critical context, standards commonly prescribe requirements that include the tasks to be performed, and resources ascribed to such tasks, i.e., personnel, work products, tools, and methods, which are also framed with essential properties. With the growing software development complexity, there is a need to adequately allocated such resources during the software development lifecycle [25]. However, this task becomes difficult due to software process diversity, i.e., the simultaneous use of multiple reference models within a single project [18]. To tackle this situation, organizations produce generic software process baselines. In the analysis of these baselines, gaps to best practices could be discovered [5], and potential improvements, based on standard's information, can be performed [4]. Thus, part of the software process improvement effort required in the safety-critical context is expended in process-based compliance.

A high level of investment in process-based compliance could result in an improvement in productivity and quality, especially when there is process diversity [18]. Process-based compliance could be supported by checking that the

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 135–147, 2020.
https://doi.org/10.1007/978-3-030-56441-4_10

process used to engineer safety-critical systems fulfill the properties set down by standards at given points. The resulting compliance checking reports can be used not only to demonstrate to auditors that process plans fulfill the prescribed requirements [10], but also to discover essential improvement aspects. Thus, in previous work [1,2], we presented a method for automated compliance checking of processes plans. Our method requires users to: 1) model a formal specification of the standard requirements by using FCL (Formal Contract Logic) [11] and 2) model a specification of the process plans that are checkable for compliance, i.e., processes augmented with compliance information, by using SPEM 2.0 (Systems and Software Process Engineering Metamodel) [16]. Thus, an essential step of our method is dedicated to creating well-formed specifications.

In this paper, we aim at facilitating the creation of the specifications required for automated compliance checking. Given the natural separation of concerns in the state of practice, we try to offer adequate means to support the separated concepts based on process structure and different standards, by using a divide-and-conquer strategy. For this, we discuss the impact of process factors in compliance assessment and justify the separation of concerns based on SPEM 2.0 concepts. SPEM 2.0 is a well-defined metamodel that not only permits the modeling of software processes but also the customization of elements to provide standards-related information. Then, we illustrate the use of the defined concerns with the requirements prescribed in the railway sector. Finally, we discuss the potential benefits and implications of our work.

The paper is organized as follows. We present essential background in Sect. 2. We discuss the separation of concerns within the regulatory space in Sect. 3. We illustrate the defined concerns in Sect. 4. We discuss our findings in Sect. 5. We present related work in Sect. 6. Finally, we conclude the work and outline future work in Sect. 7.

2 Background

This section presents essential background required in this paper.

2.1 Standards in the Safety-Critical Context

Compliance with safety standards typically involves the provision of evidence regarding process plans since standards reference frameworks contain requirements that prescribe artifacts related to the planning of activities [21]. In particular, process reference models describe a set of tasks to be performed during the development of safety-critical systems. For example, ISO 26262 [14], which is the standard that applies in automotive, proposes a V-model, in which the activities related to the development of the software are contrasted with the ones related to verification and validation. The standard DO-178C [19] describes a set of objectives that implicitly define a lifecycle model. ECSS-E-ST-40C [7], which applies in space software projects, focuses on the definition of software development phases and their characteristics. In all the standards, the detailed breakdown of the work can be inferred from the requirements. Moreover, process-related standards commonly have sections in which they describe the necessary inputs and the mandatory outputs of the safety lifecycle phases. The qualification of personnel may vary from one standard to the other. ISO 26262 mentions the importance of qualified personnel, but it leaves the decision to the company, which should have a minimum set of internal requirements in that matter. In ECSS E-ST-40C, the degree of independence between developers and reviewers is highlighted. In DO-178C, specific roles are defined for specific phases in the lifecycle. Similarly, tool qualification is required in the safety-critical context. In ECSS-E-ST-40C, supporting tools are a customer/supplier agreement that shall be documented in the plan. A specific standard annex, called DO-330 [20], defines that for Avionics, the tool qualification is in itself a process necessary to obtain qualification credit. For ISO 26262, evidence regarding the tool suitability for specific uses should be shown. All the standards prescribe methods and techniques that should be used to perform specific tasks in the form of guidance.

2.2 CENELEC EN 50128

CENELEC EN 50128 [3], which is the standard that focuses on software aspects regarding control and protection applications in railways, prescribes requirements that target the different elements described in Sect. 2.1. In Table 1, we recall a set of requirements that apply to the Architecture and Design Phase.

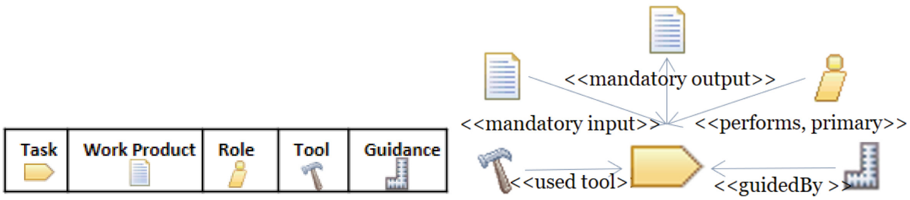
Table 1. CENELEC EN 50128-Architecture and Design Phase [3]

Element	Description
Inputs	Software Requirements Specification (SRS)
Outputs	Software Architecture Specification (SAS), Software Design Specification (SDS), Software Interface Specifications (SIS), Software Integration Test Specification (SITS), Software/Hardware Integration Test Specification (SHITS), Software Architecture and Design Verification Report (SADVR)
Tasks	1) Software Architecture Specification, 2) Software Interface Specification, 3) Software Design Specification, 4) Selection/Creation Coding Standards, 5) Software Integration Test Specification, 6) Software/Hardware Integration Test Specification, 7) Software Architecture and Design Verification Report
Roles	Designer for task 1), 2), 3) and 4). Integrator for tasks 5) and 6). Verifier for task 7). The designer shall be competent in safety design principles and EN 50128
Tools	Suitable tools with a certificate of validation (e.g., Matlab and MS Word)
Guidelines	Guidance for the Software Architecture Specification task (req-7-3-4-5), guidance for SAS (req-7-3-4-10), guidance for the SIS (req-7-3-4-19), guidance for SDS (req-7-3-4-23), guidance for the selection/creating coding standards (req-7-3-4-25), guidance for the design method selection (req-7-3-4-28), guidance for the software integration test specification task (req-7-3-4-31), guidance for the software/hardware integration test specification (req-7-3-4-36), guidance for SHITS (req-7-3-4-37), guidance for the Software Architecture and Design verification report (req-7-3-4-42)

CENELEC EN 50128 also refers to quality management and continuous improvement of the systems within the organizations. Companies may have quality assurance mechanisms that conform to different frameworks such as Software Process Improvement and Capability Determination (SPICE), also known as ISO/IEC 15504. In particular, part 5 [13] provides processes that serve primary parties during the lifecycle of software. We select the process outcome *database design*, as an example. Process outcomes are essential for determining the result of the execution of the process.

2.3 Software Processes and SPEM 2.0

A software process [8] provides a broad and comprehensive concept to frame and organize the different tasks required during the development of software. To ensure understanding, documentation, and exchange of process specifications, technological support is required [9]. In particular, SPEM 2.0 (Systems and Software Process Engineering Metamodel) [16] is a software process modeling language that provides the capability of modeling method content independently from their use in processes. Method content describes different process elements as presented in Fig. 1a. Such elements are related to each other, as presented in Fig. 1b. EPF (Eclipse Process Framework) Composer [6], which was recently migrated from Eclipse Galileo 3.5.2 to Eclipse Neon 4.6.3. [15], provides the environment for modeling SPEM 2.0-like process models.



(a) SPEM 2.0 Elements (b) Elements Relationships

Fig. 1. Content elements definitions in SPEM 2.0 [16].

2.4 Formal Contract Logic

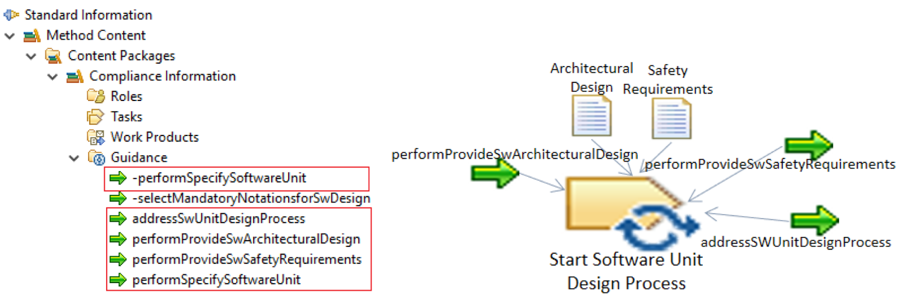
Formal Contract Logic (FCL) [11] is a logic that supports the modeling of norms representing obligations and permissions in a normative context that can be defeated by evolving knowledge. Thus, FCL is classified as a defeasible deontic logic. In FCL, a rule has the form: r: $a_1, ..., a_n \Rightarrow c$, where r is the rule identifier, $a_1, ..., a_n$ are the propositions that represent the conditions of the applicability of the norm, and c is the concluding proposition that contains normative effects.

2.5 Automatic Compliance Checking Method

Our method for automated compliance checking of processes plans [1], requires users to model processes with SPEM 2.0 (recalled in Sect. 2.3) and formalize standards requirements with FCL (recalled in Sect. 2.4). Rules in FCL are composed of propositions that correspond to the properties described in the requirements of the standard. Such properties can be annotated to the process tasks that fulfill them. Annotations reflect not only the state of the task but also the effects such task produces on subsequent tasks. For this reason, FCL propositions describe compliance effects, which annotated on process models permit the derivation of process models checkable for compliance (compliance state representation of such processes that permits automatic reasoning). We explain the modeling part of our method with an example from ISO 26262 presented in [1]. The modeled requirement is obtained from part 6 clause 8, number 8.1, which states: *"Specify software units in accordance with the architectural design and the associated safety requirements"*. The FCL representation of this requirement is presented in Eq. 1.

$$\begin{aligned} r2.1 : addressSwUnitDesignProcess \Rightarrow [O] - performSpecifySwUnit \\ r2.2 : performProvideSwArchitecturalDesign, performProvideSwSafetyRequirements \\ \Rightarrow [P] performSpecifySwUnit \\ r2.2 > r2.1 \end{aligned} \tag{1}$$

To create the process models checkable for compliance, we fist need to model the compliance effects described in the propositions of the rules. For example, the rules on Eq. 1 contains five propositions, namely addressSWUnitDesignProcess, -performSpecifySwUnit, performProvideSwArchitecturalDesign, performProvideSwSafetyRequirements and performSpecifySwUnit, which are presented in Fig. 2a. Then, we need to assign such compliance effects to the tasks that fulfill them. For example, the task Start software Unit Design Process indicates that the process is performed and has two inputs. Therefore the annotated compliance effects are addressSwUnitDesignProcess, performProvideSwArchitecturalDesign and performProvideSwSafetyRequirements (see Fig. 2b). The reader can discover more details about the previous modeling in [1].



(a) Compliance Effects. (b) Annotated Task.

Fig. 2. Method for automatic compliance checking: the modeling part.

2.6 Separation of Concerns: Divide-and-Conquer Strategy

The Romans had a strategy called divide-and-conquer, which considers that one power breaks another power into more manageable pieces to easier take control. In software engineering, this strategy is adopted as a principle to manage complexity [23]. Particularly, divide-and-conquer is seen in the principle of separation of concerns [24], which refers to the ability to separate and organize only those parts (or properties) of a system that are relevant to a particular concept or to a particular goal. A concern may be defined as a functional notion or more broadly as any piece of interest or focus.

3 Separation of Concerns Within the Regulatory Space

The relationship between the requirements imposed by safety standards (recalled in Sect. 2.1) and the targeted software processes (recalled in Sect. 2.3) is complex. The reason is that a single requirement may be impacting one element in the process, causing effects to several elements. Moreover, each element in a process may be impacted by several requirements. In addition, software process diversity, as recalled in the introductory part, may lead to problems in the understanding of what is needed for managing the compliance. Thus, we have a compact set of requirements, which we need to manage appropriately. By applying the divide-and-conquer strategy, we could break down such complexity and provide a better view of the requirements.

Separation of concerns (recalled in Sect. 2.6) applied to the regulatory space could be oriented to the process-specific factors. In particular, the aspects that requirements regulate are the tasks, their specific order, the mandatory input and outputs of the tasks, the personnel performing the tasks, the tools as well as the recommended techniques that should be used to do the tasks. Thus, the concept of a task is central, to which properties such as the definition of roles, inputs, outputs, tools, and techniques must apply.

However, requirements not only define the properties of the tasks. For example, roles and tools should have a qualification. This kind of requirements does not directly affect the tasks. They directly affect other elements, which in turn have effects on tasks. Thus, a process can be deemed compliant if we can demonstrate that the process contains the permitted tasks, such tasks have associated the prescribed roles, inputs, outputs, tools, and techniques, and if the associated elements have associated their related properties. With such consideration, dividing requirements in terms of the elements they target, as well as the specific properties defined for each element, seems to be the natural way in which concerns should be separated.

According to SPEM 2.0 (recalled in Sect. 2.3), a task is performed by a role, requires inputs and provides outputs, is guided by guidance, and a tool is used (see Fig. 1b). Thus, the tasks are the central elements, to which the other elements are allocated. Our method for compliance checking (recalled in Sect. 2.5), requires that all the properties defined by the requirements of the standard are also allocated (or annotated) to the tasks included in the process plan since such annotations describe the permitted compliance states of the tasks. An abstraction of such a concept can be seen in Fig. 2b. However, not only tasks provide compliance effects to the overall process. As we previously concluded, elements different from tasks too.

Thus, we propose a new abstraction of model annotation, in which tasks will no longer be the placeholder of the compliance effects caused by the process elements ascribed to them. Instead, every element will carry out its own responsibility in terms of compliance information (see Fig. 3). The novelty of the approach is threefold. First, we free the tasks from unnecessary annotations. Second, annotations on shared process elements should be done only once in a process model. Third, annotated elements have the potential to be reused in other processes and easily re-checked.

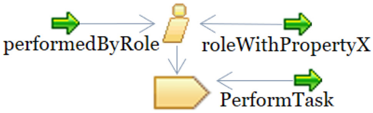


Fig. 3. Annotated role.

To facilitate the creation of compliance effects, which later can be used to form the propositions of the rules in FCL (recalled in Sect. 2.4), we propose two aspects. The first aspect is the definition of icons, which includes the description of the targeted elements, as presented in Table 2. The second aspect is the definition of templates that facilitate the creation of compliance effects, as presented in Table 3. Both, icons and templates are based on the concepts described in SPEM 2.0 in Fig. 1.

Table 2. Icons describing specific compliance effects.

Role		Work Product		Guidance		Tool		Task
Definition	Property	Definition	Property	Definition	Property	Definition	Property	
								

Table 3. Compliance effects targeting differentiated process elements

Element target	Definitional propositions	Property-based propositions
In/Output elements	provide{*Element*}	{*Element*}with{*Property*}
Roles	performedBy{*Role*}	{*Role*}with{*Property*}
Tools	used{*Tool*}	{*Tool*}with{*Property*}
Guidelines	guidedBy{*Guidance*}	{*Guidance*}with{*Property*}
Tasks	perform{*Task*}	

4 Illustrative Example

We illustrate the separation of concerns in the regulatory space by taking into account the requirements for the architecture and design phase proposed by

CENELEC EN 50128 (see Sect. 2.2). Initially, we need to classify the requirements in terms of the process elements they target. This operation is already presented in Table 1. From this division, we can describe the definitional and property-based propositions derived from these requirements by using the propositions templates shown in Table 3, and the icons described in Table 2. Then, we model them as SPEM 2.0-like elements in the guidance part of EPF Composer. Figure 4 presents the instantiation of the templates with the predefined icons. For example, the designer should be defined (definitional proposition), and the designer should have competence in safety design and EN 50128 (two property-oriented propositions). The previous propositions are highlighted in red in Fig. 4. A similar analysis is done with all the requirements.

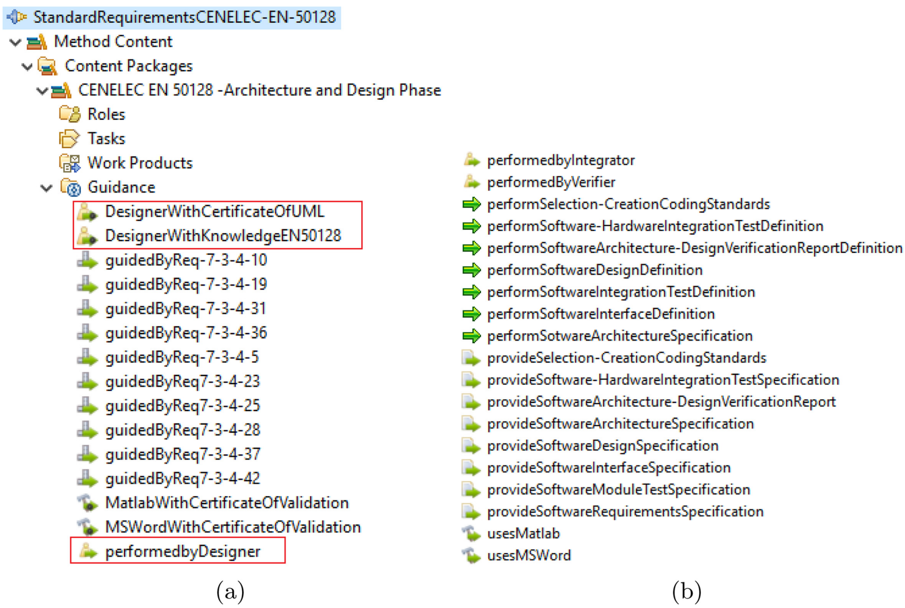


Fig. 4. CENELEC EN 50128 - Architecture and Design Phase. (Color figure online)

The next step is to annotate the compliance effects defined in Fig. 4 into a process plan. For simplicity, we described a process plan taking into account the process elements described in the standard, recalled in Table 1 (see Fig. 5). As we can see in Fig. 5, the process plan contains a series of tasks and elements ascribed to such tasks. To annotate the effect, we need to compare each element in Fig. 5 with the list of compliance effects in Fig. 4. In this case, the names of the process elements can be found in the names of the compliance effects since both models are taken from the standard. Thus, the annotation process is straightforward.

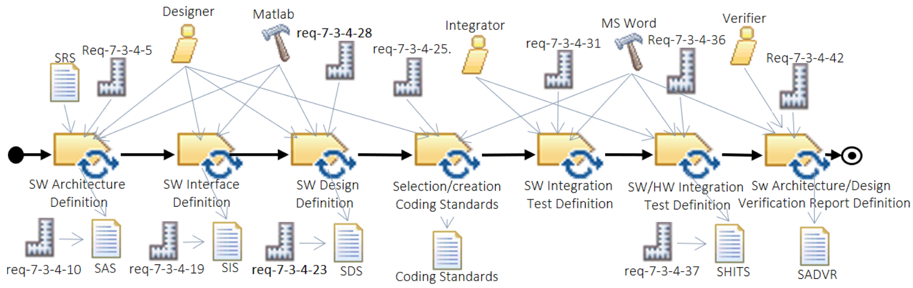


Fig. 5. Process plan targeting the Architecture and Design Phase

Figure 6 shows the annotation of the task SW Design Definition. As the figure depicts, this task has one direct CENELEC EN 50128-related compliance effect, i.e., performSoftwareDesignDefinition. The remaining eight compliance effects are allocated to the elements that directly fulfill them, e.g., the task is performed by a designer, who should have a certificate of UML, and that has knowledge of EN 50128. As we can see in Fig. 5, some tasks are done by the same role. e.g., the designer should perform the first four tasks, and the same tools should be used. Thus, our approach simplifies the annotations process since all those indirect compliance effects are not required to be annotated in each task. To make the process also compliant with ISO/IEC 15504, the outcome prescribed by the effect provideDatabase (we assume it was modeled as in Fig. 4), should be included in the modeling of the process (see the work product Database highlighted with a dotted line in Fig. 6).

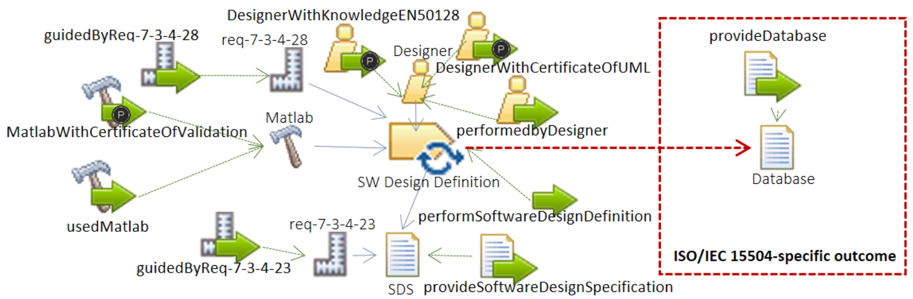


Fig. 6. Task and their ascribed elements annotated with compliance effects

5 Discussion

In this section, we present a discussion regarding our method.

5.1 Compliance-Related Process Information

Compliance management can benefit from our proposed modeling strategy. First, the icons describing definitions (see Table 2) will correspond to the software process elements required in a fully compliant process plan. Thus, such visual descriptions make the process engineer compliance-aware during software process modeling. Second, the templates presented in Table 3 aim at relating process elements with their properties. Thus, discovering the compliance effects, which the software process element produces in the context of the whole process, is facilitated. Third, as every process element carries out its compliance information, the annotation process is more efficient since it is expected that tasks share their associated elements, i.e., roles, guidance, tools, and work products (See Fig. 6). Moreover, compliance effects from different standards can be added to software process elements without limitations, helping to define multi-compliant process-checkable for compliance. Finally, we propose a standardized template-based mechanism for creating definitional and property-based compliance effects (See Table 3). Such mechanism can also be exploited for automating the generation of standard(s)-compliant process components that can be reused when assembling the processes required in different projects.

5.2 Software Process Diversity

Software process diversity is common in safety-related context, as recalled in the introductory part. Our approach implicitly takes into consideration process diversity by providing a method that facilitates the selection of compliance artifacts as needed for specific compliance frameworks. In particular, the definition of compliance effects, as presented in Fig. 4, could help in the creation of compliance artifacts from one standard, that can also be enriched with the compliance effects of related standards, as depicted in Fig. 6, for configuring process models that are multi-compliant. This aspect results in the utilization of cohesive process components that have distinctive value attributes. Besides, process components that do not receive significant levels of resource commitments in terms of compliance could be identified as potentially less useful and could be eliminated without significantly impacting project outcomes.

5.3 Relation with the SPI Manifesto

The application of standards best practices is a way to learn from the experience of the functional experts. Such experience is valuable to define and improve specific, project-oriented software processes. Our approach provides a method for deploying compliant-related pieces required for controlling knowledge across standards and projects (as presented in Fig. 4). A process engineer can play with such pieces and learn how to use them to satisfy the demands, not only of the applicable standard(s) but also the company and customer needs. In this way, our approach addresses principle 4 of SPI Manifesto [17], which states that SPI *creates a learning organization*. Moreover, having a model of the required pieces could help the definition and improvement of baseline process models (see Fig. 5). The resulting artifacts aim not only at enabling certification according to the standard but also to change existing habits in the organization, incrementing their awareness regarding best practices and therefore making the business more successful. Thus, our approach also addresses principle 6 of SPI Manifesto, which states the *use of dynamic and adaptable models as needed.*

6 Related Work

In this section, we discuss other approaches that have proposed the separation of concerns for facilitating compliance checking with FCL. In [22], four types of control tags are defined for compliance checking of business processes. These control tags consist of the state and operations of propositions about the conditions that are to be checked on a task and are typed-linked, namely control tags represent compliance effects. Such tags are: the flow tag, which represents a control that would impact on the flow of the process activities; data tag, which identifies the data retention and lineage requirements; the resource tag, which represent access, role management and authorization; finally, time tag, which represents controls for meeting time constraints such as deadlines and maximum durations. Our work, as in [22], describes the compliance effects based on the type of elements that are present in a process. However, contrary to [22], we further separate the definition of the elements from the properties allocated to these elements, i.e., we propose definitional and property-oriented compliance effects. Moreover, we provide a template for creating the compliance effect and icons that facilitate their description and its subsequent annotation in process elements. In [12], the concept of data tag described in [22] is revisited to create a methodology that permits their extraction from business process logs. Contrary to the work presented in [12], the extraction of our compliance effects is performed directly from the standards and not from process logs since we aim at having a faithful representation of the requirements prescribed by the standard at design time.

7 Conclusions and Future Work

(Process-oriented) Safety standards define process elements and their properties. Similarly, process modeling languages, such as SPEM 2.0, provide definitions that match precisely with the ones described in the standards. In this paper, we took advantage of these characteristics to offer a natural separation of concerns that could be applicable in compliance checking. From the definition of concerns, we proposed a template to describe the compliance effects that are expected from the process elements. Moreover, we proposed icons for their representation that permit their identification and annotation on the corresponding process elements. Our approach offers adequate means to support the separated concepts based on process competence and different standards, and thus it may facilitate the modeling part of our method for automated compliance checking.

Future work includes the evaluation of our approach in terms of usability. In addition, to put the approach into practice, extensions to the current algorithm used for compliance checking must be designed and implemented to permit the inclusion of co-occurrent compliance effects, which are annotated in process elements ascribed to tasks, into the compliance analysis. Moreover, to facilitate further the annotation process, algorithms that permit automatic mapping between compliance effects and company-specific processes, as well as algorithms that automatically permit the creation of process elements from the definitional compliance effects, should be created.

References

1. Castellanos Ardila, J.P., Gallina, B., Ul Muram, F.: Enabling compliance checking against safety standards from SPEM 2.0 process models. In: Euromicro Conference on Software Engineering and Advanced Applications, pp. 45–49 (2018)
2. Castellanos Ardila, J.P., Gallina, B., UL Muram, F.: Transforming SPEM 2.0-compatible process models into models checkable for compliance. In: 18th International SPICE Conference (2018)
3. CENELEC: EN 50128. Railway Applications-Communication, Signaling and Processing Systems Software for Railway Control and Protection Systems. British Standards Institution (2011)
4. Crabtree, C., Seaman, C., Norcio, A.: Exploring language in software process elicitation: a grounded theory approach. In: 3rd International Symposium on Empirical Software Engineering and Measurement, pp. 324–335 (2009)
5. Eckey, M., Greiner, C., Peisl, T.: Why do organizations focus on assessments instead of their process-improvement objectives? In: Walker, A., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 392–401. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_30
6. Eclipse: Eclipse Process Framework (EPF) Composer (2018). https://www.eclipse.org/epf/
7. ECSS: ECSS-E-ST-40C, Space Engineering Software (2009)
8. Fuggetta, A.: Software process patterns: a roadmap. In: International Conference on Software Engineering, pp. 25–34 (2000)
9. Gallina, B., Pitchai, K., Lundqvist, K.: S-TunExSPEM: towards an extension of SPEM 2.0 to model and exchange tunable safety-oriented processes. In: Lee, R. (ed.) Software Engineering Research Management and Applications, vol. 496, pp. 215–230. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-319-00948-3_14
10. Gallina, B., Ul Muram, F., Castellanos Ardila, J.: Compliance of agilized (software) development processes with safety standards: a vision. In: 4th International Workshop on Agile Development of Safety-Critical Software, pp. 1–6 (2018)
11. Governatori, G.: Representing business contracts in RuleML. Int. J. Coop. Inf. Syst. **18**, 181–216 (2005)
12. Hashmi, M., Governatori, G., Wynn, M.T.: Business process data compliance. In: Bikakis, A., Giurca, A. (eds.) RuleML 2012. LNCS, vol. 7438, pp. 32–46. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-32689-9_4
13. ISO/IEC 15504–5: Information Technology - Process assessment - An Exemplar Software Life Cycle Process Assessment model (2012)

14. ISO/TC 22/SC 32: ISO 26262: Road Vehicles Functional Safety (2018). https://www.iso.org/standard/68383.html
15. Javed, M., Gallina, B.: Get EPF composer back to the future: a trip from Galileo to Photon after 11 years. In: EclipseCon (2018)
16. Object Management Group: Software & Systems Process Engineering Meta-Model Specification. Version 2.0. (2008)
17. Pries-Heje, J., Johansen, J.: The SPI Manifesto (2009). https://2020.eurospi.net/images/eurospi/DownloadCenter/spi_manifesto.pdf
18. Ramasubbu, N., Bharadwaj, A., Tayi, G.K.: Software process diversity: conceptualization, measurement, and analysis of impact on project performance. Manag. Inf. Syst. **39**(4), 787–807 (2015)
19. RTCA/DO-178C: Software Considerations in Airborne Systems and Equipment Certification (2011)
20. RTCA/DO-330: Software Tool Qualification Considerations (2012)
21. Ruiz, A., Juez, G., Espinoza, H., de la Vara, J.L., Larrucea, X.: Reuse of safety certification artefacts across standards and domains: a systematic approach. Reliab. Eng. Syst. Saf. **158**, 153–171 (2017)
22. Sadiq, S., Governatori, G., Namiri, K.: Modeling control objectives for business process compliance. In: Alonso, G., Dadam, P., Rosemann, M. (eds.) BPM 2007. LNCS, vol. 4714, pp. 149–164. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-75183-0_12
23. Smith, D.: The design of divide and conquer algorithms. Sci. Comput. Program. **5**, 37–58 (1985)
24. Sommerville, I.: Software Engineering, 9th edn. London, Pearson (2011)
25. Yilmaz, M., O'Connor, R.V.: A market based approach for resolving resource constrained task allocation problems in a software development process. In: Winkler, D., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2012. CCIS, vol. 301, pp. 25–36. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-31199-4_3

SPI and Team Performance and Agile and Innovation

The Value of a Net Promoter Score in Driving a Company's Bottom Line: A Single-Case Study from IBM Training Services

Alexander Ziegler[1](✉) and Thomas Peisl[2]

[1] Edinburgh Napier University, Edinburgh, UK
alexander.ziegler@de.ibm.com
[2] Munich University of Applied Sciences, Munich, Germany
tpeisl@hm.edu

Abstract. The Net Promoter Score is a basic measure to assess the likelihood whether a customer would recommend an organisation to somebody else. Companies use it to drive customer multiplication. There is a vast range of academic research that details both the positive impacts and practicalities of NPS and certain instances where NPS would not be applicable. Research indicates that, although NPS isn't necessarily related to the results or success of a business, from a corporate perspective, it can provide a solid basis from which to make smart business decisions that benefit a company. We will present an exemplary case study from IBM and we will be able to demonstrate direct impacts for existing businesses. Furthermore we propose a working model that will be useful for other businesses.

Keywords: Net Promoter Score · NPS · Service business performance

1 Introduction

Created by Fred Reichheld and Bain & Company, the basic Net Promoter Score (NPS) is as follows: "On a scale of zero to 10, with 10 being highest, what's the likelihood that you would recommend us (our company) to a friend or colleague?"

There is a vast range of academic research [1–5] that details both the positive impacts and practicalities of NPS and certain instances where NPS would not be applicable. There is a lot of confusion in the industry created by articles, as recently found in Forbes magazine [6], that are only looking to the research highlighting the challenges and concluding NPS is not working.

Every business requires a trigger. The current research shows that, although NPS isn't necessarily related to the results or success of a business, from a corporate perspective, it can provide a solid basis from which to make smart business decisions that benefit a company. We will present an exemplary case study from IBM and we will be able to show direct impacts and key performance indicators of NPS for existing businesses. We will further use the results of the case study to discuss the relevance of NPS to a company's bottom line and whether there are any correlations or patterns to be found.

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 151–161, 2020.
https://doi.org/10.1007/978-3-030-56441-4_11

This means this research has two objectives to contribute to academic knowledge. We intend to:

- Create a rigorously executed single-case-study suggesting that a focus on NPS enhances service business performance
- Propose a working model that will be useful for other businesses.

In summary, the research aim of this paper is to show that NPS can be valuable if the framing of the question is extended.

The structure of the paper is as follows: After this introduction, a literature review gives an overview of the current research and concludes with the research questions. In the following section, the methods and data collection are described. The next section describes the detailed case study and the analysis and findings are summarised in the last section of this paper.

2 Literature Review

Reichheld's [7] research in 2003 is considered the foundation of NPS. According to his research, companies who want to grow their business only need one Key Performance Indicator (KPI) and not multiple. According to his research, the top-ranking question to ask clients is: How likely is it that you would recommend [company X] to a friend or colleague on a scale from 0 to 10? Then, calculate the percentage of customers who respond with nine or ten (promoters) and the percentage who respond with zero through six (detractors). The percentage of detractors is then subtracted from the percentage of promoters to arrive at your net-promoter score [7]. Reichheld states that not only is this number an indicator for growth of a company, but that no other KPIs are needed: "If growth is what you're after, you won't learn much from complex measurements of customer satisfaction or retention. You simply need to know what your customers tell their friends about you" [7].

Since 2003, several journal articles have investigated the value of NPS [2–4]. In March 2020, ProQuest shows 511 articles quoting Reichheld's article. However, the number of articles with 'Net Promoter Score' in the title in peer-reviewed sources is, according to ProQuest, below 20. There are, in general, four categories of research around NPS. Of course, there is quantitative research that states NPS works as outlined by Reichheld [4, 8, 9]. The next category is conceptual discussions that state weaknesses of NPS [3, 10] while the third category is research that questions whether NPS works, using quantitative methods [11, 12]. It needs to be highlighted that there is more than one research in this category with focus on Reichheld's assertion that NPS is the "single measurement" you need. Comparison to the American Customer Satisfaction Index (ACSI) was made by Keiningham, Cooil and Aksoy [11] who conclude that NPS is not better. However, there is also research around ACSI and NPS from East, Romaniuk and Lomax [1] who compare ACSI and NPS and conclude that both are not perfect and need an enhancement. This research [1] is, in a certain way, representative for the last category that needs attention, and shows certain weaknesses and highlights that additional questions need to be asked besides NPS [1, 2, 13, 14].

With the wide usage of NPS in the industry, it is surprising that only two academics in the above mentioned studies present case studies [9, 14]. The area of those case studies is non-profit for one of them and text-mining for the other; this means neither is in the area of software or services. This leads directly to the following research questions:

1. Are there existing data from the industry allowing to build a case study that either suggests the usage of NPS as outlined by Reichheld [7] or that indicates that a variation makes sense [1, 2, 13, 14]?
2. Is there a new business model or process that can be deduced from the findings from the first research question?

3 Research Strategy

This research follows the case study approach from Eisenhardt [15], followed by Yin [16]. From the conceptual approach, this research uses a single-case-study, following the categorisation by Yin [16] on how case studies can be structured. Going deeper into the methodology of case studies, Swanborn [17] categorised single-case studies that are based on the approach of Yin and Eisenhardt into two categories. On the one hand, there are those laying the foundation for new concepts with just a single case, while, on the other, there are single-case studies that speak on their own and, for example, underline existing hypothesis or give examples that are replicable in business. Our research, hence, uses a single case, which is an exemplary case study. The main reason for the choice of this method is the access to the data from IBM, following the approach of 'controlled opportunism' as outlined by Eisenhardt [15]. Eisenhardt justifies the use of the application of this approach if data access is available and the case is adding value, as also outlined by Swanborn [17].

On the details of the data sources, Yin [18] describes seven commonly used data sources for case studies. The authors have access to the course evaluation database of IBM, quantitative data, communication protocols and additional meeting minutes. All data are available in a case database from the authors for review on request.

There are three data sources used in this case study:

1. Quality data out of the quality database from IBM. The data are extracted out of the IBM database and stored in Excel for usage in this study and are anonymised.
2. Email exchanges around the business development meetings with the partners. The data are stored in pdf and txt format and are anonymised.
3. Meeting minutes from the author leading the business development discussions. The data are stored in pdf and txt format and are anonymised.

All data were extracted in April 2020 from the IBM sources and stored in an anonymised way in the research database and are available from the authors on request [19]. To make sure that the case is replicable, the authors followed the strict process of Eisenhardt [15]. The next section outlines in detail which questions were asked, the detailed data collection, the size of the data set, data storage as well as methods used to assure data quality.

4 The Case: Usage of NPS in IBM Training Services

IBM uses for the calculation of NPS the original definition and the question from Reichheld [7]. However, there is always in addition a second question asking for the 'why'. In detail, as second question, IBM Training Services asks: "What is the primary reason for the score you gave?"

The case study is used to draw appropriate conclusions out of IBM's adoption of NPS and some background is needed. In 2013, IBM executed a model change in their training department. In July 2013, the former profit and loss-driven training business in IBM was moved into a model based on strategic partnerships. This means IBM chose strategic partners, who are, thus, then responsible for all delivery and sales of any training. IBM's role is the creation of the course materials and the enablement of the partners. The business model, in this way, remained in place until the end of 2018 without change. The selection of the strategic partners was done in a request for proposal (RFP) process. The key prerequisites to become a training partner are the coverage of at least three continents with a legal entity and the capability of delivering training in more than 40 countries. Sub-partners are allowed in this model as long as there are legal entities on three continents of each of the strategic partners. The key requirements resulted in five partners, all having their strengths and weaknesses regarding country coverage, subject matter expertise and delivery methods. However, they are comparable from their overall delivery capability.

From a quality point of view at the introduction of the model, the key performance indicators (KPI) are the instructor and material quality measured in percent (details see Fig. 2). The material quality is not influenced by the training partners, as all materials are coming from IBM. It is a KPI for IBM to improve course material quality. The important quality factor for the partners is the instructor quality. The training partners are responsible for training and management of the instructors. Prior to introduction of the model, the instructor quality was always between 91% and 94%. Figure 1 shows the used business process for the strategic partners to manage quality based on the KPI instructor quality. The initial target set was 94%.

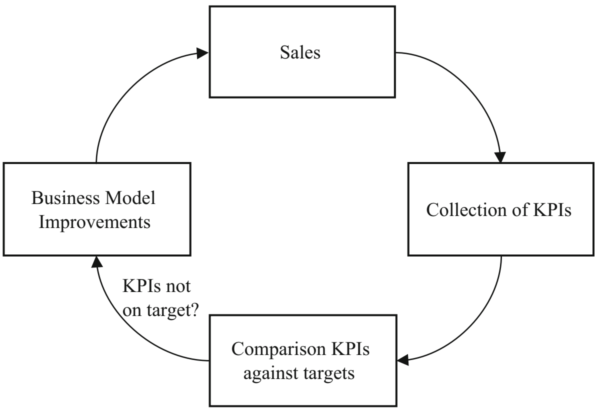


Fig. 1. Feedback model using heritage KPIs

Figure 2 shows the quality results since 2015. Partner B has been below the target, but, overall, the quality is constant on average above the 94%. Using the instructor quality as KPI, the statement can be made that there is nothing to be improved from partner point of view around quality.

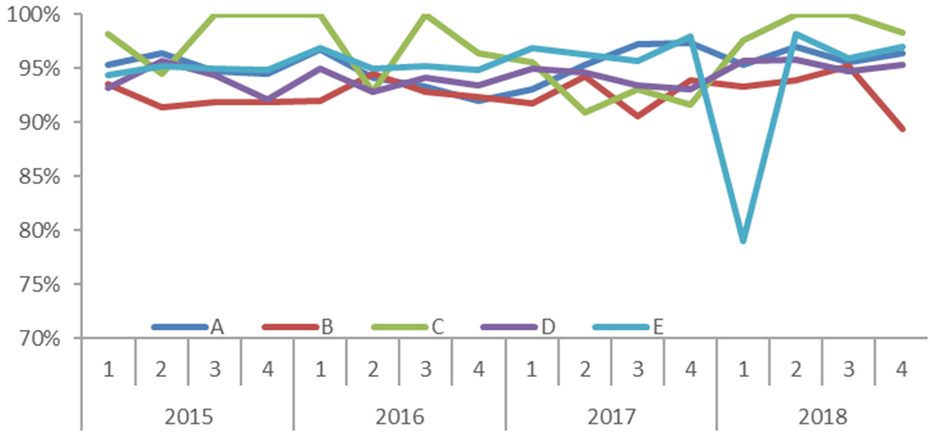


Fig. 2. Average instructor rating per quarter for training partner A-E. The source consists of the evaluation database of IBM from 2015 until 2018 with 135, 284 courses. The instructor quality is a value between 0 and 100%. At the end of a class, students are asked the question "Overall, how satisfied are you with instructor of this class (e.g. presentation, knowledge, class management)?" The following rating terminology and scale is implemented: Strongly Agree = 5, Agree = 4, Neither Agree nor disagree = 3, Disagree = 2, Strongly Disagree = 1, Not Applicable = 6. Instructor rating is calculated as: add the number of very satisfied and satisfied responses together, divide this by the total number of valid responses (only count submitted surveys with answered question).

Contrary to the quality on constant and expected level shown in Fig. 2, the data in Fig. 3 show a completely different picture. The graph shows the volume of students trained. Except for a mistake in data from B in 2015, an ongoing decline is visible. There were certain influence factors from the industry, suggesting a possibility of a decline. However, the constant decline suggested to the business leaders a deeper investigation on partner side, even with other KPIs, like instructor quality, not showing any anomalies.

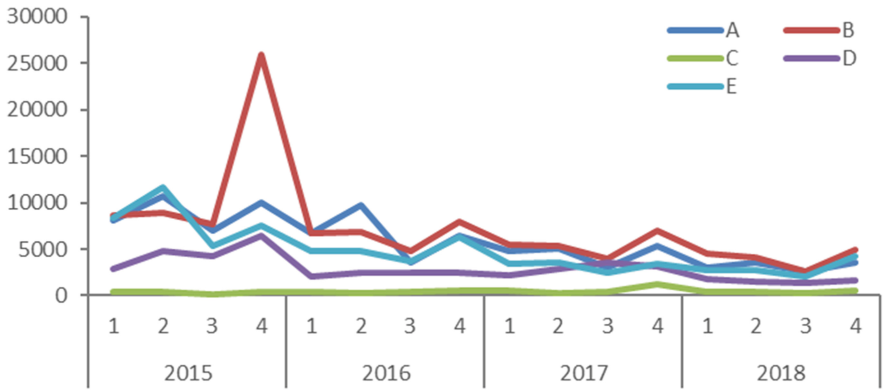


Fig. 3. Delivered training units per quarter for training partner A-E

4.1 Introduction and Management of NPS

To get more insight into the decline, the training department decided to introduce NPS in mid-2017. Based on Fig. 2, there was no need based on the instructor quality. But, as other units in IBM reported good success with improving, even while not having known problems, the introduction seemed to make sense. Even if the result would just be to confirm that everything is working well.

As all client training was delivered at that time by the training partners, there was a discussion inside the IBM training team as to how the question for NPS should be structured. On the one hand, the question could have been "Would you recommend IBM Training?" or "Would you recommend company X?" assuming the training was sold and delivered by company X. Taking the initial definition and research from Reichheld [7], which always relates the question back to the company with whom the end-user is dealing, the decision was the second question. But all the five training partners are asking the same question right at the end of the course. In addition, the question "If you would not recommend company X, could you explain why?" is added (in analogy to the question "What is the primary reason for the score you gave?" that IBM is using with their direct clients). However, to verify the results of NPS, the decision was to keep the questions for the old KPIs in place (the old KPIs were instructor and material quality, for calculation see Fig. 1) and to add NPS. The former course evaluation had many more questions, which were all removed to follow as close as possible the NPS idea that nothing else is needed.

The data collection happens in the following way. IBM is prescribing the exact wording and the evaluation scale to the partners. The partners can use any tool to collect the data. But all partners have to submit the data independent on the data collection method in an electronic way via a standardised XML feed to IBM. IBM is regularly running audits making sure that the data are correct. There is usually at least one audit during the runtime of the two year contract, but also random smaller audits are scheduled to drive maximum of data quality. The data are collected in a centralised database on a per course level. The original content of the database is used in this case study in anonymised format.

4.2 Analysis - IBMs Business Development Approach After Introduction of NPS

The business development in the first quarters was straightforward from an IBM point of view. IBM consolidated the results, and anonymised the data so that each training partner could see their own results, but could also see the NPS of the other partners in general. The data were provided per region and per quarter. In addition, the partners had access to their own comments from the additional question "What could company X improve, so that you would recommend it?" Figure 3 shows the data from the first quarters.

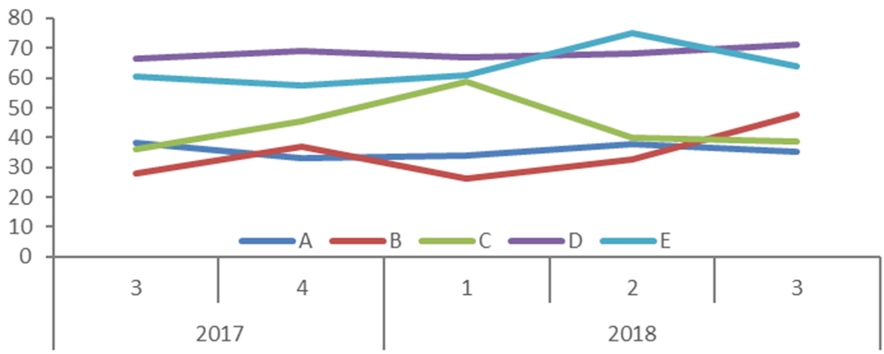


Fig. 4. NPS data per quarter for training partner A-E

IBM highlighted during regular business development discussions to partners A, B and E that their numbers were below the average of 48 of the whole timeframe, and especially lower than the average of 59 of the two top performers (C and D).

Partner B was seriously concerned around their low NPS, which did not match the very positive feedback around materials and instructors. They performed a detailed analysis for patterns in the NPS comments. Quotes like "Web book was horrible and wouldn't load properly for me to print out. Big pain, especially for exercises" or "Also the complicated way handling the material (website limited view only, though pdf format would have been more convenient)" lead to the root cause. The partner implemented a very fast change of their eBooks format, which led the NPS to increase above the average within two quarters (Fig. 4). Not only the NPS increase, but also comments like "Satisfied with the materials" were visible. However, the number of positive comments in regard to NPS is, in general, lower than the negative ones. In this case, the only change was the format of the eBooks, so that the increase of the NPS relates to this.

An interesting case regarding NPS in this data analysis is partner D. They started on a surprisingly high level. A deep dive interview resulted in the finding that the partner always asked a question around "would you recommend this company?" but the clients only had two choices: yes or no. And in the case of a 'no' the partner called the client to understand this. So, all existing clients were familiar with an NPS-like question, and went in with very high numbers. On detailed discussion, it was already visible that new clients gave NPS evaluations that were below the average of NPS from this partner.

5 Findings and Discussion

We aimed with this research to contribute to academic knowledge in two ways. On one hand creating a rigorous executed case study, and on the other hand a working model based on NPS that will be useful for other businesses.

From the point of view of the first item, the presented data of this single-case-study shows the value of confirming the findings from various researchers. On one hand, the case highlights that the number of NPS alone is not helpful [13]. The addition of the question of “If you would not recommend company X, could you explain why?” is, in this case, the key to business improvements. This is in line with the findings of [2, 13, 14].

Furthermore, the outlined approach also showed that even the pure number can be a helpful indicator. But, in this case, there must be data from comparable units (in our example five partners with similar company profiles and same targets) and the same geography. Even without the question ‘why’, it is possible to compare and judge. But complexity comes up without the ‘why’ when looking to how to improve. A company could run an investigation once NPS is indicating a problem. This means that, without an additional question, NPS can still be helpful.

In addition, the current case is an example where NPS is a much stronger indicator for needed improvements than existing industry-wide use of KPIs (KPIs rating instructor and course material quality as outlined in Fig. 2). Assuming that the instructor quality is similar to ACSI, then this case study is questioning the research from [1] who do not see NPS as being superior to existing KPIs. However, further research would be needed to compare instructor quality to ACSI. The current research shows that NPS can be superior to other KPIs, as long as it is used together with at least one additional question. As such, not fully confirming the initial research from Reichheld.

Looking into the second aim, the case study shows that NPS can be a solution in all situations where the business is not evolving as expected and existing KPIs are not giving indications for reasons. The model to improve based on normal KPIs is shown in Fig. 5.

As demonstrated in the case study of this research, NPS can give, in this context, additional detailed insight into the business and drive change [20]. Figure 5 shows when and how NPS was introduced in the current case study using generalised terms.

There are certain limitations to this research. The first one to mention is the fact that all observations are based on thousands of data records per partner. For businesses with smaller amount of data the minimum amount of data required needs to be verified. Second, all partners were very similar, as they were selected based on a business model change. There are limitations once there are partners that are not doing exactly the same, as this would put forward the question as to how comparable the data are.

Finally, there are some open questions as result of the current work. Further research could replicate the existing case and confirm the recommended model. When reflecting upon the diagram, the question could also arise as to what happens once a company is optimising everything that NPS is indicating, and then reaches a platform with the NPS number and receives no further comments. Does this mean it is the

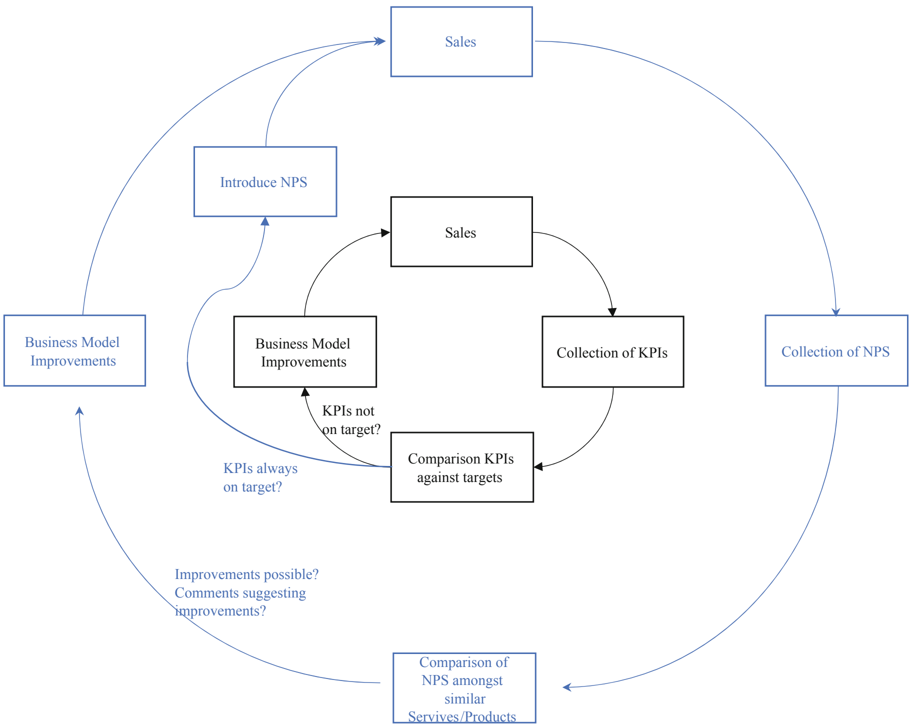


Fig. 5. Generalised model using NPS in the presented case-study.

perfect business? Or do we need the next indicator that is enhancing NPS as NPS is enhancing current KPIs? Researchers could create case studies or quantitative research that investigates what companies are doing in the above situation.

In summary, the readers need to take away that we answered research question one by presenting a case study that suggests that usage of NPS as outlined by Reichheld [7] makes sense as long as an additional question is asked, as already indicated by other research [1, 2, 13, 14]. Furthermore, we build a new business model as shown in Fig. 5 that suggests a model for others to use.

6 Conclusion and Outlook

This current research shows that NPS can be helpful to drive change in any service environment. It could even be seen as one part of the innovation strategy of a company, as suggested by Pisano [21] as it gives new client-driven insights leading to innovation. Currently, the current corona pandemic and the resulting trend toward increased online learning could be an additional proof for NPS as indicator for innovation. It could be interesting to see if research can validate that NPS is helping to drive new innovation in the current situation.

Looking into the future there are interesting challenges arising from this research. One question is to investigate which kind of innovation is triggered by NPS. Is it just incremental innovation or is there potential for, for example, disruptive innovation? Is there a categorisation possible? And, more strategically, we need to address the question what happens once NPS has reached a high and constant level in a company, meaning clients are satisfied with the performance of the company and/or the existing system has adapted to satisfy this KPI. Then, NPS will then no longer bring new ideas and will no longer act as initiator of innovation. This leads to a new fundamental research question: What is the next trigger for innovation following Reichheld's NPS [7], the NPS 4.0?

References

1. East, R., Romaniuk, J., Lomax, W.: The NPS and the ACSI: a critique and an alternative metric. Int. J. Mark. Res. **53**, 327–346 (2011)
2. Florea, N., Tănăsescu, D., Duică, A.: Enabling customer-centricity and relationship management using net promoter score. Valahian J. Econom. Stud. **9**, 115–126 (2018)
3. Sharp, B.: Net Promoter Score Fails the Test. Marketing Research: a quarterly business management publication of the American Marketing Association 20, 28 (2008)
4. Decker, C., Hartman, K.: Brand personality, congruency, and net promotor score: a university case study. J. Mark. Perspect. **1**, 23–37 (2016)
5. Krol, M.W., Boer, D., Delnoij, D.M., Rademakers, J.J.D.J.M.: The net promoter score - an asset to patient experience surveys? Report 18, 3099 (2015)
6. Shevlin, R.: It's Time to Retire the Net Promoter Score (And Here's What To Replace It With) (2019). https://www.forbes.com/sites/ronshevlin/2019/05/21/its-time-to-retire-the-net-promoter-score/. Accessed 15 May 2020
7. Reichheld, F.F.: The one number you need to grow. Harvard Bus. Rev. **81**, 46 (2003)
8. Wilberforce, M., Poll, S., Langham, H., Worden, A., Challis, D.: Measuring the patient experience in community mental health services for older people: a study of the net promoter score using the friends and family test in England. Int. J. Geriatr. Psychiatry **34**, 31–37 (2019)
9. Ghosh, M.: Case study: text-mining customers view point and perceived value about brand. Int. J. Bus. Anal. Intell. **4**, 1–4 (2016)
10. Mandal, P.C.: Net promoter score: a conceptual analysis. Int. J. Manag. Concepts Philos. **8**, 209 (2014)
11. Keiningham, T., Cooil, B., Aksoy, L.: A longitudinal examination of net promoter and firm revenue growth. J. Market. **71**, 39 (2007)
12. Schulman, K., Sargeant, A.: Measuring donor loyalty: key reasons why Net Promoter Score (NPS) is not the way. Int. J. Nonprofit Voluntary Sect. Market. **18**, 1–6 (2013)
13. Korneta, P.: Net promoter score, growth, and profitability of transportation companies. Int. J. Manag. Econom. **54**, 136–148 (2018)
14. Burnham, T.A., Wong, J.A.: Factors influencing successful net promoter score adoption by a nonprofit organization: a case study of the Boy Scouts of America. Int. Rev. Public Nonprofit Market. **15**(4), 475–495 (2018). https://doi.org/10.1007/s12208-018-0210-x
15. Eisenhardt, K.M.: Building theories from case study research. (Special forum on theory building). Acad. Manage. Rev. **14**, 532 (1989)
16. Yin, R.K.: Case Study Research: Design and Methods. SAGE, Thousand Oaks (2003)

17. Swanborn, P.G.: Case Study Research: What, Why and How?. SAGE, Los Angeles (2010)
18. Yin, R.K.: Case Study Research: Design And Methods. SAGE, Los Angeles (2009)
19. Ziegler, A.: Case study database: The value of a Net Promoter Score in driving a company's bottom line: a single-case study from IBM Training Services (2020). https://www.researchgate.net/publication/340512333_Case_study_database_The_value_of_a_Net_Promoter_Score_in_driving_a_company's_bottom_line_a_single-case_study_from_IBM_Training_Services. Accessed 15 May 2020
20. Johansen, J., Jan-Pries-Heje, J.: Software process improvement (SPI) manifesto. Softw. Qual. Prof. **12**, 12–19 (2010)
21. Pisano, G.P.: You Need an Innovation Strategy (2015). https://hbr.org/2015/06/you-need-an-innovation-strategy. Accessed 15 May 2020

Recognising the Types of Software Assets and Its Impact on Asset Reuse

Kathleen Wolfram[1(✉)], Michael Martine[2], and Paul Sudnik[3]

[1] Heriot-Watt University, Edinburgh, Currie EH14 4AS, UK
kw74@hw.ac.uk

[2] Dept/Bldg/Office, 3039 Cornwallis Rd., Research Triangle Park, P.O. Box 12195, Durham, NC 27709-2195, USA

[3] Munich University of Applied Sciences, Am Stadtpark 20, 81243 Munich, Germany
paul.sudnik@hm.edu

Abstract. Early software reuse success triggered immense euphoria culminating in many publications on the topic. However, documented successes could not keep pace. Some commentators even called for the end of the reuse.

This paper examines what can be reused and how. Interviews indicated the need for a more precise asset definition and an enhanced asset description. It is suggested that identified types of assets should be treated individually, thus increasing their chance of reuse. This will assist in overcoming the biggest hurdle of reuse – the asset discovery. Here the so far manual asset identification/selection will be replaced by a triggered, systematic and cognitive asset suggestion via "Reuse Suggestion Engine" based on (a) sufficiently structured asset records and (b) continuously maintained asset entries.

This interim research report provides business management guidelines of how both the assets as well as the organisational set-up should be shaped to be prepared for the era of automation which will propel us into a future, more abstract chapter of continuous software reuse resulting from an automated solutioning. This update will have a positive impact on the increasing number of traditional hardware products whose heart and soul is now software.

Keywords: Software asset reuse · Asset suggestion engine · Automated solutioning · Systematic reuse · Continuous reuse

1 Introduction

The ever-increasing amount of software in high-end hardware, e.g. Tesla cars are "computers on wheels" (Loh 2019), determines the quality, the security, and the life span of these products. Software reuse could be of help (McIlroy et al. 1968; Prieto-Díaz 1993) but has not delivered on its promise, yet. Some even calling for the ending of efforts in the field of software reuse (Prieto-Diaz 1994). Is there still hope?

Ideally, if there are people who create assets and put them into asset repositories, this should result in high-quality reuse at a low cost. In reality, (1) the "producer activities" (Frakes and Terry 1996), the creation and the sharing of assets by the *asset owners* work surprisingly well. (2) As a result of this, there is plenty of *assets* in

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 162–174, 2020.
https://doi.org/10.1007/978-3-030-56441-4_12

(3) crowded, always outdated Asset Repositories undergoing endless centralisations and updates. (4) The shortfall in systematic *reuse* "consumer activities" (Frakes and Terry 1996) is little due to mal-intent perception of re-users (Ryan et al. 2019) or missing reuse target setting but a lack of operationalisation and documented reuse (Prieto-Diaz 1994). Despite these challenges, there is excellent ad-hoc reuse. "The problem we face in software engineering is not a lack of reuse, but a lack of widespread, systematic reuse" (Prieto-Díaz 1993) is still valid.

Systematic reuse (Fig. 1) starts during pre-sales and stops when the contract gets fully terminated. It requires three fundamental approaches: (a) Each existing asset should be considered for reuse in each newly emerging opportunity (maximise "Opportunity Assetisation"). (b) All existing contracts should get continuously assetised (maximise "Contract Assetisation"). The applicability of assets in contracts can change as new assets are continuously added to the asset repository. (c) At the same time, the reusability of each asset should be continuously expanded (maximise "Asset Scope").

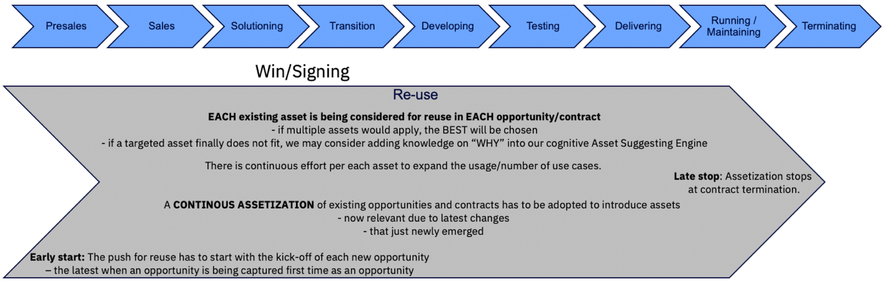


Fig. 1. Systematic reuse.

While business benefit is measured short-term, *reuse benefit is realised mid- to long-term*. Creating an asset is, like planting a tree, an investment into the future. Software metrics are vital for software management (Ram and Devi 2019). However, it is *challenging to define metrics* measuring the reuse success along the life cycle of the contract(s) reusing the asset. We can conclude from this that we have not at all reached what (Prieto-Diaz 1994) called the "end of the reuse". Instead, we got to a sound level of ad hoc reuse. Traditional approaches will not bring significant improvement. Therefore, we should now look into alternative ways.

This paper will explain in Sect. 2 what a software asset is before Sect. 3 reveals key interview findings. Section 4 proposes derived key concepts towards systematic reuse which will be summarised in Sect. 5.

2 The Asset

2.1 What Is an Asset?

Traditionally, IT Companies provide tailored IT solutions to their clients (top of Fig. 2): One client gets one feature he wants, another client gets another. There is no link between the two projects. Each client receives, often seller-driven, a unique software solution. Enhancements are possible but on an individual basis. There is a 1:1 business relationship between the IT provider and his client.

IT providers may shape an IT product that they then sell via license to millions of clients (see bottom of Fig. 2): Every client receives the same set of software. New features will be rolled out via a new release to all clients. There is a 1: many business relationship between the IT provider and his vast set of clients. Marketing makes clients aware of the IT provider's software products. The business relationship is typically initiated by the client who turns to the IT provider asking for a license (client-driven).

Assets, however, (see middle-layer of Fig. 2) are *a hybrid construct* in a way that none of the clients licences the same set of software. Crucial parts of the software are licensed to ca. a dozen of clients. Open Source is similar to Asset-based IT-Service (but for free).

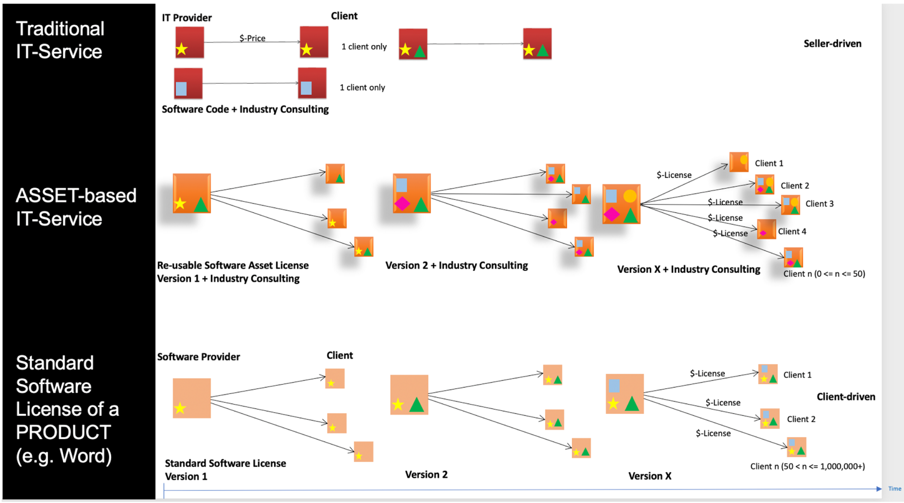


Fig. 2. Difference between a product, custom IT service and asset-based IT service.

2.2 What Can Be Reused?

There is a wide range of digital assets (or artefacts) that can be reused in business (Herrin et al. 2020), e.g. offerings, sales material, software, requirements, documentation, architectures, estimates, test cases, test data, business process knowledge, services.

Reusable software (Fig. 2) can be either a product or an asset: Software products can be internally owned products or third-party products. Software assets can be internally owned or open-source code[1].

This paper is focusing on the reuse of software assets in combination with related artefacts, i.e. the software system (Imoize et al. 2019) such as code in a bundle with sales material, documentation, architectural charts, implementation plans, test plans, test cases, or test data. These *Software Services Assets* are typically more prominent in size justifying the effort required to spot them, commercialise them, upload them into an Asset Repository, promote them, and retrieve them for reuse.

2.3 Benefits of Reuse

Reuse offers tremendous benefits (Table 1). Reused software is "hardened" and therefore of significantly higher quality. This quality improvement is essential whenever software is used in conjunction with expensive hardware made for a more extended use period.

Table 1. Benefits of reuse.

Benefits of Reuse for the …	Client	IT company
Speed of go-to-market	x	x
High quality	x	x
Low price	x	x
Long-term software maintenance (stable team over a long time)	x	x
Experts/Centre of competence	x	x
Key differentiator in the market		x
Reduced testing		x
Productivity increase		x
Triggers the closure of deals much bigger than the size of the asset		x

Reusability is one of the eleven quality attributes defined by (McCall et al. 1977). The reuse of mature Services Assets maximises its impact (1) in situations that require specialist work and (2) for complex solutions whose development would be resource-intensive and require massive funding.

3 Findings

Nine Asset Owners participated in a semi-structured interview. Based on the pinpointed topics, the following themes got identified, brought into a logical order, and analysed: (1) Asset Management, (2) Asset Marketing, (3) Asset Sale, (4) Solutioning, (5) Asset

[1] Neglecting the option that an IT provider could reuse the asset of a competitor.

Development, (6) Asset Delivery, and (7) Asset Maintenance. The first three topics were (in the order of appearance) of particular importance to the Asset Owners. They will now get discussed in more detail.

3.1 Asset Management

Several Asset Owners rated their asset as a "very old and *mature* asset". One notably successful asset is older than 20 years. It started in Media & Entertainment and is now being reused in various geographies and industries. Follow up revealed: The asset itself, the team's strong expertise and dedication ensure continuous market success. Every time this asset team got engaged, they did not start from scratch but transformed and expanded their asset, making it reliable and robust.

As far as the *alignment of assets against offerings* is concerned, most Asset Owners agree this helps assets "getting considered and may be included" into the deal. Nevertheless, not all assets could benefit from mapping against suitable offerings ("The asset … doesn't have a direct … offering that it always goes with…"). Some asset owners favour cross-selling as a way to promote the reuse of their assets. It is, in fact, a surprise to see established assets not always benefiting from their mapping. This is an area worth exploring.

The third major topic asset owners talked about was *Asset Attributes*. Whether it is cloud or on-prem, provides microservices or not. Whether it is an asset relevant purely for a particular industry or a specific service line. One thing that gets obvious from the answers is that assets cannot be labelled and put in a box: Most of them started as on-prem assets and are now available on the cloud. Some of them started as industry assets and expanded their reach via cross-selling. Unlike products, assets are much more agile, changing over time in their technical set-up as well as their target clients. This has an impact on their documentation. It needs constant updates to reflect these changes.

3.2 Asset Marketing

All nine interviewed Asset Owners talked in detail about "Asset Marketing", especially about their *company internal asset marketing* to enable sellers of other account teams to offer their asset for reuse. Statements like "it's not an easy path to make other(s) … knowledgeable on how to apply …" the asset, and "The assets are amazing but not being utilised." describe best what would happen if the Asset Owner would take no action. "I have done whatever I could do, but that is definitely not enough." is one assessment. Some call for a "visionary salesperson … making things happen".

"We need both internal-facing and external-facing communication around assets …" or "I'm thinking about how to promote more, but I need more resources and lacking funding on that." or "No one taking control from an offering perspective …" indicates that *Asset Owners are implicitly expected to cover multiple roles* – from pre-sales/marketing over sales, solutioning, transformation, to delivery.

The second-largest focus area in Asset Marketing can be tagged with "Client Interaction". Here the asset owners state "We don't have a dedicated … asset sales team." It is being compensated by involving the "*Personal Network*" to cover the pre-sales activities.

One Asset Owner thinks ahead coming up with the idea of an "intelligent analysis of the content … that is something that can be done using AI and using our pipeline and *recommending assets based on the content* of sales tools" referring to some kind of an Asset Suggestion Engine which will be picked up and further discussed in Sect. 4.4.

Most others list in detail typical marketing activities such as an Asset Repository, Client interaction, collaboration with Sectors and Service Lines, Design Thinking, external Trade Shows, the use of an external website, internal documentation, learning about competitors in the market, asset mock-ups, tracking of opportunities or others. There is a substantial ask for marketing assistance which maybe could be addressed by merely sharing an Asset *Marketing Guidance* inspiring the Asset Owners and leveraging marketing approaches.

While the typical reuse promotion aims at this particular branch of the organisation (Global Business Services), there is at least one Asset Owner bringing up the noteworthy point of *cross-branch marketing* of the assets. His asset has not only been reused in all sectors and all geographies but also in 3 distinct branches (Global Business Services, Software Group, and Global Technology Services).

3.3 Asset Sale

Asset Sale is the third most commented theme. The topics broadly comprise three areas (1) Go-to-Market, (2) Contracting, and (3) Opportunities.

As part of the Go-to-Market findings, one Asset Owner highlights that he uses his asset to expand the business rather than expanding the profit. Further, since an asset is not a standardised product but a continually developing, highly customisable piece of software, *sellers have to involve the asset owner during sales* ("Before a seller sells the deal, they first call us"). Appointing asset dedicated sellers would be hard to justify - no seller has time to endlessly catch up with the permanent changes of a "niche product" called asset ("We don't have … sellers out there who can discuss the context and content of (asset)."). Instead, the individual asset owning team supports the sellers every time they are called into a deal.

As far as contracting, the most commented Asset Sale topic, is concerned, the asset owners strongly commented on pricing. It became apparent that by reusing assets, *teams can price much faster and better* because (a) they received upfront pricing support and advice, (b) their previous experience with that particular asset, and (c) the asset team is very much impacted by the price out of which both the (typically very stable) asset team, as well as asset investment, will be funded. However, teams were asking for Software as a Service (SaaS) pricing guidance ("there is not so much maturity when it comes to selling SaaS. And this is more about the pricing and legal constructs around it … I find some sort of assistance on that part can also be helpful to the teams …").

Finally, the asset owners talked about their opportunities. It became clear that the asset team is typically technically profoundly involved in every reuse deal. While there is a keen interest to track the reuse in the system, there is *no agreed tracking process* ("We would like to have the ability to track where the asset is being included in deals"). As a consequence, the actual reuse is not or not fully reflected in the management system leading executives to assume that reuse is not taking place. All Asset Owners mentioned that they bring their asset into global (not just local) opportunities and are involved in related client discussions right from the beginning - due to the aforementioned Asset Marketing. Asset and asset owner must be known to the sellers.

3.4 Solutioning and Asset Development

Solutioning is an area, the Asset Owners consider business as usual. Worth mentioning is one Asset Owner untapped new reuse potential – suggesting expanding reuse by *partnering with related assets* ("there is a solution here you should partner with. There is an opportunity here, you could position your solution … just point us to similar assets. That would be very, very useful.") to establish an ecosystem.

Regarding Asset Development, Asset Owners commented that Asset Investment into new development is an issue if neither central funding is provided nor enough funding is being generated out of the asset sale. Here at least one team spotted an exciting option: develop the asset by *reusing other asset*s ("giving us assets that originated somewhere else … and bring them to us.").

3.5 Asset Delivery and Maintenance

Some Asset Owners mentioned the support the Asset Teams are providing varies - there is *no standard, no common support agreement*: "I think if they had more confidence around the process of using … assets and our product versions of our software it would be easier to deploy those technologies across the companies."; "From a support perspective, my time is limited".

Only a few asset owners presented their assets as *platform assets* (although they are). There is no agreed definition of a platform asset. This is an area that certainly demands more management attention and guidance.

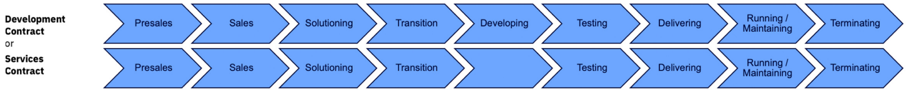
Concerning Asset *Maintenance*, an asset needs both maintenance strategy and maintenance execution ("An asset is not something that happens in a moment in time. There has to be a maintenance view of it.").

4 Discussion

4.1 Types of Software Assets

Asset reuse can be recommended to industrial clients. It provides stability and continuity which are very much needed since the design of, e.g. a car takes ca. 7 years until it reaches the market where it will be in use for ca. 15 years. So, we need software that will be around for 20+ years in our cars – in the absence of software products software assets are here well suited.

Software Assets can be applied throughout all phases of the software life cycle, not just during sales. Figure 3 maps Software Services Assets against the project phases (Sales, Development, Transition, Test, Delivery, etc.) during which they can get applied.



Reuser of Software Services Asset	Example	Sellers	Solutioners	Transition team	Development team	Test team	Delivery team	Maintenance Team	Account Team
Offerings incl. assets		x	x						
Sales Assets	AREMA	x	x	x	x	x	x	x	
Development asset/ Accelerator	SAP …		x		x				
Transition/acquisition assets	Transition assets			x					
Test assets	IGNITE Testing					x			
Productivity/ Automation assets	Digital Worker			x				x	
Delivery assets	Ticket Analytics						x		
Terminating assets	Merger & Acquisition Accelerator								x

Fig. 3. Types of assets – and when they can be applied during the software life cycle.

Some assets will not generate revenue but savings. All of them help teams to perform their work better, innovating from a higher level. "Technology reuse…, is the easiest to achieve and has the least impact. Reusing information and business processes are more challenging but deliver the greatest bottom-line impact." (Woerner et al. 2013).

Since Open Source Assets grow in importance, reuse should not be limited purely to internally owned assets – but include Open Source assets, too.

IT providers could look into options of making their smaller assets (e.g. pure code assets) open source. Especially if assets are less key to competition, making them open-source could increase their maturity and robustness while hosting costs stay minimal.

4.2 Asset Attributes

Up till now, all assets are described with the same narrow set of attributes in the central asset repository. However, as per each Software Services Asset Type, different attributes are relevant (Table 2). For example, the "Average Revenue generated per Reuse" may be a great attribute of a Sales Asset, but not of a Transition Asset which typically triggers savings. Important conclusions are: (1) Depending on the Asset Type, *different* Asset Attributes are relevant. (2) This dedicated set of asset attributes improves picking of the most suitable asset (Kalia and Sood 2014). (3) Each of these *Asset Attributes reflects past reuse* experience (future reuse experience may be different). (4) *Metrics depend on the type of asset*. The success of an asset may be misinterpreted if the type of asset is not adequately considered.

Table 2. A selection of Asset Attributes that can be relevant for specific Asset Types.

1-pager;	Automation tower	Company internal deck	Company benefit/RoI
Architecture	Which cloud vs. on-prem	Internal site, wiki, slack	External website
Code	Project lifecycle phase	Competitive position	Implementation time
Owner	Client mock-up	Cost/Price/License	Platform asset
Support	Commercialisation	Client benefit	Describing keywords
Related offering	Requirements	Sector, Industry	Technology
Maturity	Sales marketing video	Service line/Practice	Use case, Purpose
Microservices	Sales presentation deck	Technical video	

4.3 Reuse Process

"Reuse just doesn't happen naturally." (Woerner et al. 2013) Reuse is the conscious process of replacing ahead of time (Fig. 4) upcoming work with work done previously. It requires (1) a *Trigger* (Sect. 4.6) to get the reuse started earliest at the *Entry Point of Reuse* and latest when the *Window of Reuse* closes, (2) an effort to understand the piece of work considered for reuse, (3) an effort to locate a suitable asset ideally via a Reuse Suggestion Engine (Sect. 4.4), (4) an effort to prioritise one asset out of a selection of many, (5) a final decision for reuse or against it, (6) administrative work to initiate the reuse, (7) technical interaction with the asset owner, (8) harvesting of the new feature to incorporate the outcome of the current project back into the main asset (new features could be offered to existing customers).

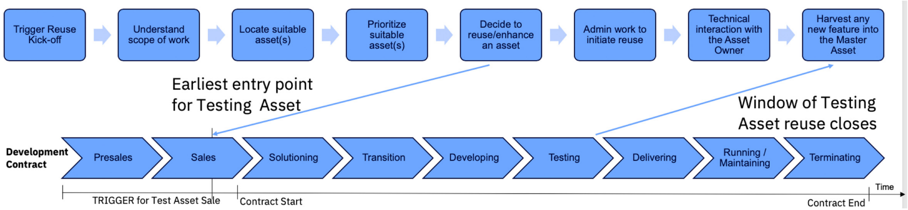


Fig. 4. Example of a "Testing Asset"-reuse process along the software life cycle.

4.4 Reuse Suggestion Engine

Triggers reminding teams to reuse assets at certain project stages are not enough. There are simply too many assets for a human to keep up with. In order to move from an ad hoc reuse towards systematic reuse, a cognitive "reuse suggestion engine" is needed. It applies Artificial Intelligence and deep learning to the business operation

(Tarafdar et al. 2019) of precise asset fetching (Ali et al. 2020) thus expanding the concept of Eldgeway's asset "discovery engine" (Elgedawy 2015) and Hummel's "hybrid semantics-driven retrieval engine" (Hummel and Atkinson 2006). It acts like a Cognitive Helpdesk: Input is the user request (via Word or PDF document), and the output is an elaborated solution suggestion.

The Reuse Suggestion Engine automates solutioning, a job so far done without any cognitive support. It allows solutions to be fully sketched out and priced. Similarly, based on the newly introduced software services asset types and their just explain type-specific sets of attributes, it will now be possible to instantly come up with the right assets for project phases like transition, development, testing, delivery, maintenance, and termination – areas so far less in focus for asset reuse. The introduction of a Reuse Suggestion Engie has multiple benefits: (1) It avoids the difficulty of manually locating suitable assets. (2) It automatically presents located (internal and open-source) assets in a ranked order (spotting overlapping assets that burn resources). (3) It cuts time for solutioning from weeks to hours. (4) It allows to cross-company optimise significant proposals that previously would have been worked on by separate teams. (5) It is available 24/7. (6) It works independently of the availability of specific experts since it incorporates and builds up solutioning experience per each project stage. Consequently, the systematic use of a cognitive Reuse Suggestion Engine allows providing clients speedy with high-quality input. As demonstrated years ago in Jeopardy!, cognitive decision taking can beat human capabilities (Tesauro et al. 2013).

4.5 How to Keep the Asset Records in the Asset Repository up to Date

Assets are more agile than products. It makes sense to "tell" some kind of *Cognitive Asset Repository* were documents about the asset, its usage, or its pricing are to be found. This way, a Cognitive Asset Repository can continuously conclude on the values that populate the vast and expanded set of Asset Attributes explained in Sect. 4.2. While this approach requires perfect discipline in the way the input documents (e.g., presentation charts, opportunity records, contracts, solution design documents) are being stored, there is no need to store all of these documents duplicate and in precisely the same place. The Asset Repository will get virtually updated if the distributed sources of input are made known to the cognitive engine. Besides, it will stipulate inspirational reuse.

4.6 Effect on the Business

There has been much discussion about measuring the success of reuse (Imoize et al. 2019). If the business would be seen as a helpdesk, experts could focus on the Level 3 topics (new) while all L1 (simple) and L2 (challenging) tickets can be resolved by reusing past approaches. Reuse unfolds its most significant effect if it spans the entire business process, not just IT. Digital reuse includes the reuse of business process knowledge, IT, and data (Woerner et al. 2013). There are options for auto-harvesting.

"…, reuse remains an underutilised strategy in most companies because it requires a high level of organisation, discipline and enterprise-wide exchange." (Woerner et al. 2013). As a consequence, reuse should be made part of the daily business (Fig. 5).

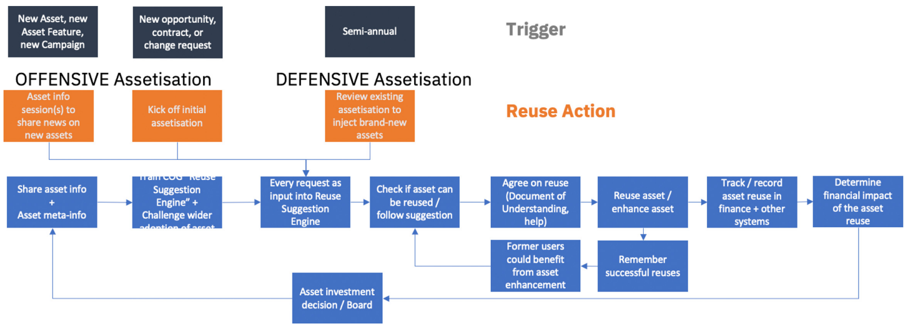


Fig. 5. Systematic asset reuse process.

This will help the teams, as they complained: "Have been requesting for a more structured, pipeline-driven approach." Figure 5 explains how this process looks like.

5 Conclusions

5.1 Contribution to Knowledge

Our grandfathers have been right suggesting asset reuse. However, "to achieve significant payoffs a reuse program must be systematic" (Frakes and Terry 1996). Working harder does not help – we need to work *differently*. Technological investment in **Cognitive Automation** is needed to (i) ease the asset onboarding, management, continuous asset attribute updating and the reuse suggestion while acknowledging **the different kinds of assets** and reuse and (ii) make reuse a part of the **organisation's business processes**.

Explicit new findings are: (1) Assets are a *hybrid construct* located between fully tailored services and products. (2) Assets can become *pretty old* while providing stability and continuity, having the asset owning team act as a competence centre. (3) The *mapping of assets against offering* seems to be less relevant than previously thought and, as the Reuse Suggestion Engine is emerging, may lose even further in importance. (4) Unlike products, assets are *much more agile* in their Asset Attributes. This has effects on (a) the Asset Management/Marketing and (b) the update of Asset Attributes as they notoriously reflect only the past. (5) *Asset Attribute categories* depend on Asset Types. This impacts asset reuse metrics. (6) A more detailed description of the *Asset Reuse Process* has been shaped – responding to why teams struggle with reuse unless *Reuse Triggers* are systematically put in place. (7) Suggesting the introduction of an *Asset Suggestion Engine* that goes beyond what" (Elgedawy 2015) described.

This interim research focusses on statements of experienced Asset Owners owning mature Sales Assets. Future research should consider younger Asset Owners, younger assets, and saving-generating Assets (e.g. Testing Assets).

5.2 Vision

Cognitive knowledge support will boost the positive effect of reuse on the quality of services and products. As “new problems are likely to have been solved in other areas of the business, at least in part”, reuse allows to “deliver more capability in less time” (Woerner et al. 2013). There is almost no limit: Reuse what can be reused immediately, systematically, and globally going beyond software to include “business processes, services, products, and data” (Woerner et al. 2013).

Reuse Automation will lead to a higher, more abstract reuse maturity level. We will then spend more time dealing with the different kinds of assets, metrics and so on focusing on controlling the process much more actively once we have it systematically implemented. Reuse is more than a mindset. It becomes THE “source of competitive advantage” (Woerner et al. 2013).

References

Ali, N., Daneth, H., Hong, J.E.: A hybrid DevOps process supporting software reuse: a pilot project. J. Softw. Evol. Process e2248 (2020)

Elgedawy, I.: USTA: an aspect-oriented knowledge management framework for reusable assets discovery. Arab. J. Sci. Eng. **40**, 451–474 (2014). https://doi.org/10.1007/s13369-014-1428-5

Frakes, W., Terry, C.: Software reuse: metrics and models. ACM Comput. Surv. (CSUR) **28**(2), 415–435 (1996)

Herrin, B.C., Johnson, M.S., McGuigan, M.T., Stein, J.: System and method to identify, gather, and detect reusable digital assets. In: Google Patents (2020)

Hummel, O., Atkinson, C.: Using the web as a reuse repository. In: Morisio, M. (ed.) ICSR 2006. LNCS, vol. 4039, pp. 298–311. Springer, Heidelberg (2006). https://doi.org/10.1007/11763864_22

Imoize, A.L., Idowu, D., Bolaji, T.: A brief overview of software reuse and metrics in software engineering. World Sci. News **122**, 56–70 (2019)

Kalia, A., Sood, S.: Characterisation of reusable software components for better reuse. Int. J. Res. Eng. Technol. **3**(05) (2014)

Loh, E.: Interview: Elon Musk Reflects on Significance of Tesla Model S (2019). https://www.motortrend.com/news/tesla-elon-musk-interview/

McCall, J.A., Richards, P.K., Walters, G.F.: Factors in software quality. volume i. concepts and definitions of software quality (1977)

McIlroy, M.D., Buxton, J., Naur, P., Randell, B.: Mass-produced software components. In: Paper presented at the Proceedings of the 1st International Conference on Software Engineering, Garmisch Pattenkirchen, Germany (1968)

Prieto-Diaz, R.: The disappearance of software reuse. In: Paper presented at the Proceedings of 1994 3rd International Conference on Software Reuse (1994)

Prieto-Díaz, R.: Status report: software reusability. IEEE Softw. **10**(3), 61–66 (1993)

Ram, B., Devi, A.: Code reuse & reusability of the software (2019)

Ryan, T.J., Walter, C., Alarcon, G., Gamble, R., Jessup, S.A., Capiola, A.: The influence of personality on code reuse. In: Paper Presented at the Proceedings of the 52nd Hawaii International Conference on System Sciences (2019)

Tarafdar, M., Beath, C.M., Ross, J.W.: Using AI to enhance business operations. MIT Sloan Manage. Rev. **60**(4), 37–44 (2019)

Tesauro, G., Gondek, D.C., Lenchner, J., Fan, J., Prager, J.M.: Analysis of watson's strategies for playing Jeopardy! J. Artif. Intell. Res. **47**, 205–251 (2013)

Woerner, S., Weill, P., McDonald, M.: Turn time into money: faster growth through digital reuse. Eur. Bus. Rev. **25**(3), 38–42 (2013)

SPI and Agile

Exploring Technology Acceptance and Planned Behaviour by the Adoption of Predictive HR Analytics During Recruitment

Thomas Peisl[1] and Raphael Edlmann[2(✉)]

[1] Department of Business Administration, University of Applied Science, Am Stadtpark 20, 81243 Munich, Germany
tpeisl@hm.edu

[2] Edinburgh Business School, Heriot-Watt University, Edinburgh, Scotland, UK
raphael-edlmann@web.de

Abstract. This research aims to investigate the technology acceptance and use behaviour of hiring mangers when it comes to the adoption of predictive human resources analytics during recruitment. Additionally, this paper discusses the identification of dishonest behaviour to increase the job offer success during algorithm-based data screening. In the age of digital transformation, researchers and practitioners explore the possibilities of predictive analytics in human resource recruitment. Predictive data modelling enables hiring managers to discover attrition, reduce cognitive bias, and identify the compatibility between job candidates and organizational environments. The unified theory of technology acceptance and usage (UTAUT) will be used to identify the intention and use behaviour of hiring managers when it comes to the application of predictive HR analytics. It will also be explored how the actual system use impacts key recruitment performance indicators. The structural relationships of the UTAUT model will be examined by an empirical questionnaire and a partial least square structural equation model (PLS-SEM). To predict the misrepresentation and dishonesty practised by job candidates during algorithm-based data screening, the theory of planned behaviour is applied in conjunction with semi-structured interviews. This research uncovers to what degree human resource managers trust, accept, and integrate predictive HR analytics in daily routine. Further, data modellers and researchers should be able to test, improve, and optimize future machine-learning algorithms based on the dishonest behavioural themes identified in this research study. Finally, this research will show how software process improvement (SPI) initiatives can be constantly improved by machine learning algorithms and user group requirements.

Keywords: HR analytics · Machine learning · Big data

1 Introduction

Digitization and the emergence of new technologies in the past fundamentally reshaped operational and decision-making processes for organizations. During the digital age boom, one of the major technological trends in human resource (HR) management is

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 177–190, 2020.
https://doi.org/10.1007/978-3-030-56441-4_13

the analysis of large datasets based on a computational foundation [1, 2]. In a highly volatile environment, where markets continuously demand different work ethics and skills, HR managers are encouraged to optimize their hiring decisions during the recruitment process [3].

Previous research emphasized the negative effects for organizations whenever HR managers hired wrong candidates. Poor recruitment decisions lead to a financial loss (e.g. recruitment initiatives, salary, training, etc.) and at the same time high staff turnover rates have a negative effect on the overall productivity and performance of the organization [4]. To improve the decision-making process during recruitment, technology vendors and researchers explore the opportunities of big data analytics.

In terms of HR management, big data analytics try to identify patterns and correlations within large datasets to provide a better basis for decision-making processes [5–7]. Moreover, big data analytics aim to reduce efforts such as candidate screening, selection, and retention. While the screening of candidates during the applicant-selection process for each vacancy is very costly and time-consuming, HR professionals can enhance their productivity by using predictive HR analytics [8].

By using statistical, data mining, machine learning, and AI techniques, predictive analytics use current and historical data to predict future events. Organizations can use these insights to remain proactive through all stages of the employee lifecycle [9]. If, for example, hiring managers try to occupy open positions, data mining and AI algorithms can be used to autonomously find a match between the resumes and job requirements of applicants [10]. Moreover, by the analytical evaluation of current and historical data, such as experience, role maturity, manager ratings, and promotion status, organizations can calculate the probable risk of attrition. In this regard, the HR department can decide to offer appealing job positions to retain employees who are willing to shift [11].

While the opportunities relating to predictive analytics during recruitment has been explored, technology acceptance, usage of user groups, and the actual impact on business performance are yet to be examined. This gap leads to the following research questions: RQ1: What drives the intention of HR managers to use predictive HR analytics during recruitment? RQ2: What is the impact of predictive HR analytics on key HR performance indicators?

Despite the opportunities that come along with predictive HR analytics during the hiring process, potential barriers should be considered. At a low job offer success rate, the candidates can intentionally corrupt their data to increase the probability of getting the job. This leads to unreliable results, which might be considered during the hiring process. Machine learning, which aims to optimize the process performance based on past information from datasets, must continuously be improved to deliver reliable results [8]. While different application scenarios and potentials of predictive HR analytics have been explored, dishonest actions by job candidates to overcome algorithms during recruitment are yet to be investigated further. This gap, in turn, leads to the following research questions—RQ3: What drives the intention of job candidates to perform a behaviour when it comes to profiling algorithms? RQ4: How can machine-learning processes be optimized when it comes to misrepresentations of job candidates on resumes and career-oriented social media profiles? While digital transformation brings a wind of change inside the organization, it also requires commitment and engagement of the people involved in the process [12]. To improve software design methods, researchers and practitioners show a great interest in how people respond to new technology [13]. While algorithm-based HR decision-making might reduce the potential biases of HR managers during recruitment, challenges arise in terms of ethical awareness, personal integrity, and compliance [14]. Therefore, the level of technology acceptance of HR managers plays an important role when it comes to hiring decisions based on computational algorithms.

2 Data-Driven Human Resource Recruitment

HR recruitment is the process of attracting and hiring the best possible job candidates for a job position in a cost-effective and timely manner. The process consists of attracting, screening, selecting, hiring, and integrating applicants into the environment of an organization [15]. While external recruiting methods focus on integrating new job candidates from the labour market, internal recruiting tries to fill open vacancies by people who already work for the organization [16].

Predictive HR analytics is a systematic predictive modelling approach that applies sophisticated statistics and quantitative analyses techniques on existing data relating to HR people. This predictive modelling approach is sometimes linked with computational data science techniques such as data mining, AI, and machine learning [17]. To support evidence-based decision-making during recruitment, these data-driven techniques also apply social science theories to reduce the biases and wrong perceptions of hiring managers [18–20]. Finally, the superior goal of predictive HR analytics in the context of recruitment is to improve decision-making, increase efficiency, and create a positive impact on key performance indicators.

2.1 Predictive Algorithms to Identify Personality Traits from Online Social Networks

Information system researchers have developed and evaluated predictive algorithms that try to identify personality traits from career-oriented social media profile data. Through empirical research it is investigated if there exist linkages between personality traits from the five-factor model and social media attributes. Afterwards, the empirical results are used to evaluate which algorithmic method (e.g. decision-tree process, artificial neural networks etc.) is the best to identify and predict personality traits from social media profile data. The most suitable algorithm has an accuracy between 31.4 and 46.2 [21].

Gou, Zhou and Yang tried to automatically derive personality traits from Twitter messages. This experimental research allowed the study participants to log in to a designed system which collected the most recent 200 public tweets. Next, a computation algorithm inferred personality traits from their messages based on a lexicon-based approach that calculated correlations between personality traits and word categories. Finally, the participants were asked how well each derived personality trait matched their own perceptions. The mean values of all ratings for the big five personality were above 3 (“somewhat”) [22]. These results showed that social-media platforms can have valuable data when it comes to the prediction of personality traits. In terms of recruitment, these metrics can be used to identify a match between organizational culture and people's personalities.

A person–organization environment (P-OE) fit is given when people and work environment characteristics are well matched [23]. The P-OE model consists of a person–organizational (P-O) fit, which is the compatibility between people and the organization, a person–group (P-G) fit, which is the compatibility between people and their work groups, and a person–job (P-J) fit, which is the match between people's abilities and job requirements [24, 25]. Based on this, the design of recommender systems for online social network recruiting has become important in recent years [26]. E-Recruitment systems can apply the findings of predictive analytics in terms of personality traits to achieve a higher fit between the organization environment and the job candidate. In 2017, Buettner created an electronic recruitment framework for online social network recommender systems which considers the social context in the form of the personality traits of job candidates (Fig. 1).

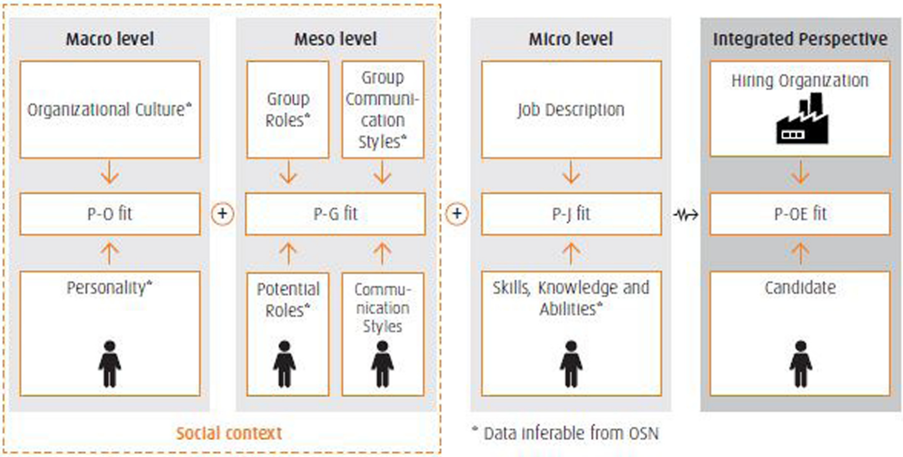


Fig. 1. E-recruitment framework [21]

Based on the P-OE score a candidate ranking can be created. Depending on the requirements of the job position, recruiters can set filter criteria to shortlist the candidates. This shortlist would support recruiters to compare the candidates in a rational and consistent way [27]. Moreover, it would support them to make an unbiased hiring decision.

Following this approach, Buettner emphasizes, other than recruitment, the value-added potential. Therefore, the P-OE score can also be used to identify gaps for employee development and retention. Additionally, organizations can plan initiatives for team and organizational development [28].

2.2 Employee Retention by Predicting Risk of Attrition

Predictive analysis in terms of attrition aims to identify employees who are willing to relocate. Moreover, HR managers can improve their internal recruitment by offering these employees valuable open job positions. In this relationship, HR departments can reduce staff turnover rates by the adoption of predictive HR analytical and machine-learning approaches.

It is important to identify employee attrition in an organizations' predictive data modelling. Since the reasons for which an employee chooses to leave the company can vary between organizations, datasets should be analysed individually. Vasa and Masrani conducted an experimental research study, where a sample dataset of 15,000 employees has been analysed [29]. This dataset contained details of all the employees who stayed as well as left the company. The main attributes examined in this context were the number of working hours, the number of projects worked on, and salaries. In all, 24% of the employees left the company and 76% stayed. Based on this sample data, the researchers applied and compared the accuracy of different machine-learning algorithms (e.g. logistic regression, decision tree etc.) to choose a predictive model for the future [29].

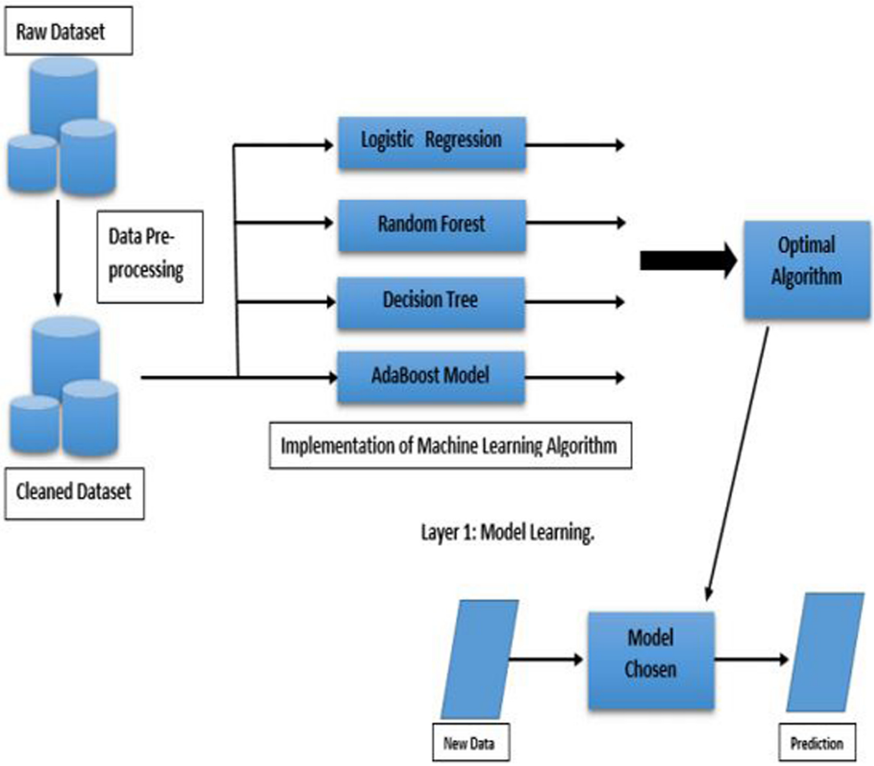


Fig. 2. Predictive modelling approach to foresee the employee turnover [29].

The predictive conclusion of this experimental research was that the turnover rate increased if the employees were overworked, under-performed, or had better job opportunities. Moreover, it was identified that most of the employees with low and medium salaries left the company. Finally, the random forest algorithm has been proved as the most reliable classifier [29].

This predictive modelling example shows that organizations can fill vacant positions by identifying people with the risk of attrition. Moreover, decision-makers can take appropriate measures to confront possible problems that may lead to attrition.

2.3 Technology Acceptance and Use Behaviour of Predictive HR Analytics

Recent research has shown the opportunities of predictive modelling for internal and external recruitment activities. HR predictive analytic tools are designed to optimize performance, thereby generating a higher return on investments based on better decision-making [30]. However, it has not been explored how much HR practitioners trust, accept, use, and integrate predictive HR analytics in their daily routine.

User acceptance is defined as the willingness of a user group to apply information technology for tasks for which it has been designed. Several research observations have shown that there is often a deviation between the actual and planned usage of information technology [31]. To improve software creation and implementation, it is important to note how technology is being used by a user group and to what degree it affects its key performance indicators.

2.4 Predicting Dishonest Behaviour of Job Candidates to Improve Machine- Learning Algorithms

A conducted survey among 3,500 employers showed that 88% of recruiters, hiring managers, and HR staff caught misrepresentations presented through applicant's resume [32]. According to the participants, the most common lies were information regarding skills and abilities, embellished responsibilities, employment dates, job title, academic degree, companies worked for, and awards [33].

Henle, Dineen and Duffy reported that resume fraud negatively impacts job performance [34]. To improve machine-learning algorithms in terms of predictive HR analytics, it is important to know how user groups adjust their behaviour if they know that resume and social media profile screening is done by computational algorithms. The identification of dishonest behavioural patterns of job candidates enables information systems and computer scientists to test the accuracy of different machine-learning algorithms in terms of predictive HR data modelling.

2.5 Contribution to Software Process Improvement (SPI)

One core value of SPI is the active involvement of people. Furthermore, it is recommended to have a right mix of competences on board (e.g. functional experts) to improve parts of the process which affects daily activities of the people involved [12].

Technology acceptance and use behaviour research enables SPI teams to consider user group requirements during the SPI planning phase. Since usability can be considered as an essential factor of quality, return on investments (ROI) of SPI initiatives can be realized earlier as expected.

However, while SPI encourages practitioners to involve people of the organization to increase competitiveness it is also important to consider external stakeholders which might have an indirect impact on the software improvement process [12]. In times of machine learning and AI, software processes can be constantly improved by detecting meaningful data patterns of external stakeholders. In this context, predictive HR data modelling can be further developed by considering behavioural patterns of potential job candidates.

3 Theoretical Framework

The unified theory of technology acceptance and usage (UTAUT) model is widely used in the field of information systems and technology adaption research [35, 36]. This theoretical framework will be used to identify the driving factors for HR managers when it comes to technology acceptance and use of predictive HR analytics.

The theory of planned behaviour has often been used by researchers to predict human behaviour in terms of technological adoption [37]. This theoretical framework will be used to predict the dishonest behaviour of job candidates when it comes to misrepresentations in resumes to increase job offer success.

3.1 Unified Theory of Technology Acceptance and Usage

The goal of the UTAUT model is to describe technology acceptance and use from the user's perspective. The model unifies eight theories in the field of individual acceptance and technology use [38, 39].

The similarities between the unified models and theories resulted in the derivation of four constructs that are essential for this study. Performance expectancy is defined as the expected benefits received by an individual through the use of a technology for a certain purpose, while effort expectancy is conceptualized as the degree of ease. In this regard, this study wants to find out if predictive HR analytics are used because of better performance and usability. Social influence describes if an individual gets influenced by their social environment (e.g. peer groups, competitors etc.) while using a technology. Finally, facilitating conditions are defined as the degree to which resources and support are provided. Behavioural intention describes the intention of users to use a technology, while use behaviour illustrates how the technology is being used in practice [35, 36]. Time to hire, hire cost, and quality of hire are the main HR recruitment key performance indicators—these are considered together to investigate the impact of use behaviour on recruitment activities.

Age, gender, and experience have been chosen as moderator variables for this study. Gender and age moderate performance expectancy—this means that performance expectancy varies according to the differing expectations of men as well as women. Gender, age, and experience moderate effort expectancy. In terms of effort expectancy, usability and the ease of use depend on the experience of an individual when they use a certain technology. The social environments of user groups are influenced by age, gender, and experience. Finally, facilitating conditions only get moderated by age and experience. Intangible resource assets, such as knowledge and the current level of education, depend on age and experience [35, 36].

Each structural relationship in the model will be tested by hypotheses (H_1 – H_8). Each driving factor is assumed to be positively associated with behavioural intention (H_1 – H_4). Behavioural intention is positively linked with use behaviour (H_5), which again is positively linked with time to hire, hire cost, and quality of hire (H_6 – H_8). Figure 3 illustrates the structural equation model (SEM) based on an extended version of the UTAUT framework.

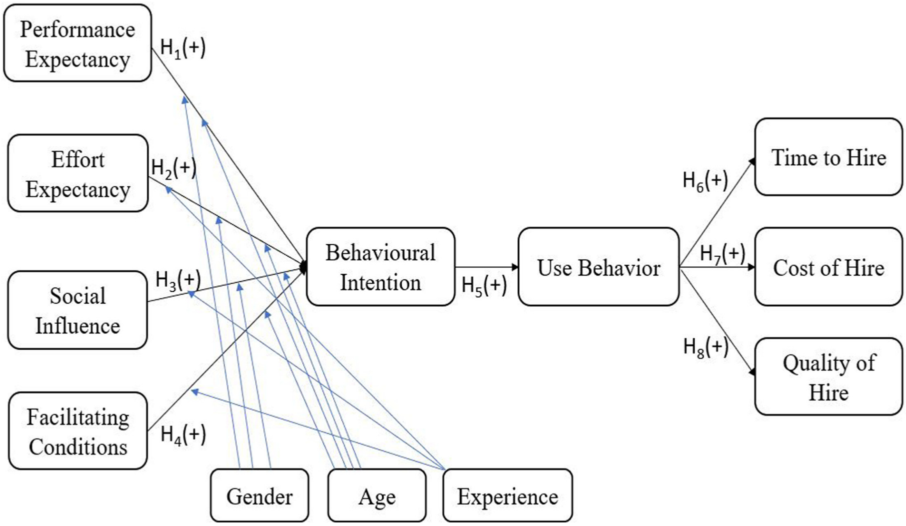


Fig. 3. SEM based on an extension of the UTAUT framework [36].

The structural relationship between performance expectancy and behavioural intention (H_1) will provide information if hiring managers and recruiters expect a benefit by using predictive HR analytics during recruitment. The impact of effort expectancy on behavioural intention (H_2), on the other hand, will be tested to find out what kind of effort hiring managers and recruiters expect by using predictive HR analytics.

Additionally, the impact of social influence on behavioural intention (H_3) will provide insights if hiring managers and recruiters are influenced by important others (e.g. competitors, peer groups etc.). The impact of facilitating conditions on behavioural intention (H_4) shows if hiring managers and recruiters expect to have sufficient resources and support during the usage of predictive HR analytics.

The relationship between behavioural intention and use behaviour (H_5) shows if predictive HR analytics are used as intended. Finally, the relationship between use behavior, time to hire, cost of hire, and quality of hire (H_6 – H_8) shows how the actual usage of predictive HR analytics impacts HR key performance indicators.

3.2 Theory of Planned Behaviour

Central to the theory of planned behaviour is the intention of an individual to perform a given behaviour. The theory consists of three independent determinants that are linked with intention and used for this study. The first is the attitude towards the behaviour. This describes the degree to which a person has a favourable or unfavourable evaluation when it comes to planned behaviour. In the context of this study, job candidates can evaluate their accessible beliefs and link their planned behaviours with specific outcomes (e.g. actions to increase job offer success rate) [40].

The second predictor is the subjective norm and describes the perceived social pressure of an individual whether to perform or not perform a certain behaviour. The last predictor is called perceived behavioural control and refers to non-motivational factors like availability of requisite opportunities and resources (e.g. abilities, collaboration with others, money, etc.). This non-motivational factor represents the perceived ease or difficulty of performing a behaviour; it describes the behavioural control of a job candidate. Finally, intentions are described as the motivational factors that influence a certain behaviour. Here behaviour refers to misrepresentation of resumes and social media profiles by job candidates to increase job offer success [40] (Fig. 4).

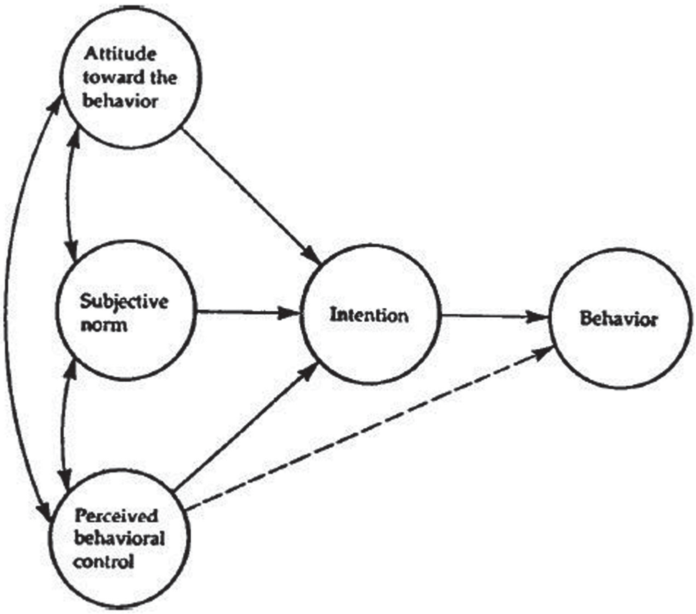


Fig. 4. Theory of planned behaviour.

4 Methodology

The research study will follow a quantitative and qualitative mixed-method approach. The quantitative approach is deductive—the UTAUT model serves as a theoretical foundation to analyse the structural relationships between the latent variables based on a partial least square's structural equation model (PLS-SEM). The empirical data matrix for the PLS-SEM model will be provided by a questionnaire distributed among hiring managers and recruiters. The reason for choosing PLS-SEM is the possibility to

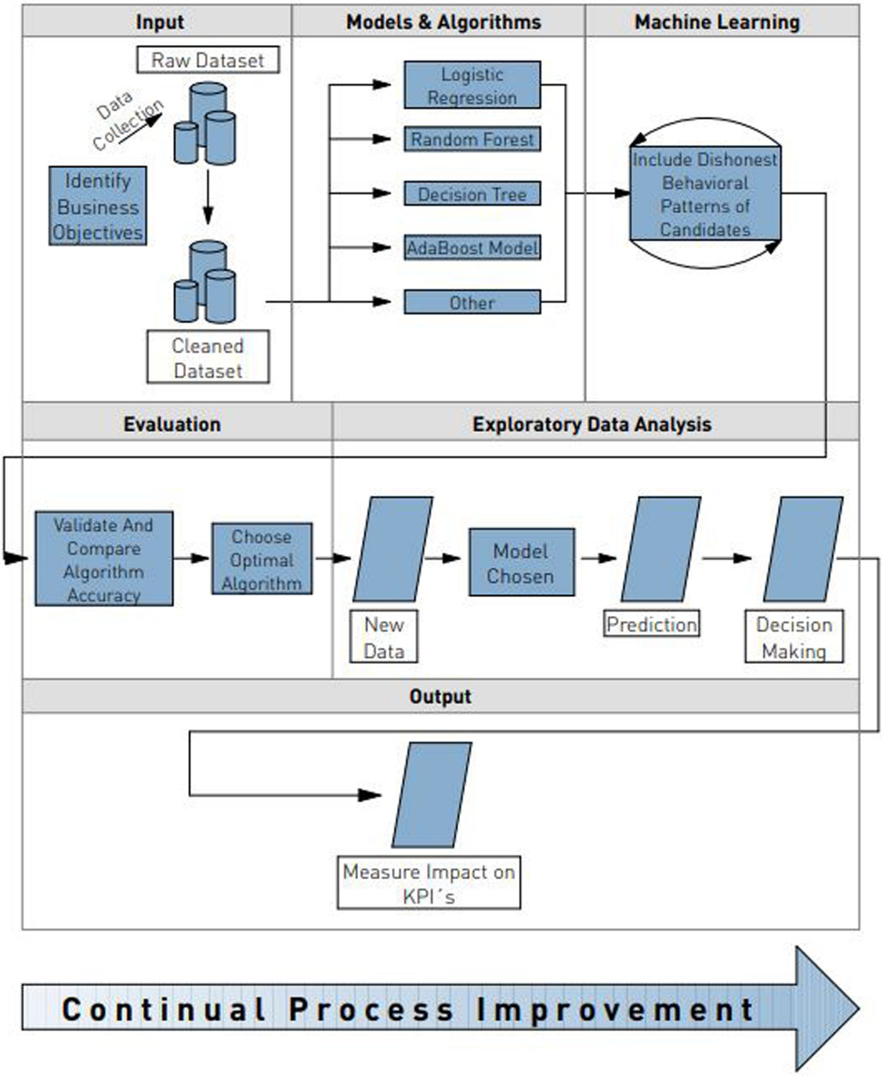


Fig. 5. Extension of the predictive modelling approach

evaluate formative and reflective constructs within a single observed measurement model. While the basic UTAUT model has reflective constructs, key HR performance indicators can be formative. Finally, based on PLS-SEM calculation, the hypotheses (H_1 – H_8) will be tested to answer what driving factors influence the behavioural intention of hiring managers (RQ1) and how does the actual use behaviour influence the key HR performance indicators (RQ2).

To identify planned behaviours by job candidates when it comes to misrepresentation of resumes on career-oriented social media profiles, semi-structured interviews will be conducted. The interview questions are based on the indicators that define the constructs of the theory of planned behaviour. By interviewing job candidates, a better understanding of the motivational and non-motivational factors leading to misrepresentation to increase job offer success (RQ3) is gained. As a data analysis technique after the conducted interviews, thematic analysis can be applied to identify and summarize planned behaviours as themes. The identified themes can be used to test and optimize machine-learning algorithms for future predictive models (RQ4).

As a result of this research study the predictive modelling approach of Fig. 2 can be extended. At the beginning, business objectives for the recruitment use case can be identified. Based on this, relevant data can be collected, filtered, and pre-processed. Then, the accuracy of algorithms regarding the identified business objectives must be validated and compared. In this stage, the identified themes of dishonest behaviour can be considered to improve machine-learning. Afterwards, the optimal algorithm with the highest accuracy is selected. Finally, the exploratory data analysis model can be chosen to predict future events. After decision-making, the impact on KPIs should be measured to check if there has been an improvement to the recruitment performance (Fig. 5).

5 Conclusion

To improve software creation, implementation, and data-modelling processes in terms of predictive HR analytics during recruitment a contribution to research is necessary. Researchers and practitioners can consider the user intention and system usage in their target group during software design and data modelling processes. Further, the relationship between use behaviour and key HR performance indicators show if predictive HR data modelling must be adjusted to achieve a desirable impact on the business.

By identifying dishonest behavioural patterns of job candidates when it comes to adjustments in resumes and career-oriented social media profiles to increase job offer success, machine-learning algorithms can be tested against these patterns. This enables researchers and data modellers to improve the analysis techniques when large datasets are analysed in relationship with recruitment activities.

Finally, SPI can be enhanced by this research from two perspectives. By considering technology acceptance and driving factors of the affected user groups, it enables practitioners to set SPI projects in the right direction when it comes to system usability. Subsequently, potential benefits in terms of usability can appear earlier.

Additionally, by the implementation of machine learning algorithms, software and data modelling processes can be constantly improved. In this relationship, SPI initiators should also consider stakeholders which are indirectly involved. In terms of predictive

HR analytics and machine learning, external stakeholders such as potential job candidates can deliver valuable data to continuously improve software and data modelling processes. By detecting dishonest behavioural patterns of job candidates, data modelling approaches can be refined and improved.

References

1. Likhitkar, P., Verma, P.: HR value proposition using predictive analytics: an overview. In: Patnaik, S., Ip, Andrew W.H., Tavana, M., Jain, V. (eds.) New Paradigm in Decision Science and Management. AISC, vol. 1005, pp. 165–171. Springer, Singapore (2020). https://doi.org/10.1007/978-981-13-9330-3_15
2. Wright, A.: Top 6 HR Technology Trends for 2018. AI, bots and digital twins will shape the year (2018). https://www.shrm.org/hr-today/news/hr-magazine/0218/pages/top-6-hr-technology-trends-for-2018.aspx. Accessed 06 Mar 2020
3. Singh, T., Malhotra, S.: Workforce analytics: increasing managerial efficiency in human resource. Int. J. Sci. Tech. Res. **9**(1), 3260–3266 (2020)
4. O'Connell, M., Kung, M.: The cost of employee turnover. Ind. Manage. **49**(1), 14–19 (2007)
5. Faroukhi, A.Z., El Alaoui, I., Gahi, Y., Amine, A.: Big data monetization throughout Big Data Value Chain: a comprehensive review. J. Big Data **7**(1), 1–22 (2020). https://doi.org/10.1186/s40537-019-0281-5
6. Aswale, N., Mukul, K.: Role of data analytics in human resource management for prediction of attrition using job satisfaction. In: Sharma, N., Chakrabarti, A., Balas, V.E. (eds.) Data Management, Analytics and Innovation. AISC, vol. 1042, pp. 57–67. Springer, Singapore (2020). https://doi.org/10.1007/978-981-32-9949-8_5
7. Kakkar, H., Kaushik, S.: Technology driven human resource management – a strategic perspective. Int. J. Emerg. Technol. **10**(1a), 179–184 (2019)
8. Mahmoud, A., Shawabkeh, T., Salameh, W. et al.: Performance predicting in hiring process and performance appraisals using machine learning. In: International Conference on Information and Communication Systems, ICICS, Irbid, pp. 110–115. IEEE (2019)
9. Kumar, V., Garg, M.L.: Predictive analytics: a review of trends and techniques. Int. J. Comput. Appl. **182**(1), 31–37 (2018)
10. Zehir, C., Karaboğa, T., Başar, D.: The transformation of human resource management and its impact on overall business performance: Big Data analytics and ai technologies in strategic HRM. In: Hacioglu, U. (ed.) Digital Business Strategies in Blockchain Ecosystems. CMS, pp. 265–279. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-29739-8_12
11. Sivathanu, B., Pillai, R.: Smart HR 4.0 – how industry 4.0 is disrupting HR. Hum. Resour. Manag. Int. Dig. **26**(4), 7–11 (2018)
12. Pries-Heje, J., Johansen, J.: SPI Manifesto, Version A.1.2.2010 (2010)
13. Morris, M., Dillon, M.: The influence of user perceptions on software utilization: application and evaluation of a theoretical model of technology acceptance. IEEE Trans. Softw. Eng. **14**(4), 58–65 (1997)
14. Leicht-Deobald, U., et al.: The challenges of algorithm-based HR decision-making for personal integrity. J. Bus. Ethics **160**(2), 377–392 (2019). https://doi.org/10.1007/s10551-019-04204-w
15. Renuka Devi, B., Vijaya Banu, P.: Introduction to recruitment. SSRG Int. J. Econ. Manage. Stud. **1**(2), 5–8 (2014)

16. Devaro, J.: Internal hiring or external recruitment? The efficacy of internal or external hiring hinges on other policies that a firm uses simultaneously. IZA World of Labor, p. NA (2016)
17. Edwards, M., Edwards, K.: Predictive HR Analytics. Mastering the HR Metric. 2nd edn. Publisher, New York (2019)
18. Bohnet, I.: How to take the bias out of interviews. Harv. Bus. Rev. (2016). https://hbr.org/2016/04/how-to-take-the-bias-out-of-interviews. Accessed 20 Mar 2020
19. Huselid, M.: The science and practice of workforce analytics: introduction to the HRM special issue. Hum. Resou. Manage. (Special Issue: Workforce Analytics) **57**(3), 679–684 (2018)
20. Greasley, K., Thomas, P.: HR analytics: The onto-epistemology and politics of metricised HRM. Hum. Resour. Manage. J. 1–14 (2020). https://doi.org/10.1111/1748-8583.12283
21. Buettner, R.: Prädiktive Algorithmen zur Persönlichkeitsprognose auf Basis von Social-Media-Daten. Personalquartaly Wissenschaftsjournal für die Personalpraxis **3**, 22–27 (2017)
22. Gou, L., Zhou, M., Yang, H.: KnowMe and ShareMe: understanding automatically discovered personality traits from social media and user sharing preferences. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, SIGCHI, CHI, Toronto, pp. 955–964 (2014)
23. Kristof-Brown, A., Guay, R.P.: Person-environment fit. In: Zedeck, S. (eds.) APA Handbook of Industrial and Organizational Psychology, vol. 3, pp. 3–50. American Psychological Association (2011)
24. Kristof, A.: Person-organization fit: an integrative review of its conceptualizations, measurement, and implications. Pers. Psychol. **49**(1), 1–49 (1996)
25. Seong, J.Y., Kristof-Brown, A., Park, W.W., et al.: Person-group fit diversity antecedents proximal outcomes and performance at the group level. J. Manage. **41**(4), 1184–1213 (2015)
26. Buettner, R.: A framework for recommender systems in online social network recruiting: an interdisciplinary call to arms. In: 47th Hawaii International Conference on System Science, Hawaii, pp. 1415–1424. IEEE Computer Society (2014)
27. Faliagka, E., Tsakalidis, A., Tzimas, G.: An integrated e-recruitment system for automated personality mining and applicant ranking. Internet Res. **22**(5), 551–568 (2012)
28. Buettner, R.: Abschlussbericht zum BMBF Forschungsprojekt. Effizientes Recruiting von Fachkräften im Web 2.0 (EfficientRecruiting 2.0): Hoch-automatisierte Identifikation und Rekrutierung von Fachkräften durch Analyse internetbasierter sozialer Netzwerke mittels intelligenter Softwareagenten. Technical report (2017). https://www.prof-buettner.com/downloads/buettner2017b.pdf. Accessed 31 Mar 2020
29. Vasa, J., Masrani, K.: Foreseeing employee attritions using diverse data mining strategies. Int. J. Recent Tech. Eng. **8**(3), 620–626 (2019)
30. Mohammed, A.Q.: HR analytics: a modern tool in HR for predictive decision making. J. Manag. **6**(3), 51–63 (2019)
31. Seuwou, P., Banissi, E., Ubakanma, G.: User acceptance of information technology: a critical review of technology acceptance models and the decision to invest in information security. In: Jahankhani, H., Carlile, A., Emm, D., Hosseinian-Far, A., Brown, G., Sexton, G., Jamal, A. (eds.) ICGS3 2017. CCIS, vol. 630, pp. 230–251. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-51064-4_19
32. Hireright Homepage. https://www.hireright.com/news/press-release/hireright-survey-finds-88-percent-of-employers-have-found-a-misrepresentati. Accessed 03 Apr 2020
33. Clark, J.: The perfect resume. Air Med. J. **36**, 13–15 (2017)
34. Henle, C.A., Dineen, B.R., Duffy, M.K.: Assessing intentional resume deception: development and nomological network of a resume fraud measure. J. Bus. Psychol. **34**(1), 87–106 (2017). https://doi.org/10.1007/s10869-017-9527-4

35. Sohn, K., Kwon, O.: Technology acceptance theories and factors influencing artificial Intelligence-based intelligent products. Telematics Inform. **47**, 1–14 (2020)
36. Venkatesh, V., Thong, J., Xu, X.: Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q. **36**(1), 157–178 (2012)
37. Lai, P.: The Literature Review of Technology Adoption Models And Theories For The Novelty Technology. J. Inf. Syst. Tech. Manage. JISTEM **14**(1), 21–38 (2017)
38. Deng, S., Liu, Y., Qi, Y.: An empirical study on determinants of web based question-answer services adoption. Online Inf. Rev. Bradford **35**(5), 789–798 (2011)
39. Venkatesh, V., Morris, M., Davis, G., et al.: User acceptance of information technology toward a unified view. MIS Q. **27**(3), 425–478 (2003)
40. Ajzen, I.: The theory of planned behavior. Organ. Behav. Hum. Decis. Process. **50**(2), 179–211 (1991)

Systematic Agile Development in Regulated Environments

Alexander Poth[1(✉)], Jan Jacobsen[2(✉)], and Andreas Riel[3(✉)]

[1] Volkswagen AG, Berliner Ring 2, 38436 Wolfsburg, Germany
alexander.poth@volkswagen.de
[2] Volkswagen Financial Services AG, 38112 Brunswick, Germany
jan.jacobsen@vwfs.com
[3] Université Grenoble Alpes, CNRS, Grenoble INP, G-SCOP, 38031 Grenoble, France
andreas.riel@grenoble-inp.fr

Abstract. For established domains within highly regulated environments, a systematic approach is needed to scale agile methods and assure compliance with regulatory requirements. This paper presents a structured method to designing such a systematic approach which works adequately in small agile teams. It is independent of the underlying method such as Scrum, Kanban, etc., and is scalable to more and bigger teams or even entire subsidiaries. It is based on a compliance and a quality risk dimension which are both needed to fit regulatory requirements. The approach has been validated in the financial IT context with more than 100 developers in one subsidiary.

Keywords: Software development management · Agile software development · Regulation compliance · Large scaling agile

1 Introduction

Established industry sectors are more or less regulated. Less regulated sectors solely have to incorporate basic requirements like European Union regulation, i.e. the General Data Protection Regulation (GDPR), and/or national requirements such as the German Commercial Code (HGB). In highly regulated sectors however, products and services have to comply with further extensive standards and regulations. The financial sector for example has to fulfill regulations imposed by the EU countries' national supervisory authorities, as well as Minimum Requirements for Risk Management for financial institutions (MaRisk) in Germany.

Many regulations are domain-specific like medical, finance or automotive. However, regulations have some common aspects like quality assurance evidences for verification and validation which demand a more or less stringent traceability and risk management [1]. A huge body of documentation exists to handle regulation and compliance. However, these works mostly focus on a specific solution or aspect within the respective domain. This leads to partial [2] and inconsistent [3] agile adoptions [4] like ScrumBut.

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 191–202, 2020.
https://doi.org/10.1007/978-3-030-56441-4_14

Our research objective is to design a framework that can be used to derive a specific compliance guideline offering as much autonomy to agile teams as possible by fitting the required specific regulations of the product or service with its organization. As for evidences for the effectiveness of the framework with its derived specific instantiation in the form of a "compliance guideline", we want to meet the following three core requirements. First, the **external confirmation** by audits with focus on compliance shall be facilitated. Second, **the delivery** of the demanded business value shall not be hampered and remain an essential part of the outcome flow. Third, the framework shall be initially set up for the specific domain and later be **adapted to new regulations** over time.

This article presents the framework and its building blocks based on practical examples in the finance domain. It demonstrates the framework's application for deriving a specific "compliance guideline", and subsequently shows the evidences for the established example. The generalization of the framework is assured by design thanks to its independence of any specific regulation.

Section 2 gives an overview about existing approaches, Sect. 3 introduces the Level of Done, Sect. 4 combines the approach with Product Quality Risks, Sect. 5 presents a case study, Sect. 6 describes the continuous improvement of the approach instantiation, Sect. 7 shows the contributions and Sect. 8 concludes.

2 Agile Approaches and Artifacts for Handling Regulation

Many regulations are domain-specific like medical, finance or automotive. However, regulations have some common aspects like quality assurance evidences for verification and validation which demand a more or less stringent traceability and risk management [1]. A huge body of literature and documentation exists to handle safety regulations and compliance in the context of agile development like [5–8]. For other domains some works exist like [9–12] for finance. However, these works mostly focus on a specific solution or aspect within the respective domain. We could not find a published generic approach to handling regulatory requirements within an agile organization setting. To address this gap, we identified additional approaches [13, 14] and artifacts [15, 16] with some relevance to solve this generic problem for handling the systematic generation of outcomes demanded by regulations in a way as agile as possible. The focus is on making sure that regulatory requirements are treated systematically [17]. The product or service engineering artefacts like [18–20] are required by regulation as outcomes for evidences. Furthermore, to show sustainability in operational excellence [21, 22] about effectivity and risk-management, continuous improvement is relevant [23–26]. A basic approach for rigor task handling from open to done/closed is Kanban [27] which can be usable to ensure systematic regulation task handling. The current gap which our work wants to address is that no generic approach is established to derive a systematic business domain framework for handling regulations systematically and offer as much as autonomy to teams as possible.

3 Scaling Conformity to Regulations via Levels of Done

The development process has to address two dimensions. The domain dimension handles the organizational and procedural compliance requirements. It has to assure that the compliance requirements be fulfilled at least at the latest required point in the product or service life cycle. Earlier assurance of regulatory requirements is possible and a part of the team's self-organization. The product specific dimension helps teams identify and realize their product specific quality-risk requirements. Within this dimension, the team handles product or service specific quality-risks in a structured and transparent manner to assure an adequate risk management. For handling the product specific quality risks, the Product Quality Risk (PQR) [28] approach is used. As the two dimensions are independent, we considered handling them by two individual approaches too - the LoD and PQR approaches make it explicit that a teams have to care about them explicit.

To ensure a development process as lean and agile as possible, it is important to set only a minimum predefined framework and take over as much outcome specific refinement as possible to the teams. However, the process has to assure a systematic handling of the team's refinement work. To realize this, handling a systematic refinement of the outcome's value is needed. The outcome's value is assessed by its (inherent) quality risks. The systematic product or service quality risk identification and handling is described in [28], which can be used to assure that the development process does not lose outcome focus by its process orientation. In [28], the product capabilities and features are used to derive the product specific quality risks. Based on the identified and prioritized quality risks, adequate mitigation actions are scheduled during the development to ensure a compliant and high quality outcome.

To assure that product teams incorporate both dimensions just in time, we propose a Levels of Done (LoD) approach. LoD are an enriched variant of the Definition of Done (DoD) of Scrum that is aligned with requirements [29] at defined milestones in the development process. The LoD approach uses the concept of boundaries [30]. The boundary concept is applied beyond the sprint time-box between Definition of Ready (DoR) and DoD for the LoD approach – it is applied to all take-overs in a value chain. This makes it simple and independent from any specific agile approach based on sprints, as well as sufficiently generic for different domains of regulation and the specific with their specific check-points they require. This is necessary to fulfill a systematic product and process quality approach demanded by most quality related standards and to allow agile scaling while staying effective [31].

4 The LoD-PQR Approach

While in a traditional compliance scope, the software development life-cycle (SDLC) is clearly defined by a comprehensive set of fixed requirements and deliverables prior to project start, agile environments require a more methodical way of defining LoD. Figure 1 depicts the main elements of the LoD-PQR approach we propose:

Identify all relevant regulations and standards of your enterprise for compliant products and/or services.

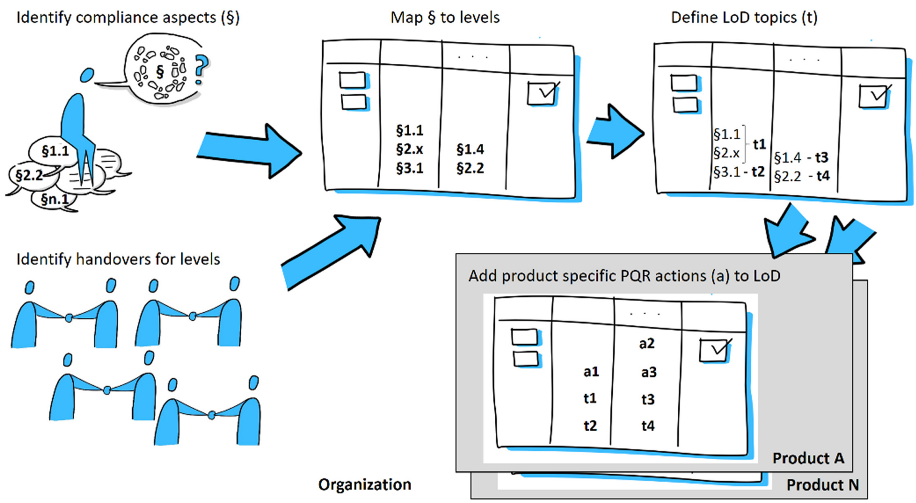


Fig. 1. Schematic picture of a practical LoD-PQR method application scenario.

Identify how many stages you have for product development via a Kanban board.

The Kanban board helps to identify handover-points in a work stream. These points are the most relevant for LoD.

According to Conway's law [32], the structure of an origination drives their outcomes. Therefore, alignment of the "planned" outcome architecture with the organization shall be considered. This should also drive future changes to an existing LoD to support the transformation in a pull-fashion. The LoD does not refine the internal team organization between two stages. The teams can apply their preferred agile approach like Scrum, Kanban etc. in their self-organized working flows to fit the next stage.

Enabling teams to choose the most effective ways to comply with regulatory relevant outcomes by mapping them to the stages of the Kanban board.

A transparent traceability from the regulation to the LoD will facilitate regulation adoption. However, finding adequate implementations should be delegated to the team to give them freedom to find solutions that fit into their particular context. The openness about the how to reach the outcomes give the teams the autonomy to work as it is best for their specific demands and the mastery (responsibility) about their implementations. The traceability from the external requirements to their internal representations – the topics in Fig. 1 – shall be established to avoid interpretations by missing "root" and to avoid non-value adding activities in a lean context.

Reduce the outcomes of "chains" to the last outcome for a shorter list.

To optimize the LoD, chains of dependencies can be reduced to the latest outcome. For example, a separate test protocol is not needed if the test result log and protocols are saved as part of the comprehensive deployment-log and stored in an auditable way. This is covered by an underlying internal control system.

Provide additional information about practices and work instructions about outcomes for assisting the teams. To help the teams for a fast instantiation, a practice collection can be provided as inspiration how other solved it. If a new practice is identified, it will be added to the practice collection to leverage continuous improvement and replacement of outdated practices.

Add the PQR dimension to assure that products and services have a comprehensive quality approach. To derive systematically the specific PQR a self-service kit for the teams is recommended as described in [33].

While the LoD covers only formal regulation requirements, the PQR method handles business risks related to deliverables by quality related mitigation actions as described in [28] and [33]. These mitigation actions are mapped to the corresponding stages and handled by the teams. Based on the regulation and quality risk dimension a holistic quality management system can be established. Figure 1 shows how the actions fit together in a product team specific instantiation. It visualizes the instantiation of the 4 LoDs, the product team specific PQRs actions (a) on top of the organization-wide valid LoD topics (t), as well as the numerous product checks.

The LoD-PQR approach is easily repeatable for the iterative and incremental development in agile product teams. It also foresees cross-team reviews conducted by technical reviewers (IT experts) providing evidence of compliance with the LoD. Quality standards covered in the reviews include: architecture, code quality, PQR, security, documentation, etc. Every topic has its own LoD acceptance criteria. Depending on the technical review result, the accountable role (e.g. Head of IT) grants technical approval for the product release (Fig. 2).

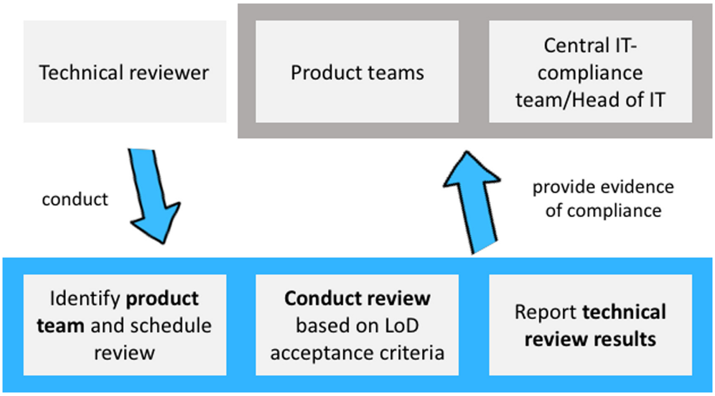


Fig. 2. LoD compliance process and involved stakeholders.

For the review and approval process as well as the LoD, internal criteria shall be derived. The control owner shall establish a monitoring on the whole process against these criteria (via preventive gates and/or detective post-checks) in order to conduct appropriate actions depending on the level of conformance and control effectiveness.

Derivations to the LoD shall be assessed and tracked to sign-off by the risk owners. Teams "pull" experts for specific standards for support in case of new or special issues. Experts for specific standards are "pulled" by the teams to support in case of new or special issues. Any regulation changes shall be integrated into the LoD as soon as possible and all teams have to ensure to fulfill the current version as soon as possible. Teams are autonomous to set synchronization points in case of inter-team dependencies. The time span between the different levels of the LoD in a team mostly depends on the team's specific delivery frequency. Some teams need weeks, others months. The LoD does not depend on the team's delivery cycle duration.

5 Case Study: Instantiation, Deployment and Its Limitations

The Volkswagen Financial Services AG Digital Unit Berlin – short DU - identified four stages for their LoD (Fig. 1). First, the business takes over the stories into the team. Second, the team implements the requirements according to compliance for security etc. Third, the product is checked for compliance and business process integration. Finally, the product's functionality is verified during operation. The last stage is interesting for the handover in cases were no DevOps is applied.

The identified regulations and standards for the financial domain are defined by the European Union and are instantiated by German governance and regulation institutions like the MaRisk, BAIT or GDPR.

To distinguish the different levels of compliance checks, an organization defines preventive and detective checks about the application of the LoD. A preventive check is conducted before a productive deployment and detective check is a post-deployment compliance check. To assure LoD compliance the DU adopted the approach from Fig. 2 with some refinements for adequate review sampling and time (pre- or post-deployment).

The decision of a pre- or post-deployment check is based on Table 1 (Check time). Relevant parameters are the product risks form the deployment, the team maturity and the accountability of the team. Based on the parameter value combination the check is done with a high sample size and preventive (bevor deployment) down to a low sample size and detective (after deployment). The sampling principle is derived from [34], which is international recognized.

Table 1. Decision-table for selection type of compliance check.

Product quality risks	Team accountability	Team maturity	Sample size	Check time
Mid, high	Low	Low	High (like 100%)	Preventive
Mid, high	Low	Mid	Mid (like 75%)	Detective
Mid, high	Mid	Low	Mid (like 75%)	Detective
Low	Low	Low	Low (like 10%)	Detective
Low, mid, high	Low, mid, high	High	Low (like 10%)	Detective

Low conformation leads (based on low team accountability and maturity) to the need of a high confidence-coefficient based on established audit practices [34]. The ratings from low to high are clear defined for the DU delivery teams – this is a domain specific detail and out of scope of the example case, which shows on a generic level how to check compliance of deliveries. A preventive compliance check "interrupts" the continuous delivery chain. A higher sample size leads to more effort for checks.

The LoD of the DU has been developed by a cross-functional team. The team incorporated experts from the headquarters compliance, headquarters security, business and development teams, as well as external experts from the Volkswagen AG. Reflections with external consults (agile coaches, auditors etc.) were done cyclically too. Throughout the almost one year development period, the team allocated approximately 6–7 experts.

In the last 3 years we established and enhanced the approach for more efficient delivery and to regulation updates with the Scrum masters and the teams. Currently more than 100 developers working with the LoD-PQR approach. These facts validate the generic approach. Currently other locations (countries) and units are in the adoption phase within the DU as a Service template as scaling model.

The application to the DU financial case revealed the following limitations of the LoD-PQR approach:

- The governance has to ensure the correct outcomes for the
  - compliance requirements, as well as
  - expected deliverable which creates the customer/user value.
- The governance has to ensure the update of the
  - LoD which has to be ensured by the regulation experts;
  - PQR which has to be ensured by the product or service experts.

These limitations are partly addressed by the review procedure (Fig. 2), which however generates a base workload scaling linearly with the delivery frequency of the products and services. To reduce this linear correlation of reviews to deliveries, a team maturity approach can be established. Higher team maturity leads to more autonomy and thus reliefs the team from having mandatory pre-deployment LoD-triggered technical reviews by team-independent reviewers.

Additional limitations are that the approach was applied in a new subsidiary without established structures which made it easier to establish new practices without being confronted with change management issues.

6 Enhance LoD as a Genuine Learning Approach

The LoD-PQR approach behaves as an infinitely genuine learning approach. Once the rollout of the LoD-PQR is established, it is kept open for enhancements. The objective was to break free from overloaded compliance processes leading to high overhead by introducing a new paradigm for achieving a new steady state apt for quick and effective response to external and internal feedback. Messing up is a necessary part of learning, so a safe-to-fail environment may help. The challenge is to encourage, precipitate, and then support that paradigm shift. To gain high acceptance during the definition and the dynamic scale-up phase of the DU, the following environment has been established:

- Train the LoD-PQR approach to the product teams and the product owners with role-specific training kits and coaching. The training is for an initial knowledge setup. The coaching and moderation is used in specific phases like the identification of the PQRs.
- As proposed in [35], the DU established a Community of Practice (CoP) LoD to assure a continuous feedback, improvement environment and to identify which work not to do. The CoP is legitimated (by the management) to change content of the LoD and release it as a new version to react fast to new demands.
- Assure "sufficient pressure" in the organization to establish the LoD-PQR approach in each product team. This was initially assured with the "invitation" of the central compliance department of the Financial Service AG to conduct a compliance audit/review of the DU Berlin. For continuous "alignment" of the product teams, randomized compliance checks are "announced" – as recommended in [36] – by the internal compliance team of the DU. Additionally, an external compliance audit could be announced at any time by external audit execution instances on behalf of the European Central Bank.

Lessons learned from the application of the LoD-PQR approach in many releases and teams are:

- Risk evaluation external – the definition of the risk ratings are made for the entire organization by the governance. The teams made the evaluation on their own and independently because they can estimate product and services best with their business partners.
- Estimations for packages done by DU external teams as "delivery modules" – to scale better, partner companies are support the DU teams. The outcomes of the external suppliers have to be the same quality standards as internal outcomes thus the LoD-PQR reviews have to be fulfilled by them as well.
- KPIs – for transparency, the governance established KPIs about the risks portfolio, the specific derivations in the teams products and services as well as their sign offs.

In the last 3 years we established and enhanced the approach for more efficient delivery and to regulation updates with the Scrum masters and the teams. Compliance audits from the headquarters compliance team and external auditors (which often work for the European Central Bank as their independent compliance auditors) confirmed the approach. Currently more than 100 developers working with the LoD-PQR approach. These facts validate the generic approach. Currently other locations (countries) and units are in the a adoption phase within the DU as a Service template as scaling model.

7 Contributions

The contributions of the LoD-PQR approach is on a theoretical view-point:

- Separation of actions for product specific quality risks from organizational regulation and compliance actions (two dimensions)
- Alignment of the organizational actions with the established value stream and its handovers (amount of levels)
- Generic approach to derive organization specific LoDs

The contribution of the LoD-PQR approach is on a practical view-point:

- Minimize the organizational requirements of the development process to give autonomy to the agile teams (comes with responsibility aligned with team capability)
- Option for the teams to optimize their specific workflows by automation pipelines
- Effort for GRC scales with amount of teams and validation effort is reduced with team capabilities and maturity.

8 Discussion and Conclusion

The LoD-PQR approach addresses the demand for a generic approach to handling regulation requirements and product specific quality management in an agile environment. While we have shown the generic LoD-PQR method application to the European finance domain, other domain specific requirements would need to be identified, e.g. for the DO-178 (avionics safety) or ISO 26262 (automotive safety). However, the amount of regulation requirements in finance was lower than initially expected, approximately 50 with direct impact to the software development. The product specific PQRs strongly depend on the outcomes, however the workload which can be handled by a team is a “limiting factor”.

The acceptance of our presented methodology within the agile-teams was encouraged by the committed degree of freedom. In our case, we have witnessed that implementing the LoD-PQR approach supported the teams to navigate through the complex compliance requirements in our domain in a lean way (conformity). Our approach enabled the product teams to realize efficiency by design and to share techniques how to implement compliance requirements in an uncomplicated way. Besides, the genuine learning character of the LoD-PQR approach leads to streamlined development processes of the approach itself, leading to a positive impact on process performance.

To reduce the number of governance-triggered compliance reviews, a team maturity model approach is currently in evaluation to make the LoD-PQR method scale more efficiently from a governance perspective. Furthermore, automated checks in the development pipeline of LoD compliance indicators are planned [37].

References

1. Fitzgerald, B., Stol, K.-J., O'Sullivan, R., O'Brien, D.: Scaling agile methods to regulated environments: an industry case study. In: 2013 35th International Conference on Software Engineering (ICSE), pp. 863–872. IEEE (2013)
2. Karvonen, T., Sharp, H., Barroca, L.: Enterprise agility: why is transformation so hard? In: Garbajosa, J., Wang, X., Aguiar, A. (eds.) XP 2018. LNBIP, vol. 314, pp. 131–145. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-91602-6_9
3. Uludag, O., Kleehaus, M., Caprano, C., Matthes, F.: Identifying and structuring challenges in large-scale agile development based on a structured literature review. In: 2018 IEEE 22nd International Enterprise Distributed Object Computing Conference (EDOC), pp. 191–197. IEEE (2018)
4. Eloranta, V.P., Koskimies, K., Mikkonen, T.: Exploring ScrumBut—an empirical study of scrum anti-patterns. Inf. Softw. Technol. **74**, 194–203 (2016)
5. Ge, X., Paige, R.F., McDermid, J.A.: An iterative approach for development of safety-critical software and safety arguments. In: 2010 Agile Conference, Orlando, FL, pp. 35–43 (2010). https://doi.org/10.1109/AGILE.2010.10
6. Wolff, S.: Scrum goes formal: agile methods for safety-critical systems. In: 2012 First International Workshop on Formal Methods in Software Engineering: Rigorous and Agile Approaches (FormSERA), Zurich, pp. 23–29 (2012). https://doi.org/10.1109/formsera.2012.6229784
7. Stålhane, T., Hanssen, G.K., Myklebust, T., Haugset, B.: Agile change impact analysis of safety critical software. In: Bondavalli, A., Ceccarelli, A., Ortmeier, F. (eds.) SAFECOMP 2014. LNCS, vol. 8696, pp. 444–454. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10557-4_48
8. Stålhane, T., Myklebust, T.: The agile safety case. In: Skavhaug, A., Guiochet, J., Schoitsch, E., Bitsch, F. (eds.) SAFECOMP 2016. LNCS, vol. 9923, pp. 5–16. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-45480-1_1
9. Vejseli, S., Proba, D., Rossmann, A., Jung, R.: The agile strategies in IT governance: towards a framework of agile IT Governance in the banking industry. Research Papers, p. 148 (2018)
10. Christou, I., Ponis, S., Palaiologou, E.: Using the agile unified process in banking. IEEE Softw. **27**(3), 72–79 (2010). https://doi.org/10.1109/ms.2009.156
11. Poth, A., Wolf, F.: Agile procedures of an automotive OEM – views from different business areas. In: Stolfa, J., Stolfa, S., O'Connor, Rory V., Messnarz, R. (eds.) EuroSPI 2017. CCIS, vol. 748, pp. 513–522. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64218-5_42
12. Sirkiä, R., Laanti, M.: Adaptive finance and control: combining lean, agile, and beyond budgeting for financial and organizational flexibility. In: 48th Hawaii International Conference on System Sciences, Kauai, HI, pp. 5030–5037 (2015). https://doi.org/10.1109/hicss.2015.596
13. Dikert, K., Paasivaara, M., Lassenius, C.: Challenges and success factors for large-scale agile transformations: a systematic literature review. J. Syst. Softw. **119**, 87–108 (2016)
14. Pries-Heje, J., Krohn, M.M.: The safe way to the agile organization. In: Proceedings of the XP2017 Scientific Workshops, p. 18. ACM (2017)
15. Kiv, S., Heng, S., Kolp, M., Wautelet, Y.: Agile manifesto and practices selection for tailoring software development: a systematic literature review. In: Kuhrmann, M., et al. (eds.) PROFES 2018. LNCS, vol. 11271, pp. 12–30. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-03673-7_2

16. Hoda, R., Noble, J.: Becoming agile: a grounded theory of agile transitions in practice. In: Proceedings of the 39th International Conference on Software Engineering, pp. 141–151. IEEE (2017)
17. Pichler, M., Rumetshofer, H., Wahler, W.: Agile requirements engineering for a social insurance for occupational risks organization: a case study. In: 14th IEEE International Requirements Engineering Conference (RE 2006), Minneapolis/St. Paul, MN, pp. 251–256 (2006). https://doi.org/10.1109/re.2006.8
18. Putta, A., Paasivaara, M., Lassenius, C.: How are agile release trains formed in practice? a case study in a large financial corporation. In: Kruchten, P., Fraser, S., Coallier, F. (eds.) XP 2019. LNBIP, vol. 355, pp. 154–170. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-19034-7_10
19. Diebold, P., Dahlem, M.: Agile practices in practice: a mapping study. In: Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering (EASE 2014). Association for Computing Machinery, New York, Article 30, pp. 1–10 (2014). https://doi.org/10.1145/2601248.2601254
20. Bowers, A.N., Sangwan, R.S., Neill, C.J.: Adoption of XP practices in the industry—a survey: research sections. Softw. Process **12**(3), 283–294 (2007)
21. Poth, A., Kottke, M., Riel, A.: Scaling agile – a large enterprise view on delivering and ensuring sustainable transitions. In: Przybyłek, A., Morales-Trujillo, M.E. (eds.) LASD/MIDI -2019. LNBIP, vol. 376, pp. 1–18. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-37534-8_1
22. Carvalho, A.M., Sampaio, P., Rebentisch, E., Carvalho, J.Á., Saraiva, P.: Operational excellence, organisational culture and agility: the missing link? J. Total Qual. Manag. Bus. Excell. **15**, 1–20 (2017)
23. Ringstad, M.A., Dingsøyr, T., Brede Moe, N.: Agile process improvement: diagnosis and planning to improve teamwork. In: O'Connor, R.V., Pries-Heje, J., Messnarz, R. (eds.) EuroSPI 2011. CCIS, vol. 172, pp. 167–178. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-22206-1_15
24. Santana, C., Queiroz, F., Vasconcelos, A., Gusmao, C.: Software process improvement in agile software development: a systematic literature review. In: 41st Euromicro Conference on Software Engineering and Advanced Applications, pp. 325–332 (2015)
25. Teece, D.: Dynamic capabilities and organizational agility: risk, uncertainty, and strategy in the innovation economy. Calif. Manag. Rev. **58**(4), 13–36 (2016)
26. Brown, A.: Managing challenges in sustaining business excellence. Int. J. Qual. Reliab. Manag. **30**(4), 461–475 (2013)
27. Kumar, C.S., Panneerselvam, R.: Literature review of JIT-KANBAN system. Int. J. Adv. Manuf. Technol. **32**(3–4), 393–408 (2007)
28. Poth, A., Sunyaev, A.: Effective quality management: risk- and value-based software quality management. IEEE Softw. **31**(6), 79–85 (2014)
29. Perkusich, M., et al.: A systematic review on the use of definition of done on agile software development projects. In: International Conference on Evaluation and Assessment in Software Engineering (EASE) (2017). https://doi.org/10.1145/3084226.3084262
30. Power, K.: Definition of ready: an experience report from teams at cisco. In: Cantone, G., Marchesi, M. (eds.) XP 2014. LNBIP, vol. 179, pp. 312–319. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-06862-6_25
31. Poth, A.: Effectivity and economical aspects for agile quality assurance in large enterprises. J. Softw. Process: Improve. Pract. **28**(11), 1000–1004 (2016)
32. Conway, M.E.: How do committees invent? Datamation **14**(5), 28–31 (1968)

33. Poth, A., Riel A.: Quality requirements elicitation by ideation of product quality risks with design thinking. In: Proceedings of the 28th IEEE International Requirements Engineering Conference (RE'20), Vienna (2020, in print)
34. ISACA audit framework: ITAFTM: A Professional Practices Framework for IS Audit/Assurance, 3 edn. Section: 2208 Sampling. www.isaca.org/ITAF
35. Paasivaara, M., Lassenius, C.: Communities of practice in a large distributed agile software development organization - case Ericsson. Inform. Softw. Technol. **56**, 1556–1577 (2014)
36. Poth, A., Kottke, M.: How to assure agile method and process alignment in an organization? In: Larrucea, X., Santamaria, I., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2018. CCIS, vol. 896, pp. 421–425. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-97925-0_35
37. Kösling, M., Poth, A.: Agile development offers the chance to establish automated quality procedures. In: Stolfa, J., Stolfa, S., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2017. CCIS, vol. 748, pp. 495–503. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64218-5_40

Agile Software Development – Do We Really Calculate the Costs? A Multivocal Literature Review

Aidan Fogarty[1], Aaron Edgeworth[1], Oishin Smith[1], Michael Dowling[1], Murat Yilmaz[1,2], Silvana Togneri MacMahon[1,2], and Paul Clarke[1,2](✉)

[1] School of Computing, Dublin City University, Dublin, Ireland
{aidan.fogarty23, aaron.edgeworth3, oishin.smith25, michael.dowling35}@mail.dcu.ie, {murat.yilmaz, silvana.macmahon, paul.m.clarke}@dcu.ie
[2] Lero – the Science Foundation Ireland Research Centre for Software, Limerick, Ireland

Abstract. Agile software development methods, in their various different forms, have become the basis for most software projects in today's world. The methodology is present in almost all organisations today. However, despite the popularity, failure rates in software projects remain high. This paper identifies why agile methodologies have become so successful. In addition, the paper discusses certain factors that may often be overlooked in organisations that have adopted agile methods, such as rework, maintainability, adoption, turnover rates and the potential costs associated with each. The research carried out was a multivocal literature review (MLR). Multiple white and grey literature which was deemed to be relevant was selected. 32 contributions from white literature were selected for use in the review as well as 8 from grey literature sources. We find that while agile has many advantages, organisations may overlook the potential downsides of using an agile methodology. If not managed or implemented correctly, organisations risk taking on more hidden and expensive costs, for example in relation to rework. It is important that organisations are sufficiently trained in agile methods in order to succeed.

Keywords: Agile software development · Rework · Agile costs · Agile success

1 Introduction

Many agile software development practices may predate the agile manifesto, but nevertheless, since its inception in 2001, agile software development (or for short, agile) has grown in popularity [43]. Many common metrics used to measure the success of a project, such as Business Value Delivered, Customer Satisfaction and Earned Value, show an improvement when agile systems are implemented [1]. From a business standpoint these are positive and in line with what the organization is seeking to

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 203–219, 2020.
https://doi.org/10.1007/978-3-030-56441-4_15

improve. However, as agile has become more ubiquitous, some organisations risk overlooking the potential costs that agile may bring.

It is undeniable that agile has proven to be an effective way to manage software projects, but this may contribute to it simplified reputation as a "one size fits all" solution to software process, even though there are many situational factors that affect the development process [44], and a need to continually adapt it has been shown to be beneficial for business success [56]. More generally, it has been suggested that agile as a concept, for example in manufacturing, should not be considered as a one-size-fit-all solution [2]. Indeed, variation in software development situational contexts also affects the software process [51–54] and the inevitable change within given contexts presents as a constant challenge to software developers [55]. Though it has been shown that agile delivers an overall improvement in project success, precisely what defines "success" in software companies is subject to variation [45]. Often companies will adopt agile in the hope of it being a general solution, but they may not adapt their organisations to allow agile to work for them [3]. The goal of this research is to investigate whether agile methodologies incur certain hidden costs that may not be accounted for (or calculated). Agile projects should be implemented with a robust level of understanding and discipline, without which they run the risk of losing time, reducing quality and creating confusion for the development team [4]. Furthermore, the way in which agile affects the culture and traditions of a company is difficult to measure, but research has shown that this has been one of the most overlooked and difficult problems to correct [5]. The agile approach requires an excellent working understanding of its processes by everyone involved, including the customers. We suggest that companies that view this development process simply as a quick way to fix production or delivery problems face hidden dangers that could result in compounding these issues. To examine this important space, we investigate what it means for agile projects to be successful, and we seek to identify aspects of agile that can be troublesome and that companies may inadvertently overlook.

Section 2 of this paper details the related agile software development literature, with Sect. 3 examining the concept of success in agile software development, and Sect. 4 investigating the challenges associated with agile. Section 5 presents research limitations and Sect. 6 concludes the paper and identified possible future research directions.

2 Related Literature

Four researchers were responsible for conducting the review over a 7 week period as part of an undergraduate assignment based Dublin city University. A multivocal literature review (MLR) [46] was employed, and it involved the use of various white (such as academic papers) and grey (such as blogs, newspapers, websites) literature. Careful consideration was taken when choosing grey literature and it was almost always used in combination with white literature to corroborate a point. The team took part in weekly meetings where any problems and questions were discussed with the module lecturer.

2.1 Research Questions

The data sources and search strategies that were thoroughly examined determined the research questions in this paper. The aim of this paper is to provide an answer to the two following research questions:

- What is "success" for agile software projects?
- What are the factors that organisations may overlook when using or adopting agile?

These two research questions were chosen as we believe that they effectively target the goal of the paper. From these research questions, the team aimed to provide information on topics such as rework, agile adoption, turnover and maintainability.

2.2 Data Sources and Search Strategy

In the first meeting, key search strings were discussed, and the literature review was then broken down into several smaller steps. First was to determine the key strings which were used in determining relevant academic papers and grey literature. Key strings such as "Agile success", "agile rework costs", "questioning agile", "rework agile" and "failure in agile" were used and careful examination of the different literature was performed before the relevant material was extracted. Google search was utilised to find grey literature. The academic papers were found through Google Scholar and the digital libraries of publications include ScienceDirect, IEEE and ACM which were used to find the academic papers used in this paper.

2.3 Inclusion/Exclusion Criteria

Prior to conducting the bulk of the research, a class was held on how to correctly carry out an MLR. Criteria for including and excluding certain literature was discussed. The title and abstract of the returned papers were briefly analysed to determine relevancy. Papers were considered relevant if they provided an interesting viewpoint on one of the topics mentioned above as well as being available in English. For grey literature, further criteria needed to be accounted for such as the validity of the source or author. Using this process, just over 40 pieces of literature were deemed relevant. The result of our findings are discussed in the following sections.

3 Agile Success

Traditional software development is plan-driven in which work begins by documenting a complete set of requirements, followed by high-level design documents before the coding and testing has begun. The emphasis on documenting comprehensive set of requirements and a design up front have while beneficial to certain aspects of software projects, they can also manifest as "a source of major contention, rework, and delay at high-change levels" [6]. It is these challenges that caused agile methods to become so popular [7], possessing as they do the ability to enable organisations to be flexible in their treatment of requirements and focused on enabling businesses to rapidly respond to changes [8]. There are also claims that agile methods achieve success as the primary focus is delivering business value while reducing costs [9], and while some costs are certainly eliminated (such as documentation costs), the authors of this work suggest that perhaps there has been insufficient treatment of other less obvious costs that are not presently subject to measurement. Prior to agile, reported project success rates were far lower than today, for example the 1994 Standish Group Chaos report, stated that software project success rates were only 16.2%. However, from this it should not be concluded that agile software development has improved project success rates, as many other technological innovations have occurred in the intervening years, and indeed, measures of project success are also subject to change [46].

3.1 Defining Success in Agile

In order to understand why agile has become so successful, it is important to define success. Defining project success has been a topic of discussion for many researchers and practitioners. Often, the definition of success comes back to the triple constraint [7]. Using this approach, a project is considered successful when it is delivered on time, within budget and on target or according to requirements. This view of success may be seen as project management success. However, this may be an over-simplified view of success, and thus it is adapted to include other perspectives, for example customer satisfaction [10]. The 2015 Standish Group Chaos report defines success as being on time, within budget and producing a satisfactory result, marking a project that passes all three as successful, challenged if one of the three measures fails and lastly, failure if the project was cancelled [11] (the report gathered data from over 10,000 projects). The report examines difference in success between agile and waterfall methods, finding that in general agile is more successful (ref. Fig. 1). However, the difference is most pronounced in medium to large size projects.

SIZE	METHOD	SUCCESSFUL	CHALLENGED	FAILED
All Size Projects	Agile	39%	52%	9%
	Waterfall	11%	60%	29%
Large Size Projects	Agile	18%	59%	23%
	Waterfall	3%	55%	42%
Medium Size Projects	Agile	27%	62%	11%
	Waterfall	7%	68%	25%
Small Size Projects	Agile	58%	38%	4%
	Waterfall	44%	45%	11%

Fig. 1. Chaos report 2015: success in agile vs waterfall [11]

While the measurements above for defining success are important, they do not take account of certain qualities of agile such as technical excellence, process improvement and sustainable development. The Art Of Agile Development outlines that success should be considered as a union of organisational success, technical success and personal success (ref. Fig. 2). The book highlights that all three are necessary. Personal success is required to ensure employees remain motivated, technical success to ensure the software created is maintainable and organisational success to ensure that the project is delivering value [9].

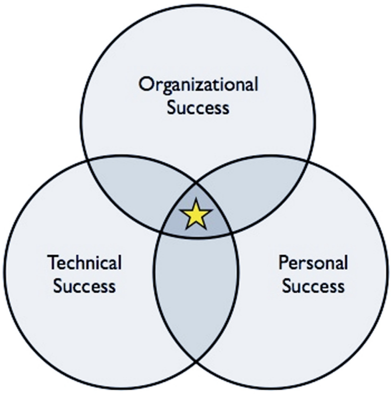


Fig. 2. Success in agile software development [9]

Agile, however, is not a silver bullet for software project success [9], with multiple grey literature sources discussing many underlying problems that have appeared in agile recently [12–14]. Kent McDonald discusses how organisations are moving to agile as they believe they will be at a disadvantage if not seen to do so. Ron Jefferies outlines that when agile is adopted poorly, developers can suffer with having less time to do the work, with increasing demand to build software faster. This can result in more defects, slower progress and developers leaving an organisation. A survey examining critical agile success factors found the following to be important: a correct delivery strategy, proper practice of agile software engineering techniques, a highly skilled team, a solid team management process, a team oriented environment and strong customer involvement [15]. Customer involvement may be particularly important, and that without it, agile projects may experience pressure to over-commit, loss of productivity, significant rework and business loss [16]. If we consider that customer collaboration, sometimes extending to the on-site presence of a customer, is highly encouraged in agile software development, we be alert to the impact of reduced customer collaboration in agile settings. Afterall, perhaps not all clients will be able to locate on-site – full time or even part time - during projects.

The 2018 State of Agile report notes that 97% of the respondents utilise agile software development methodologies, with Scrum and SAFe being the most popular agile implementations [1]. The report highlights the reasons for adopting agile methods as accelerating software delivery, managing changing priorities and increasing productivity. It is important to note that the report also mentions that 83% of the respondents were below a high level of competency with agile practices, which we suggest is cause for concern in relation to long term software project and system sustainability. A 2017 study, “An Agile Agenda”, examined 300 companies across the UK and US, finding that agile “has been adopted so enthusiastically that it is now being stretched beyond its limits” [19]. Of course, the rise of frameworks such as SAFe may work to reduce this impact, but they might be more heavyweight agile frameworks, requiring as they do an increases in process and not necessarily at a rapid delivery cadence. In this sense, scaled agile frameworks such as SAFe might, we suggest, could be classified as *semi-agile* frameworks, falling as they do between traditional lifecycle models such as the waterfall, and more recent agile themed approaches such as continuous software engineering [48].

4 Potential Challenges with Agile

In this section, we aim to outline potential challenges that may be overlooked when using agile methods.

4.1 Minimal Documentation

One of the primary values of the Agile manifesto is “working software over comprehensive documentation”, which encourages developers to focus more on delivering software rather than spending time on documentation. Often, the code is considered the documentation and as change is guaranteed in agile, spending time on documentation can be considered effort wasted [17]. Traditional methods such as Waterfall made extensive use of up front design and documentation. Agile, in contrast, aims to reduce the large up front design cost as this is considered one source of the high rework costs that may associated with traditional approaches [6]. Note that it is not just the initial creation of the big upfront design that is considered expensive, it is also the fact that the design must be revisited and modified as changes arise. Agile does not completely disregard documentation, it only emphasises that it is more important to apply knowledge rather than documenting it [18].

However, there are genuine concerns regarding the absence of appropriate levels of documentation on agile projects, especially as it has been observed that “44% of Agile projects that fail do so because of a lack of documentation” [19]. Due to agile methodologies focus on minimal documentation, it means that the primary source of knowledge within the methodology is tacit. This approach works in favour of agile methods if the team’s tacit knowledge is sufficient for the project life cycle, however, there is also a risk that the team may make critical mistakes due to the shortfalls in tacit knowledge [6]. In small agile teams or organisations, tacit knowledge may suffice, but in larger agile organisations, formal documentation is required for inter-team communication and coordination [20].

Another interesting point refers to one of the 12 principles of agile: Agile processes promote sustainable development. With relation to sustaining quality performance in software development, tacit knowledge transformation to organisational knowledge is essential. However, due to agile's fast paced environment, it can be challenging to record tacit knowledge in documentation [18]. Knowledge sharing techniques have been used in agile approaches such as pair programming, pair rotation as well scrum daily stand-up meetings which include the entire team [21]. But the hidden value of these practices in enabling a cohesive overall agile strategy may not be appreciated across the board, for example some may consider pair programming too expensive, even though Kent Beck in prescribing XP [49] advocated that all practices must be implemented, and not just some a la carte selection.

Rudimentary and partial measures of success may only measure up until the project is handed over to the maintenance stage, and therefore, further studies have examined maintainability in agile environments. One study, involving 18 organisations, observed that agile may be effective in the short term as there is less focus on documentation and more on productivity, but long term there may be problems [22]. For example, the study results showed that 50% of software engineers lost track of projects they are working on, 66% also highlighting that the loss of key engineers would be a major issue, leading to increased costs and decreased productivity. This is a good example of the challenges raised by absences of formal knowledge reification (for example in the form of documentation describing the product), whereby developers that wishes to extend the design may be puzzled by the current implementation [17].

4.2 Rework

Rework can be described as redoing a process or activity that was initially implemented incorrectly, often due to changing requirements, and it can directly impact the performance and profits of an organisation, perhaps depleting 40% to 70% of a project's budget [23]. We suggest that there are genuine challenges regarding the classification of rework, especially concerning avoidable and unavoidable rework. However, literature shows that some rework is unavoidable in software development as software is an evolving process. Multiple research efforts have described the different types of rework in software development, which may be broken down into three categories [24]:

- Evolutionary: Usually rework that occurs in response to a change in user requirements (unavoidable).
- Retrospective: Rework that occurs when developers knowingly exclude needs required by users. Thus, requiring the developer to add the needs in the next version (avoidable).
- Corrective: Often the most common type of rework. It involves fixing defects that were added in previous versions of the code (sometimes avoidable).

It is important that organisations identify the root causes of the different types of rework, as it can help with improving productivity, developer morale and customer satisfaction [24]. Reasons for rework vary within different organisations, and we suggest that some interpretation may be required in classifying rework, and not all participants may agree on the classification reached for a given item. Common reasons for rework are often due to poor communication, ambiguous requirements, inadequate testing and a lack of documentation [6]. The main causes can also be represented using the "fish-bone" model as seen in Fig. 3 [23].

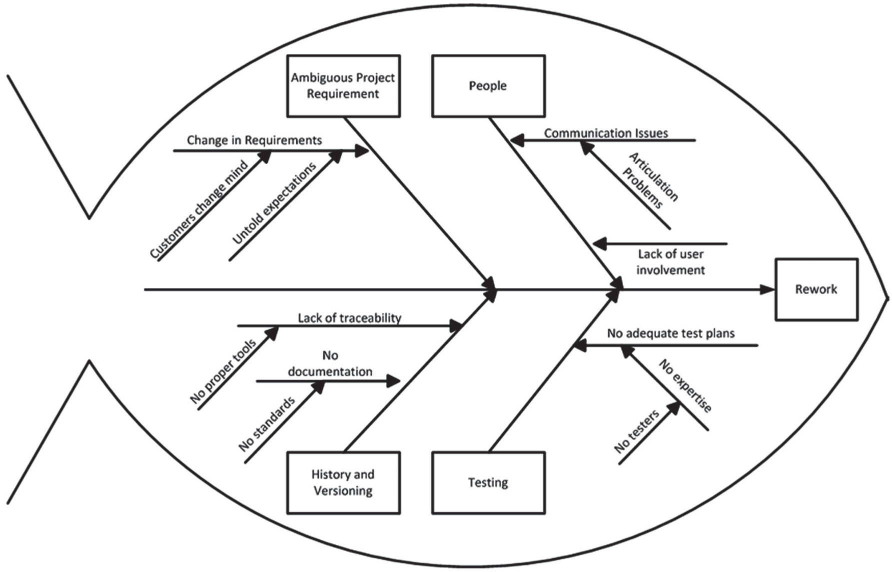


Fig. 3. Fish-bone model [23]

While some consider refactoring to be considered a form of rework [24], rework is considered by many as a negative process, whereas refactoring is positive [7]. Refactoring helps to improve code understanding, improve maintainability, easier to test and allows for easier defect detection [5]. Of course, with agile based development, rework is inevitable [24] as the fundamental principle is that of welcoming change at any time during the development lifecycle [25]. Agile principles encourage better collaboration with customers, technical excellence and releasing working software frequently, all of which target the causes of rework. Owing to the increased collaboration between users and developers in agile, the amount of rework can be reduced. A study which examined the socio-technical aspects of agile, showed rework was reduced because of frequent communication between team members [26]. Agile also promotes the use of an automated build and test system which alerts developers of errors early in the development lifecycle when new features are implemented [24]. A study which examined the effectiveness of agile methodologies versus a waterfall methodology found that the total amount of rework done was less in the agile approach (ref. Fig. 4) [27]. Our analysis suggests that most rework research focuses on minimizing the amount of rework in a given software project [23], and for an aspect of cost volatility as large as rework can potentially be, there was little concrete or substantial contemporary published material dedicated to measuring the cost of rework in agile settings.

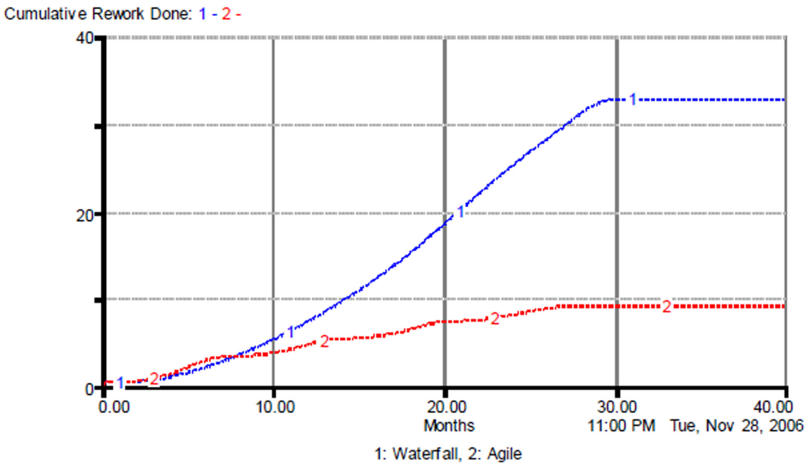


Fig. 4. Rework in agile versus waterfall [27]

It is accepted that 10% to 20% of total effort in agile is spent on rework [24], yet we are also told that it accounts for 40% to 70% of a software project's budget [23]. All types of rework should be collected and analysed if organisations wish to find the root causes. However, software developers may avoid noting rework as it can take away from perceptions of productivity and their organisations may tie rework to performance reviews. Organisations should encourage developers to speak about rework without the worry of it affecting performance reviews [24], perhaps rework reduction should be a performance parameter. It is also interesting to note the relationship of rework with the product backlog which is often used in agile methodologies such as scrum. One study focused on how product backlog changes are managed in Scrum projects [28]. In this study, it states that as a requirement changes, over 50% of practitioners write the updated requirement as a new requirement, instead of recording the change reason or motivation. This may lead to significant untracked rework and it appears that there is ripe potential for both research and measurement in this space.

4.3 Organisational Change

All software development is complex [49] and we can therefore expect that it requires detailed planning and analysis. It is therefore not surprising to discover that up to 34% of agile projects that fail are reported to do so because of a lack of planning for the project and its methods [29]. We therefore see that while agile may advocate more so-called *lightweight* developer-oriented planning, monitoring and control, this alternative to traditional project management-led planning does not necessarily yield improved project success. In the UK, over £30 billion is estimated to have been wasted on unsuccessful agile projects in 2017 [19]. These are the costs that companies hope to reduce when adopting agile but which may manifest due to its adoption.

One example of the process change that may be required relates to software architecture, with 68% of CIOs interviewed in one survey claiming that agile teams require more architects, with hiring and training of architects seen as necessary if a company wishes to scale-up using agile methods [19]. This can be considered to be a very interesting observation, as software architects have a broader view of the software system and its overall design, and an insistence of having higher numbers of architects in agile environments may reduce the impact of design decay. Perhaps, indeed, agile settings prefer or benefit from increased architectural knowledge among heretofore standard software engineers. Agile architects are expected to mentor and be directly involved with the development process and future planning as requirements change [30].

When transitioning to agile, changes in the organization's culture, structure and management can be expected. Agile is highly dependent on social interaction between team members, and a shift in organizational structure can impact the work culture of an organization. Culture has a powerful yet intangible effect on the development process, and the power shift from management to the developers which can cause discontent and disruption for a project and its members [31]. In the 9th Annual State of Agile Survey, 42% of respondents claimed one of the leading reasons for a failed agile project was company philosophy or culture at odds with core agile values, and 36% stated a lack of support for cultural transition [1]. Failed projects are a continuing problem in the IT industry, 66% of projects partially or totally fail [11], with the worldwide estimated cost of IT project failure reported to be 6.18 Trillion USD [32]. If a company is unable or unwilling to account for the cultural change that agile requires it will cost the company time, money and workers that become dissatisfied or unable to work in a conflicted system. Agile requires a shift from a "command-and-control" process to a "leadership-and-collaboration" process [31], and this might not be easily achieved in various settings.

Scrum is an agile project management approach that emphasises the interaction and cooperation of many different roles on a team, with the welcomed involvement of the client to the process. However, the various meetings required in agile, including those with client involvement and potentially other stakeholders also, can raise challenges related to differing priorities [33]. Face to face communication is regarded as the most effective form of communication and is encouraged as the primary form of communication [25]. This is typically facilitated by the daily stand-up meetings, however, as software development has become increasingly globalized in recent years, organisations have had to figure out how to accommodate teams with members that are living in different time zones. Success rates for projects with members that are located far apart drops significantly when compared to those that are co-located [34]. Large time zone differences, national and religious holidays, language barriers all can act as an impediment for agile practices [35]. If these practical realities are not properly accommodated, organisations could experience additional costs in terms of time, money and quality. The collaborative nature of agile methods means that it is at increased risk [5].

4.4 Job Satisfaction

Some research suggests that most of what we know about job satisfaction in agile is anecdotal and to be careful about the findings [36]. Job dissatisfaction can have a huge impact on staff turnover rates and stress, and high turnover rates are concerning as they can have a considerable economic effect on an organisation and its employees [36]. Where turnover is an issue, and the software development business is not immune to this phenomenon, training new replacement employees to an efficient standard is important. However training new employees can be costly especially in small software development companies [18].

Many different factors may affect job satisfaction in agile environments, with certain aspects shown to have a positive effect, including agile project management methods (such as daily stand-ups, retrospectives, iterative planning) and development practices (such as automated unit testing, pair programming, continuous integration and refactoring) [37]. Studies have proven that higher perceptions of job characteristics of a) feedback, b) task significance, c) skill variety, d) job autonomy, e) ability to complete a whole task are positively related to job satisfaction [36, 37, 40]. What can infer from these works that factors such as job autonomy, skill variety, task significance, feedback and the ability to complete a whole task are significant factors to support developer satisfaction [36, 37].

Comparing the different levels of stress that an employee may be experiencing depending on the methodology used may provide useful insights into how agile affects stress. A study performed by Mannaro et al., studied factors affecting satisfaction in software teams and specifically examined the relationship between stress and the software methodology adoption, finding that agile methods were associated with reduced stress levels [38, 39]. A similar study conducted in Switzerland performed an online survey querying managers and developers about the usage of development methods and practices such as agile, and examined possible associations with stress (ref. Fig. 5) [39]. In this study, managers and developers were asked to rate how agile had influenced their stress at work. They responded on a scale from 1 (not very stressed) to 5 (very stressed), and findings suggest that managers seem to be less stressed in agile environments compared to the developers who tend towards a neutral observed stress effect. This is perhaps not surprising as managers have less direct responsibility.

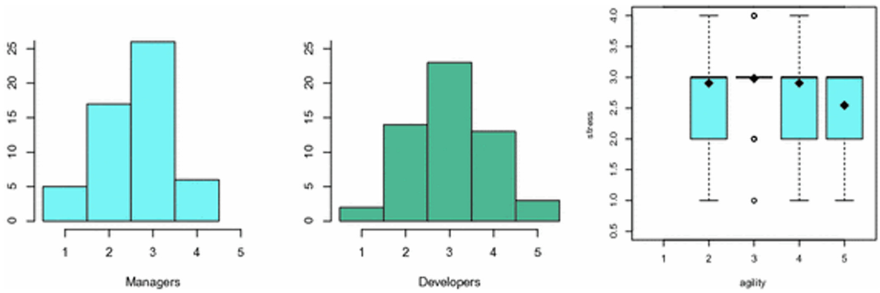


Fig. 5. Levels of stress in agile, managers vs developers [39]

4.5 Knowledge and Employee Turnover in Agile

Knowledge loss is a serious issue for any development team, the smaller the team, the greater the impact of the staff turnover rates [18]. This impact can manifest in many ways, for example on team productivity and quality. There are many different forms of knowledge loss that can happen within a company, for example, knowledge hoarding, staff turnover and knowledge loss due to the pace of technology. Knowledge hoarding is where employees want to keep their knowledge hidden for inter-personal reasons or where it may be endemic in the organisational culture, however building a knowledge sharing environment is very important for any business [18].

The pace at which technology changes is a good example of the challenge for knowledge sharing in software development companies. Developers may find it difficult to update their knowledge while trying to meet all their deadlines and because of that they will not get around to sharing knowledge with their colleagues [18]. Due to the agile methodologies emphasis on minimal documentation, employee turnover rates are perhaps especially concerning to a company [18, 25, 42]. If the original development team cannot be maintained, knowledge about the software will be lost and extra resources will go into training new employees [18, 25]. These costs might not be insignificant the knowledge intensive software development ecosystem where employee turnover has been a perennial challenge.

The opportunity to work on interesting projects, being able to influence decisions and the relationship with their users are statistical factors in satisfaction [36]. One study found that job dissatisfaction can lead to high voluntary staff turnover [36], with turnover costing about 20 percent of all manpower expenses [41]; other studies have estimated turnover costs to be as much as 70–200% [36]. Software development firms will understandably want to avoid high turnover rates and reports agile positively affects employee satisfaction, might be aligned with improved strategic outcomes including staff retention [36, 38]. It should be noted that calculating the true cost of replacing an experienced worked is a non-trivial exercise in its own right [41].

Agile adapts to employee turnover rates by introducing two knowledge sharing approaches, which create strong enthusiasm in software engineering [18]. Pair programming is the first one and it allows knowledge sharing between two developers. Pair rotation is the second knowledge sharing technique and it is used to break the ice between software development team members, however, the downsides to pair programming and pair rotation are that they double the manpower and cost on a given problem [18]. So while there are mechanisms in place to support knowledge sharing in agile settings, they are not free and perhaps not consistently adopted in industry, and the very concept of knowledge retention in software engineering is a broad and complex field [50].

5 Limitations of Research

The field of agile software development is vast, it is employed around the world and is documented extensively across both academic and grey literature. To investigate the true costs of agile software development is to some extent an impossible task, especially given the extent of situational variation that can arise in software development [44]. In an effort to better understand the costs associated with agile software development, four primary researchers investigated different costs themes in agile software development over a seven week period. While this was sufficient to gain some insight into the phenomena of interest, it can be considered quite limited in terms of a full evaluation of all of the material that has been documented on these topics.

Furthermore, the four primary researchers were four final year undergraduate students in the B.Sc. in Computing Applications (Hons.) based in Dublin City University, Ireland. These researchers are not formally trained in research, and it was their first opportunity to conduct academic research. To offset this, the researchers received training on performing multivocal literature reviews, and were furthermore supported weekly on the process of performing robust academic research.

6 Conclusion and Future Work

Awareness of the strengths and successes of agile are spoken about at large in the software industry. When implemented into an organisation's process effectively, the benefits to customer and employee satisfaction can be substantial; with product improvement and cost reductions reported across both academic and grey literature. However, the ease at which agile can be effectively adopted can be overestimated, and the concept of "success" may be overstated and over simplified, giving a false impression that agile is the standard, simple fix to any project regardless of the business context. Agile is a way of thinking and a culture that must be clearly explained to all those who are involved and creating awareness of this point is massively important. Without mature familiarity with the agile process and all its supporting practices, both internal and external disruption can arise which can directly or indirectly affect the product and its clients. Those responsible for strategic decision making in software development firms should invest in understanding agile more deeply, the ultimate realisation from which may be that agile is more complex than some simple scaffolding provided in some agile method. It is in fact a cocktail of interrelated practices and techniques that drive an organisational culture and which can prove difficult to master.

Clear and direct communication are essential to an effective agile strategy, and issues with scaling and distantly located team members can be exacerbated if they are not carefully accommodated in the agile process. Often, management can look to agile to fix general issues with a project at the time of adoption, but as an organization has more success, they risk scaling without properly preparing their process.

There are a significant number of intangibles that are difficult to assess when it comes to examining project success in agile. With work culture playing a large role in the level of an employee's job satisfaction, the need for a culture shift can cause dissatisfaction for workers. Our research found that the agile paradigm shifts some of the power from management and grants it to the developers, and that this can raise developer job satisfaction. Although agile may increase perceived work pressure among developers, overall their engagement with their role can increase and stress levels can be reduced. Somehow, the increased pressure that comes with increased empowerment may in fact serve to reduce overall stress load. But our findings here are not conclusive and given the role of job satisfaction in employee retention, we recommend further attention to understanding the reasons for high staff attrition and if these may be related to agile adoption. Other reasons for this attrition clearly exist, for example the high demand on skilled software developers can drive wage inflation. But not every employee leaves a company for financial reasons, and with knowledge workers in particular, job satisfaction may be a key factor to consider. And a key concern in job satisfaction relates to the work processes, and therefore, this should be examined further through the lens of agile software development.

In relation to rework, our findings suggest that it is a complicated concept. Put simply, what one person considers to be rework, another person may classify as requirements discovery through an evolutionary feedback process. This is perhaps the most important finding from our research: it seems that there is very limited research into the costs of rework in agile settings. We also observe what might be considered a dangerous practice of introducing wholly new product backlog features when eliciting user feedback on already-implemented features. Left unchecked, this practice can disguise rework as new product backlog items, and it could be the source of cost hemorrhage in some settings. It may be that effective agile implementation should insist on linking new product backlog items to earlier backlog items to police against runaway rework costs. But as much of this appears unmeasured – or at least unreported – today, we cannot know the true position. This, we suggest, should be a major concern for software companies.

While our research brought forward no dedicated studies on the impact of minimal documentation in agile, multiple studies have suggested that a lack of documentation may cause issues. Agile may be beneficial for short term goals, but if ineffectively implemented, it may exacerbate long term maintainability costs. This point cannot be overemphasised and it plays back into the long-established guidance not to take an a la carte approach to agile software development practices. Supporting practices such as refactoring and pair programming are important ingredients in building a long-term sustainable product and team. Without an emphasis on software design as supported through refactoring, products can become very expensive to maintain, and without practices for training and knowledge sharing such as pair programming, companies can silently accumulate dependencies on transient workers, all the while perhaps also working against a truly agile philosophy. And finally, let us not forget that while the Agile Manifesto values working software over comprehensive documentation, and individuals and interactions over processes and tools, it also recognises that there is *value* in documentation and processes. This is perhaps the small print that some businesses miss.

Acknowledgements. This work was supported, in part, by Science Foundation Ireland grant 13/RC/2094 and co-funded under the European Regional Development Fund through the Southern & Eastern Regional Operational Programme to Lero - the Irish Software Research Centre (www.lero.ie).

References

1. Annual State of Agile Survey. www.stateofagile.com/. Accessed 10 June 2020
2. Shewchuk, J.P.: Agile manufacturing: one size does not fit all. In: Bititci, U.S., Carrie, A.S. (eds.) Strategic Management of the Manufacturing Value Chain. ITIFIP, vol. 2, pp. 143–150. Springer, Boston (1998). https://doi.org/10.1007/978-0-387-35321-0_16
3. Agile budgeting: How much will it cost? Agilest.org. Accessed 10 June 2020
4. Stoica, M., Marinela, M., Bogdan, G-M.: Software development: agile vs. traditional. Inf. Econ. **17**(4) (2013)
5. Korkala, M., Maurer, F.: Waste identification as the means for improving communication in globally distributed agile software development. J. Syst. Softw. **95**, 122–140 (2014)
6. Boehm, B.: Get ready for agile methods, with care. Computer **35**(1), 64–69 (2002)
7. Serrador, P., Pinto, J.K.: Does Agile work?—A quantitative analysis of agile project success. Int. J. Proj. Manage. **33**(5), 1040–1051 (2015)
8. Highsmith, J.: Agile Software Development Ecosystems. Addison Wesley, Boston (2002)
9. Shore, J.: The Art of Agile Development: Pragmatic Guide to Agile Software Development. O'Reilly Media, Inc., Newton (2007)
10. Van Der Westhuizen, D., Fitzgerald, E.P.: Defining and measuring project success. In: Proceedings of the European Conference on IS Management, Leadership and Governance 2005. Academic Conferences Limited (2005)
11. Standish Group. 2015 Chaos Report
12. McDonald, K.: Agile Q&A: Why do Organizations Adopt Agile? www.agilealliance.org/why-do-organizations-adopt-agile/. Accessed 10 June 2020
13. Jefferies, R.: Developers should abandon agile. https://ronjeffries.com/articles/018–01ff/abandon-1/. Accessed 10 June 2020
14. Nicolette, D.: Questioning Agile Dogma. www.leadingagile.com/2019/02/questioning-agile-dogma/. Accessed 10 June 2020
15. Chow, T., Cao, D.-B.: A survey study of critical success factors in agile software projects. J. Syst. Softw. **81**(6), 961–971 (2008)
16. Hoda, R., Noble, J., Marshall, S.: The impact of inadequate customer collaboration on self-organizing agile teams. Inf. Softw. Technol. **53**(5), 521–534 (2011)
17. Knippers, D.: Agile software development and maintainability. In: 15th Twente Student Conference (2011)
18. Ersoy, I.B., Mahdy, A.M.: Agile knowledge sharing. Int. J. Softw. Eng. (IJSE) **6**(1), 1–15 (2015)
19. An Agile Agenda, 6Point6 Technology Services, April 2017. https://6point6.co.uk/insights/an-agile-agenda/. Accessed 10 June 2020
20. Dikert, K., Paasivaara, M., Lassenius, C.: Challenges and success factors for large-scale agile transformations: a systematic literature review. J. Syst. Softw. **119**, 87–108 (2016)
21. Chau, T., Maurer, F.: Knowledge sharing in agile software teams. In: Lenski, W. (ed.) Logic versus Approximation. LNCS, vol. 3075, pp. 173–183. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-25967-1_12
22. Kajko-Mattsson, M.: Problems in agile trenches. In: Proceedings of the Second ACM-IEEE International Symposium on Empirical Software Engineering and Measurement (2008)
23. Ramdoo, V., Huzooree, G.: Strategies to reduce rework in software development on an organisation in mauritius. Int. J. Softw. Eng. Appl. **6**(5), 9–20 (2015)
24. Fairley, R.E., Willshire, M.J.: Iterative rework: the good, the bad, and the ugly. Computer **38**(9), 34–41 (2005)

25. The Agile Manifesto, Agile Alliance (2001). https://agilemanifesto.org/. Accessed 10 June 2020
26. Inayat, I., Marczak, S., Salim, S.S.: Studying relevant socio-technical aspects of requirements-driven collaboration in agile teams. In: 2013 3rd International Workshop on Empirical Requirements Engineering (EmpiRE). IEEE (2013)
27. Chichakly, K.: Modeling agile development: when is it effective? In: Proceedings of International Conference of the System Dynamics Society (2007)
28. Alsalemi, A.M., Yeoh, E.-T.: A survey on product backlog change management and requirement traceability in agile (Scrum). In: 2015 9th Malaysian Software Engineering Conference (MySEC). IEEE (2015)
29. Saran, C: Agile development, an 'IT fad' that risks iterative failure, May 2017. https://www.computerweekly.com/news/450418205/Agile-development-an-IT-fad-that-risks-iterative-failure. Accessed 10 June 2020
30. Johnston, A.: The role of the agile architect. https://www.agilearchitect.org//agile/role.htm. Accessed 10 June 2020
31. Nerur, S., Mahapatra, R., Mangalaraj, G.: Challenges of migrating to agile methodologies. Commun. ACM **48**(5), 72–78 (2005)
32. Krigsman, M.: Worldwide cost of IT failure, December 2009
33. Coram, M., Bohner, S.: The impact of agile methods on software project management. In: 12th IEEE International Conference and Workshops on the Engineering of Computer-Based Systems (ECBS 2005). IEEE (2005)
34. Geographically Distributed Agile Teams, PMI Disciplined Agile. https://www.pmi.org/disciplined-agile/agility-at-scale/tactical-agility-at-scale/geographically-distributed-agile-teams. Accessed 10 June 2020
35. Kajko-Mattsson, M., Azizyan, G., Magarian, M.K.: Classes of distributed agile development problems. In: 2010 Agile Conference. IEEE (2010)
36. Melnik, G., Maurer, F.: Comparative analysis of job satisfaction in agile and non-agile software development teams. In: Abrahamsson, P., Marchesi, M., Succi, G. (eds.) XP 2006. LNCS, vol. 4044, pp. 32–42. Springer, Heidelberg (2006). https://doi.org/10.1007/11774129_4
37. Tripp, J.F., Riemenschneider, C., Thatcher, J.B.: Job satisfaction in agile development teams: agile development as work redesign. J. Assoc. Inf. Syst. **17**(4), 267 (2016)
38. Meier, A., Kropp, M., Anslow, C., Biddle, R.: Stress in agile software development: practices and outcomes. In: Garbajosa, J., Wang, X., Aguiar, A. (eds.) XP 2018. LNBIP, vol. 314, pp. 259–266. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-91602-6_18
39. Mannaro, K., Melis, M., Marchesi, M.: Empirical analysis on the satisfaction of IT employees comparing XP practices with other software development methodologies. In: Eckstein, J., Baumeister, H. (eds.) XP 2004. LNCS, vol. 3092, pp. 166–174. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-24853-8_19
40. Tessem, B., Maurer, F.: Job satisfaction and motivation in a large agile team. In: Concas, G., Damiani, E., Scotto, M., Succi, G. (eds.) XP 2007. LNCS, vol. 4536, pp. 54–61. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-73101-6_8
41. DeMarco, T., Lister, T.: Peopleware: Productive Projects and Teams. Addison-Wesley, Boston (2013). pp. 17, 118
42. Documentation in agile: how much and when to write it? InfoQ, January 2014. https://www.infoq.com/news/2014/01/documentation-agile-how-much/. Accessed 10 June 2020
43. Clarke, P., O'Connor, R.V., Yilmaz, M.: In search of the origins and enduring impact of agile software development. In: ACM Proceedings of the International Conference of Software and System Processes (ICSSP 2018), Gothenburg, Sweden, 26–27 May 2018, pp. 142–146 (2018)

44. Clarke, P., O'Connor, R.V.: The situational factors that affect the software development process: towards a comprehensive reference framework. Inf. Softw. Technol. **54**(5), 433–447 (2012)
45. Clarke, P., O'Connor, Rory V.: The meaning of success for software SMEs: an holistic scorecard based approach. In: O'Connor, R.V., Pries-Heje, J., Messnarz, R. (eds.) EuroSPI 2011. CCIS, vol. 172, pp. 72–83. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-22206-1_7
46. Garousi, V., Felderer, M., Mäntylä, M.V.: Guidelines for including grey literature and conducting multivocal literature reviews in software engineering. Elsevier J. Inf. Softw. Technol. **106**, 101–121 (2019)
47. O'Connor, R.V., Elger, P., Clarke, P.: Continuous software engineering - a microservices architecture perspective. J. Softw.: Evol. Process **29**(11), 1–12 (2017)
48. Beck, K.: Extreme Programming Explained: Embrace Change. Addison Wesley, Boston (2000)
49. Clarke, P., O'Connor, R.V., Leavy, B.: A complexity theory viewpoint on the software development process and situational context. In: Proceedings of the International Conference on Software and Systems Process (ICSSP), pp. 86–90 (2016)
50. Rashid, M., Clarke, P., O'Connor, R.V.: A systematic examination of knowledge loss in open source software projects. Int. J. Inf. Manag. (IJIM) **46**, 104–123 (2019)
51. O'Connor, R.V., Elger, P., Clarke, P.: Exploring the impact of situational context: a case study of a software development process for a microservices architecture. In: proceedings of the International Conference on Software and Systems Process (ICSSP), Co-Located with the International Conference on Software Engineering (ICSE), pp. 6–10 (2016). https://doi.org/10.1145/2904354.2904368
52. Clarke, P.M., O'Connor, R.V., Solan, D., Elger, P., Yilmaz, M., Ennis, A., Gerrity, M., McGrath, S., Treanor, R.: Exploring software process variation arising from differences in situational context. In: Stolfa, J., Stolfa, S., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2017. CCIS, vol. 748, pp. 29–42. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64218-5_3
53. Giray, G., Yilmaz, M., O'Connor, R.V., Clarke, P.M.: The impact of situational context on software process: a case study of a very small-sized company in the online advertising domain. In: Larrucea, X., Santamaria, I., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2018. CCIS, vol. 896, pp. 28–39. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-97925-0_3
54. Marks, G., O'Connor, R.V., Clarke, P.M.: The impact of situational context on the software development process – a case study of a highly innovative start-up organization. In: Mas, A., Mesquida, A., O'Connor, R.V., Rout, T., Dorling, A. (eds.) SPICE 2017. CCIS, vol. 770, pp. 455–466. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-67383-7_33
55. Clarke, P., O'Connor, R.V.: Changing situational contexts present a constant challenge to software developers. In: O'Connor, R., Umay Akkaya, M., Kemaneci, K., Yilmaz, M., Poth, A., Messnarz, R. (eds.) EuroSPI 2015. CCIS, vol. 543, pp. 100–111. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24647-5_9
56. Clarke, P., O'Connor, R.V., Leavy, B., Yilmaz, M.: Exploring the relationship between software process adaptive capability and organisational performance. IEEE Trans. Softw. Eng. **41**(12), 1169–1183 (2015). https://doi.org/10.1109/tse.2015.2467388

On the Development of a Model to Support the Combined Use of Agile Software Development with User-Centered Design and Lean Startup

Cassiano Moralles[1(✉)], Maximilian Zorzetti[1], Ingrid Signoretti[1], Eliana Pereira[2], Matheus Vaccaro[1], Bruna Prauchner[1], Larissa Salerno[1], Cássio Trindade[1], Sabrina Marczak[1], and Ricardo Bastos[1]

[1] MunDDoS Research Group – PPGCC – School of Technology, Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Porto Alegre, RS, Brazil
{cassiano.mora,maximilian.zorzetti,ingrid.manfrim,matheus.vaccaro, bruna.prauchner,larissa.salerno}@acad.pucrs.br, {cassio.trindade,sabrina.marczak,bastos}@pucrs.br

[2] Instituto Federal do Rio Grande do Sul (IFRS), Porto Alegre, RS, Brazil
eliana.pereira@restinga.ifrs.edu.br

Abstract. Despite the benefits of Agile Software Development, organizations still do not have a clear understanding of the problem to be solved by the software product-to-be. This problem can solved by having a closer engagement with users, and the combined use of Agile, User-Centered Design, and Lean Startup has been pointed out as a strategy to achieve this. However, this novel development approach does not have publicly available tools or instruments that aid organizations in using it in any way. We describe the first steps in building a literature and practice-informed acceleration model to aid in the adaptation to this combined approach. The model is based on a set of principles, activities, and practices derived from Agile, User-Centered Design, and Lean Startup pillars. These will serve as guidance for software teams to determine the competences and skills they need, by accounting for forces that might influence their work processes. We discuss how the acceleration model could help teams adapt to the combined approach and point out our future research plan to define the next steps to fully realize our model.

Keywords: Agile software development · Lean Startup · User-Centered Design · Maturity model · Process transformation

1 Introduction

Agile arose as an alternative to traditional software development methods and is widely adopted by companies today [7]. However, there are some limitations in a pure agile approach, namely the lack of user involvement [22] and clear

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 220–231, 2020.
https://doi.org/10.1007/978-3-030-56441-4_16

identification of added value [11], which indicate that agile may need to be combined with other approaches.

A combined use of Agile Software Development, User-Centered Design (UCD), and Lean Startup can be a great way to tackle the aforementioned limitations: UCD [1] puts the user at the center of the discussion to incentivize creativity and empathy; while Lean Startup [18] focuses on adding value to business stakeholders by searching for the best solution through experimentation, leading to a reduced waste of resources, time, and financial investments.

However, an organizational transformation to such an approach has several challenges regarding structural, technical, and cultural aspects [14]. These challenges are made more difficult when dealing with a large-scale development setting as new issues arise (e.g., inter-team coordination [14]). Aside from these challenges, there is a lack of understanding on how to properly integrate these three pillars, and of how teams should adapt themselves skill-wise in order to use this combined approach—an issue that can be resolved, for instance, with the use of maturity models, which aspire to gauge the transformation in a not overly expensive and time-consuming manner [12]. Motivated by this, we aim to propose an acceleration model to help teams adapt to this combined approach.

This paper reports on the development of the model through a mixed-methods approach. We looked at previous literature to build a conceptual foundation of the three pillars and conducted a case study to experimentally examine the adoption of the combined approach by a large-scale company. We used both to fuel iterative design sessions to create an initial version of our model, which aims to support teams towards the full adoption of the combined approach by taking into account any team's context.

2 Background

2.1 The Combined Use of Agile Development, User-Centered Design, and Lean Startup

In the past two decades, agile methods have taken over in popularity over more traditional methods for software development [7]. A crucial agile advantage is the ease of interaction and collaboration among team members and customers, which is supported by a range of ceremonies that promote such interactions, such as daily meetings, stakeholders meetings, and so on [8]. Still, despite the incentives to team-user collaboration, there is still a lack of user involvement [22] and difficulty in addressing business value into the product. Vilkki [30] argues that agile needs to be combined with other approaches so agile teams can improve their understanding of the problem at hand, thus keeping the customer engaged and leading to the development of better aligned solutions.

Given this need to complement agile with other methodologies, some studies have reported on the combination of Agile, UCD, and Lean Startup (i.e., the "combined approach"). Grossman-Kahn and Rosensweig report on the evolution of the Nordstrom Innovation Lab [6], in which an agile team shifts its focus to the users' needs and starts to use rapid experimentation and prototyping; turning into an acclaimed innovation team that embraces Agile, Design Thinking, and the Nordstrom-inspired software core. Ximenes, Alves, and Araújo present the Nordstrom-inspired software development process model Converge [31], while Dobrigkeit, de Paula, and Uflacker present InnoDev [5]; both describing a list of activities, roles, deliverables, and techniques for each phase of their model. However, they do not detail how to integrate practices from each pillar and how to measure the team's understanding of the rationale behind the combined approach.

2.2 Process Transformation and Maturity Models

Motivations for an agile transformation are distinct. Paasivaara et al. [14] highlight the need to align software development with corporate strategies, the need to rapidly respond to market changes, and the teams' dissatisfaction with the current work process and culture. The transformation process comes with a set of changes (e.g., structural, technical, and cultural changes [14]) that the organizations need to employ. For this reason, the use of a strategy to guide the transformation process (generally either a "big bang" approach, adopting all practices by-the-book; or a "gradual" approach, gradually integrating agile practices into the organization), is most important [8].

Aside from the need of a strategy to guide a transformation, it is crucial that teams have a mechanism, typically a maturity model, that sets up and supports an improvement process during the adoption phase, while also enabling the full benefits of agile practices and techniques to take place. These maturity models are often organized in a 2 by 2 matrix that defines the specific aspects that the model will map (e.g., attributes and dimensions of a formal software development process [15]) and are oftentimes divided in levels or stages [23] while emphasizing key characteristics that result in a successful agile adoption (e.g., people, process, project, and product characteristics [27]).

These models often have the adoption of agile practices as a goal, determining the level of agility of an organization by measuring the use of agile practices [15] and defining specific circumstances for each level [23]. This, however, is using a definition of maturity determined by the model itself, disregarding the actual needs and context of development teams, which makes it harder for them to achieve a specific goal or follow a defined plan.

Despite the studies that describe how to use the combined approach (Sect. 2.1), we did not find any model that helps organizations to accelerate the adoption process and mature in the understanding of such new way of working.

3 Research Method

In order to build our acceleration model, we conducted a mixed-methods research (Fig. 1) to understand what the combined approach consists of and what should be considered when helping an organization adopt it. We started off by building a knowledge foundation based on existing literature (Step 1 – Systematic

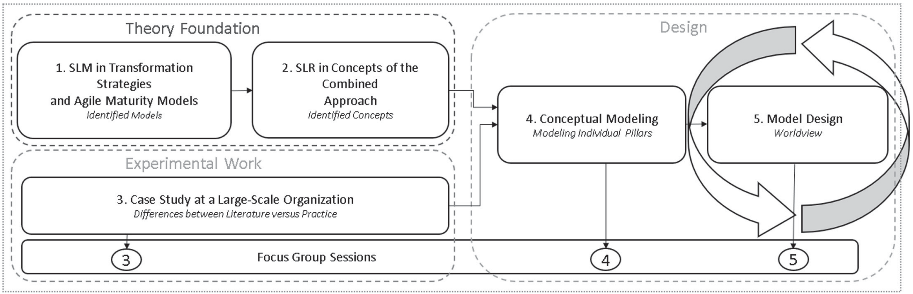


Fig. 1. Diagram of our research method.

Literature Mapping (SLM) on Maturity Models for Agile, Lean Startup, and User-Centered Design in Software Engineering and Step 2 – Systematic Literature Review (SLR) on Concepts of the Combined Approach). In parallel, we sought to identify and understand how a large-scale company undergoes such a transformation (Step 3 – Case Study at a Large-Scale Organization). Next, the results from Step 2 and 3 were expressed in the form of a conceptual model (Step 4 – Conceptual Modeling). Finally, we used an iterative design process to design the model (Step 5 – Model Design), in which we used the conceptual model and some agile maturity models we identified in Step 1 to determine the theoretical basis of our model proposal. During this whole process, we performed several Focus Group Sessions with the participants of our case study and other ORG staff to iteratively validate our findings from Step 3 to 5.

This work is the culmination of a series of previous studies centered on the combined approach. Each step in the research effort of this paper corresponds (or partially corresponds) to an existing study of our research group. As such, in-depth details and results for each step are available in each respective studies.

3.1 Steps 1 and 2. Literature Mapping and Review

We conducted two tertiary studies, a SLM and a SLR. The SLM aimed to identify maturity models for software development approaches comprised of the three pillars, but resulted in no studies that contemplate such combined approach being found. This prompted us to perform a SLR to identify concepts pertaining to the individual pillars of the combined approach, so as to build a knowledge base to assemble our model with. Both studies were conducted following specific guidelines (Petersen's [16] for the SLM and Kitchenham's [9] for the SLR).

SLM on Maturity Models for the Combined Approach. We conducted an electronic search on studies regarding the use and structure of maturity models for a software development approach consisting of the three pillars in the following electronic databases: ACM Digital Library, IEEExplore, Science Direct, Scopus, and Springer Database. The following search string was replicated as accurately as possible[1] and used in each database's search engine:

Agile AND User-Centered Design AND Lean AND Maturity Model

Each keyword (Agile, User-Centered Design, Lean[2], and Maturity Model) stands for the logical disjunction of itself and several of its synonyms, similar wordings, or related terms (e.g., "UCD", "User-Centred Design", or "Human-Computer Interaction" for the User-Centered Design keyword).

At the end of the search process, we did not retrieve any studies that reported on maturity models for such an integrated approach. In order to find mechanisms and inspirations for our model, we broadened our search by including studies that contemplate just one or two of the three pillars, for a total of 6 additional searches (Agile and UCD, Agile and Lean, UCD and Lean, Agile, UCD, and Lean). The details of this study are found in the work of Zorzetti et al. [32].

SLR on the Concepts of the Combined Approach. We conducted an electronic search on studies on Agile, UCD, and Lean Startup concepts in the following electronic databases: ACM Digital Library, IEEExplore, Science Direct, Scopus, and Springer Database. Since our model focuses on software development, we chose Extreme Programming (XP) [2] as our agile pillar, given that it is the most popular [29] agile methodology that directly addresses engineering, as it is dedicated exclusively to software development (as opposed to Scrum, which can be applied in other fields). In short, we performed 3 searches by crafting and applying research strings in the following format:

Subject OR (Subject-related keywords)

In this format, "Subject" stands for the topic of each search (namely, Extreme Programming, Lean Startup, and User-Centered Design) while "Subject-related keywords" stands for a logical disjunction of several synonyms, similar wordings, or related terms to the subject at hand (e.g., "XP" for Extreme Programming and Build-Measure-Learn for Lean Startup).

Our results contained several papers that pointed towards well-known books. We concluded that these books were integral to explore the pillars' concepts sufficiently. Given that books were an exclusion criterion in our search, we decided to remove such criterion and have books included as well.

3.2 Step 3. Case Study

We conducted a multiple case study [20] in a multinational company named ORG (name omitted for confidentiality reasons). ORG has software product

[1] some search engines have restrictions, such as Science Direct's limits on the number of boolean terms.

[2] in addition to Lean Startup, we looked for studies on Lean Software Development; a result of our understanding of the subject matter at the time.

development sites in the USA (headquarters), India, and Brazil. With over 7,000 employees and responsible for about 1,200 software products, its IT department started its agile transformation in 2015 and moved to a combined use of Agile, UCD, and Lean Startup inspired by the Pivotal Labs methodology[3] in late 2017.

We observed *in loco* two teams (team A and team B) from ORG's financial area, both with different product characteristics. The two teams are mostly located in Brazil and comprise a total of 14 participants, 7 for each. Both have the same structure: 4 software engineers, 1 product designer, and 2 product managers. Team A members have on average 13.5 years of experience in software development and 6 at ORG, while Team B members have on average 11 years of experience in software development and 7 at ORG. These teams spent 8 months working in a dedicated software development lab at PUCRS specifically prepared for ORG teams undergoing the transformation.

We used multiple data sources: a questionnaire to collect their profile (name, role, responsibilities, time in years working in IT and at ORG, and whether the person participated in an immersion training about Pivotal Labs in the company headquarters or is being trained by those who received the training), semi-structured interviews to gather information on their perceptions about the transformation to the combined approach, daily observations of team ceremonies, and shadowing of roles for in-depth knowledge. Interviews were recorded and transcribed for analysis and lasted 30 min on average).

To identify the transformation strategies employed by ORG and to understand how the teams adopt the combined approach in terms of activities conducted, techniques used, work products developed, and so on; we followed a content analysis procedure [10] organized into the following steps: organization and pre-analysis, reading and categorization, and recording the results. We first read the dataset, extracted text excerpts and marked them as codes. These codes were revisited and grouped into larger codes, forming categories.

This case was addressed in several studies spanning a multitude of aspects: on the team members' roles [21], on the transformation to the combined approach [24], on the concepts of the combined approach [13], on continuous experimentation [28], and on mindset and operational-level changes [25]. Certain phenomena in the study prompted us to consider contextual factors (e.g., resistance to change) that interfere with development work in our model proposal.

3.3 Step 4. Conceptual Modeling

In order to understand how the concepts of the approach interact, supplement, and influence one another and to establish a common body of knowledge to build the acceleration model with, we converted the results of Steps 2 and 3 into a Conceptual Modeling scheme. The process was as follows:

i We gathered the results from the Concept Review and simplified it into three glossaries (one for each pillar) containing definitions for principles, activities,

[3] https://pivotal.io/Labs.

techniques, and roles to establish a common understanding of the pillars' concepts for our research team;

ii Each glossary was then mapped to a single Unified Modeling Language (UML) class diagram [19] and then developed upon to identify the relationships between the concepts themselves (e.g., the "Build" phase of Lean Startup leads into the "Measure" phase; or a "Test Suite" is comprised of multiple "Test Cases" in XP);

Throughout both steps, we had several meetings take place to discuss how each concept was to relate to one another and if there were any overlapping concepts between each pillar. These meetings were attended by two senior Software Engineering professors, a PhD, a PhD candidate, and two Master candidates. Although not shown in their paper, the diagrams were used to guide the discussion in the work of Moralles et al. [13]. To address some of our needs in the development of our model, we further elaborated upon the diagrams:

iii The class diagrams were converted to use the Software & Systems Process Engineering Metamodel (SPEM)[4] to provide a process-oriented view into our model, helping us in identifying additional relationships between concepts and in grouping similar elements into broad categories (e.g., principles).

3.4 Step 5. Model Design

We gathered all of our findings from Steps 1 to 4 to develop the first draft of our model. We used key ideas from specific maturity models found in Step 1 that were supported by what we observed in Step 3 as inspiration for initial designs, and we grounded our model on the conceptual model of Step 4.

The design of the model came to fruition through an iterative design process which consisted of a series of weekly meetings that encompassed brainstorming activities and in-depth discussion of the data we had collected so far.

3.5 Focus Group Sessions

We conducted focus group sessions from Step 3 onward to validate our findings and collect more data to enrich our results. Initially, the focus group participants were the teams' enablers (see Sect. 3.2): two software engineers, one product manager, and one product designer from each team. On Step 3, we had 3 sessions to validate the case study's data and capture their collective vision of the transformation and adoption process. On Step 4, we had 6 sessions (2 for each pillar) to validate the concepts we identified and to determine which are actually used in practice. On Step 5, we had a session to validate the mechanisms of our model with two senior managers of the ORG transformation initiative and a senior developer leading the transformation of the case study's teams. All sessions were recorded and transcribed for analysis and lasted 1.5 h on average.

[4] A free UML profile/meta-model used to describe software and systems development processes and their components. A specification document for SPEM is available at https://www.omg.org/spec/SPEM/2.0/PDF.

4 An Acceleration Model for the Combined Approach

Our acceleration model aims to provide a supporting tool for companies that want to adopt a software development approach that consists of a combination of Agile, User-Centered Design, and Lean Startup, providing means to scale the adoption of the approach. In this sense, we defined two components that make up our model: an appraisal package and the worldview it is based on.

The appraisal package consists of two elements. The first is a team assessment model that enables the assessment of teams in order to diagnose their stage of maturity and qualification in the various aspects of the combined approach. This assessment could be applied by teams, the company, or other people to identify possible improvement points in a team's skillset and other competencies to accelerate its adoption of the combined approach. The second element is a project evaluation model that enables its user to identify the required expertise to develop the deliverables of a given project successfully by analyzing the contextual and influencing factors acting upon it and thus determine whether an already-assessed team is capable of executing it. Together, they allow the analysis of an organization's current projects and staff assignment through lenses optimized for the combined approach, enabling a confident hiring or shifting of personnel to meet the success criteria of more projects, or to suggest specialized training to improve skills in which employees are lacking.

Our acceleration model proposal garnered positive feedback when presented to both operational and managerial ORG staff (Sect. 3.5). As the specifics of the team assessment and project evaluation model are still in development, we present the worldview of our model, which is grounded in our conceptual modeling, case study findings, and select maturity model studies.

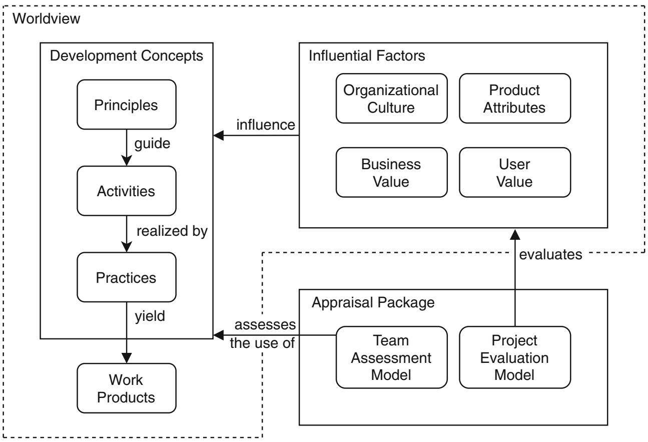


Fig. 2. Overview of the acceleration model.

4.1 Worldview

As previously stated, the model enables its user to identify whether a team is adequately qualified to fulfill the required expertise to successfully develop the deliverables of a project. It does so by analyzing the context and nature of a project and determining what are the qualifications required in order for a team to develop the project adequately, and assessing teams to see if they meet these criteria. As such, the model bases itself on the worldview depicted in Fig. 2.

To clarify the worldview, the use of the combined approach results in the development of a series of project deliverables, but its use is influenced by the project's context and nature, which directly impacts the project deliverables themselves. The elements of the worldview are described next.

Development Concepts Elements. In order to determine how the concepts in the combined approach interact with one another, we grouped them into distinct categories. These high-level elements arose from the conceptual modeling efforts for the combined approach and are measurable "points of interest" to be used in the team assessment model.

- **Principles:** the principles that serve as the foundations for each pillar, guiding the whole development process (e.g., *customer satisfaction* for XP, *be people-centered* for UCD, and *validated learning* for Lean Startup).
- **Activities:** abstract actions that produce expected outcomes key to the development process (e.g., Lean Startup's *measure* activity, which generates quantifiable and actionable data).
- **Practices:** also known as techniques or concepts, practices are concrete actions that directly result in tangible or intangible artifacts that contribute to the resolution of a specific activity (e.g., *pair programming* results in code and shared understanding of a system, contributing to XP's *coding* and *designing* activities). Practices can also implement a change in behavior to better enact principles (e.g., *collective code ownership* and *sustainable pace*).
- **Work Products:** work products are tangible or intangible artifacts consumed, produced, or modified by practices (e.g., source code, system usage metrics, understanding of a system).

As per what we observed on the case study, we emphasize that contemplating the principles of each pillar is essential for a successful adoption of the combined approach, as they guide a team's behavior and mindset, leading to the their better understanding of the rationale behind specific practices, and thus resulting in better decisions made during development. We will be careful for our team assessment model not to become too "prescriptive" [17] as in regards to practices, as this often hinders an individual's internalization of principles.

Influential Factors Elements. Several factors often negatively impact the use of any given development approach during and after the adoption process. The challenges effected from these factors are of several natures (e.g., change resistance, hierarchical management, and organizational boundaries), which makes understanding these influential factors crucial to facilitate a team's usage of the combined approach. Building on some maturity models we identified in Step 1 [15,23,27], we determined the following influential factors:

- **Organizational Culture:** the values, expectations, underlying assumptions, collective memory, and definitions shared by organization members that affect employee behavior [4].
- **Product Attributes:** the attributes of the product being developed, which directly influences the amount of effort that goes into each development activity, as each product requires different competences to be realized [26].
- **User Value:** an experience the user can partake in through means of a product or service that can be characterized in 3 dimensions: intrinsic-extrinsic, self-oriented versus other-oriented, and active-reactive [3].
- **Business Value:** strategic decisions and projects that aim to increase the health and well-being of the organization in the long term [26].

Essentially, these factors compose the context in which the team and the project is inserted in, and some can be directly addressed by using Agile, UCD, and Lean Startup (e.g., UCD explores user value directly while Lean Startup does the same for the business side). As such, given an environment with specific influence factors, one can examine them to determine which parts of the pillars should be focused on in order to obtain better results, thus helping the organization focus on its needs [17].

5 Conclusion

We envision our model as being able to support teams not by prescribing how things should work, but by enabling them to adapt to contextual factors and thus be flexible enough to adhere to their current needs. However, each organization has its own policies and compliance concerns that might require enforcement of specific practices that go against the optimal use of the combined, and dealing with this issue is out of scope for our model.

As the model is largely shaped on the inherent subjectivity of the design process, its validity is threatened by the non-repeatability of the creativity process, as well as the "correcteness" of our interpretation of the data used to develop the model. In addition, the heavy influence of the ORG environment might lead the model to not be as generally applicable as we would like. We plan to counter this through extensive testing of the model in other distinct organizations.

Our model proposal deals with context-specific factors, an idea that is proposed in existing work [15,23,27], but is not thoroughly dissected to fit our purposes. As such, our immediate next step is to discover what are the possible (and measurable) factors for each influential factor category. Afterwards, the proper design of the team assessment and project evaluations models is next.

As for other future work, we highlight that perceptions and understandings of the combined approach, which is of growing interest, are not yet set in stone and continue to be a fertile field of research.

Acknowledgements. We acknowledge that this research is sponsored by Dell Brazil using incentives of the Brazilian Informatics Law (Law no. 8.2.48, year 1991).

References

1. Abras, C., Maloney-krichmar, D., Preece, J.: User-centered design. In: Bainbridge, W. Encyclopedia of Human-Computer Interaction. Sage Publications, Thousand Oaks (2004)
2. Beck, K., Andres, C.: Extreme Programming Explained: Embrace Change, 2nd edn. Addison-Wesley, Boston (2004)
3. Boztepe, S.: User value: competing theories and models. Int. J. Design **1**, 55–63 (2007)
4. Cameron, K., Quinn, R.: Diagnosing and Changing Organizational Culture: Based on the Competing Values Framework. Wiley, Hoboken (2011)
5. Dobrigkeit, F., de Paula, D.: The best of three worlds: the creation of innodev, a software development approach that integrates design thinking, scrum and lean startup. In: Proceedings of the 21st International Conference on Engineering Design, Vancouver, Canada, pp. 319–328 (2017)
6. Grossman-Kahn, B., Rosensweig, R.: Skip the silver bullet: driving innovation through small bets... Lead. Through Design 815–830 (2012)
7. Hoda, R., Salleh, N., Grundy, J., Tee, H.M.: Systematic literature reviews in agile software development: a tertiary study. Inform. SW Tech. **85**, 60–70 (2017)
8. Julian, B., Noble, J., Anslow, C.: Agile practices in practice: towards a theory of agile adoption and process evolution. In: Kruchten, P., Fraser, S., Coallier, F. (eds.) XP 2019. LNBIP, vol. 355, pp. 3–18. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-19034-7_1
9. Kitchenham, B., Charters, S.: Guidelines for performing systematic literature reviews in software engineering (2007)
10. Krippendorff, K.: Content Analysis: An Introduction to Its Methodology. SAGE, Thousand Oaks (2018)
11. Kuusinen, K., et al.: Knowledge sharing in a large agile organisation: a survey study. In: International Conference on Agile Software Development, Germany, pp. 135–150 (2017)
12. Maier, A.M., Moultrie, J., Clarkson, P.J.: Assessing organizational capabilities: reviewing and guiding the development of maturity grids. IEEE Trans. Eng. Manage. **59**(1), 138–159 (2012). https://doi.org/10.1109/TEM.2010.2077289
13. Moralles, C., et al.: On the mapping of underlying concepts of a combined use of lean and user-centered design with agile development: the case study of the transformation process of an IT company. In: Meirelles, P., Nelson, M.A., Rocha, C. (eds.) WBMA 2019. CCIS, vol. 1106, pp. 25–40. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-36701-5_3
14. Paasivaara, M., et al.: Large-scale agile transformation at ericsson: a case study. Empir. Softw. Eng. **23**, 2550–2596 (2018)
15. Packlick, J.: The agile maturity map a goal oriented approach to agile improvement. In: Proceedings of the AGILE, pp. 266–271. IEEE Computer Society, Washington, DC (2007)

16. Petersen, K., Vakkalanka, S., Kuzniarz, L.: Guidelines for conducting systematic mapping studies in software engineering: an update. Inf. Softw. Technol. **64**, 1–18 (2015)
17. Pries-Heje, J., Johansen, J., Messnarz, R., et al.: The SPI manifesto. Technical report, EuroSPI (2010)
18. Ries, E.: The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses. Crown Business, New York (2011)
19. Rumbaugh, J., Jacobson, I., Booch, G.: Unified Modeling Language Reference Manual, 2nd edn. Pearson Higher Education, London (2004)
20. Runeson, P., Höst, M.: Guidelines for conducting and reporting case study research in software engineering. Empir. Sofw. Eng. **14**, 131 (2008)
21. Salerno, L., Signoretti, I., Marczak, S., Bastos, R.: Repensando papeis em equipes Ágeis: Um estudo de caso no uso de uma abordagem combinada de desenvolvimento Ágil, user-centered design e lean startup. In: Anais Estendidos da X Conf Brasileira de Software: Teoria e Prática, pp. 72–77. SBC, Porto Alegre (2019)
22. Schön, E.-M., Winter, D., Escalona, M.J., Thomaschewski, J.: Key challenges in agile requirements engineering. In: Baumeister, H., Lichter, H., Riebisch, M. (eds.) XP 2017. LNBIP, vol. 283, pp. 37–51. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-57633-6_3
23. Sidky, A.S., Arthur, J.D., Bohner, S.A.: A disciplined approach to adopting agile practices: the agile adoption framework. Innov. Syst. Softw. Eng. **3**, 203–216 (2007)
24. Signoretti, I., Marczak, S., Salerno, L., de Lara, A., Bastos, R.: Boosting agile by using user-centered design and lean startup: a case study of the adoption of the combined approach in software development. In: Proceedings of the International Symposium on Empirical Software Engineering and Measurement, pp. 1–6. IEEE, Porto de Galinhas (2019)
25. Signoretti, I., Salerno, L., Marczak, S., Bastos, R.: Combining user-centered design and lean startup with agile software development: a case study of two agile teams. In: Proceedings of the International Conference on Agile Software Development (2020, in press)
26. Sliger, M., Broderick, S.: The Software Project Manager's Bridge to Agility, 1st edn. Addison-Wesley Professional, Boston (2008)
27. Soundararajan, S., Balci, O., Arthur, J.D.: Assessing an organization's capability to effectively implement its selected agile method(s): an objectives, principles, strategies approach. In: Agile Conference, pp. 22–31, August 2013
28. Vargas, B., Signoretti, I., Zorzetti, M., Bastos, R., Marczak, S.: On the understanding of experimentation usage in light of lean startup in software development context. In: Proceedings of the International Conference on Evaluation and Assessment in Software Engineering (2020, in press)
29. VersionOne: 13th annual state of agile report. Technical report, CollabNet VersionOne (2019)
30. Vilkki, K.: When agile is not enough. In: Abrahamsson, P., Oza, N. (eds.) LESS 2010. LNBIP, vol. 65, pp. 44–47. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-16416-3_6
31. Ximenes, B.H., Alves, I.N., Araújo, C.C.: Software project management combining agile, lean startup and design thinking. In: Marcus, A. (ed.) DUXU 2015. LNCS, vol. 9186, pp. 356–367. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-20886-2_34
32. Zorzetti, M., et al.: Maturity models for agile, lean startup, and user-centered design in software engineering: a combined systematic literature mapping. In: Proceedings of the International Conference on Enterprise Information (2020, in press)

Agile Team Work Quality in the Context of Agile Transformations – A Case Study in Large-Scaling Environments

Alexander Poth[1(✉)], Mario Kottke[1(✉)], and Andreas Riel[2(✉)]

[1] Volkswagen AG, Berliner Ring 2, 38436 Wolfsburg, Germany
{alexander.poth,mario.kottke}@volkswagen.de
[2] Grenoble Alps University, Grenoble INP, G-SCOP, CNRS, 38031 Grenoble, France
andreas.riel@grenoble-inp.fr

Abstract. The maturity of organizations is measured with process assessment models like the ISO/IEC 33001. The product quality is aligned with internal and external product quality characteristics based on models like the ISO/IEC 25010. With the shift from the Tailorism-driven process orientation to a more people centric organization, the two dimensions process and product quality have to be extended by the people or team quality dimension. The presented approach offers aspects for agile Team Work Quality (aTWQ), as well as related measurement indicators. The approach is evaluated in the large enterprise context of the Volkswagen AG. The indicators of aTWQ have been integrated and established in the agile project review for a sustainable agile transition of the company's Group IT.

Keywords: Agile Team Work Quality (aTWQ) · Large-scaling agile · Quality Assurance (QA) · Agile transformation

1 Motivation and Context for the Demand

Several big enterprises are in the process of agile transformation. Cisco [1], Ericsson [2], and Volkswagen [3] are only very few examples. They apply top-down or bottom-up approaches, depending on their organizational culture and individual demands and constraints [4]. Both approaches have to handle teams, multi-teams (often called *scale*), as well as entire organizations (often called *large-scale*). These demands and constraints may evolve during the agile transition process.

In the Volkswagen Group IT case, a bottom-up approach has been applied:

– Start on individual team level (initial spark).
– Scale to multi-teams projects (scale).
– Scale to organizations (large-scale).

The IT governance accompanies these phases with the agile project review which has to scale from individual project teams to bigger organizational entities [5]. To determine the individual teams' progress and maturity levels, a systematic evaluation

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 232–243, 2020.
https://doi.org/10.1007/978-3-030-56441-4_17

approach is needed. This approach, which we will call aTWQ (agile Team Work Quality) in the following, has to support the agile mindset in which teams can improve by themselves without external assessments. It has to address the following requirements and constraints for application in the legislative/compliance and cultural context that is typical for European enterprises:

– The smallest check level shall be the team in order to avoid individual performance assessments for alignment with the workers' councils mindset.
– The approach shall be appropriate for integration in project and program reviews to measure transition progress from a governance perspective.
– The approach shall be applicable as a self-service by the teams to ensure scaling without centralized coaching etc.

In [6], a key observation is "although a theoretical distinction is possible between TWQ and team members' success, we found no empirical distinction between the two concepts". This further justifies the team focus on the lowest level rather than the individual team member focus.

The lean and agile approaches frequently used in industry, such as Scrum and SAFe®, do not address TWQ explicitly. In SAFe®, one of the four core-values is "Build-in Quality" [7]. In the deep dive documentation [8], however, the focus is product quality and "Flow" as a generic construct for all other aspects of quality. The process quality is implicitly addressed by links to other topics. TWQ is not mentioned at all, and therefore much more implicit. On the other hand, the consequence of this observation is: everything that is needed for quality is done inherently. Missing explicit transparency about what should be done and how does not fit to the SAFe® approach, however. In Scrum, the heart of the value creation is the team, which is supported by the Definition of Done (DoD) for achieving product quality, as well as the team retrospectives for process improvement. The team itself does not get any kind of explicit quality-related instructions and tasks. Instead, the daily, open communication and commitments are essential parts of TWQ. These incompatibilities among different agile approaches have been our drivers for developing an aTWQ approach that is applicable to many different lean and agile approaches and emphasizes the most important dimension of an agile organization: the team, who creates the value with their processes and products.

The particular challenge related to addressing TWQ is in the fact that it is more abstract than value perceived by users. Any user can give feedback about any product quality aspect, as long as it affects the externally perceivable product quality (ISO/IEC 9126:2003 [9]). Internal quality aspects are typically hidden and invisible from the outside, which makes it difficult in lean and agile environments to identify and explicitly "spend effort" to address internal product quality factors related to the product creation process. The ISO/IEC 25010:2011 [10] makes this more transparent with the terms "quality in use" and "product quality". The latter, however, is often directly addressed by regulation and compliance requirements like security or reliability. This often makes the process quality to a "first class citizen", because there are powerful and influential (external) stakeholders for legal compliance etc. TWQ is even below this first abstraction layer, since it affects the team itself and the organization striving for higher performance. Without some explicit measures and metrics related to TWQ, a systematic development is difficult from the organizational point of view. Hence, with the aTWQ approach, we want to support the organization in developing teams and their performances.

In addition to the previously explained practical aspects, the following questions are interesting from a research perspective:

- Which aspects have to be addressed by aTWQ?
- Which aspects are generic and relevant every time?
- Which aspects have to be prioritized depending on the teams' products and services, their life-cycles, etc.?
- How can these aspects be measured?
- Is a maturity model needed/useful?
- Are team profiles useful to address SAFe® or other common agile approaches? Is there a need for product-specific profiles?
- What impact does the company culture have?

In order to address the practical requirements and constraints as well as the research questions listed above, this article is structured as follows: the subsequent Sect. 2 explains the research method and approach and elaborates on the aTWQ modeldesign. Section 3 demonstrates the intended application and evaluation of aTWQ in the corporate context of the Volkswagen Group IT. Finally, Sect. 4 discusses the approach's limitations and gives an outlook.

2 The Design of an Approach to aTWQ

Our approach is based on Design Science Research [11] and iterative improvement within the use case (UC) of application. The iterations are used to implement the UCs. UC1 is the self-assessment of teams, UC2 is the self-assessment of larger entities like organizational units. UC3 is an external review by the corporate governance.

Team work aspects have been treated to a large extent in literature, e.g. [12–14]. Some of this previous work addresses agile team work quality explicitly [15, 16], some also propose organizational models fostering team work quality [17] or [18]. The three most interesting approaches are the Team Work Quality (TWQ) [16], Team Climate Inventory (TCI) [19] and Group Development Questionnaire (GDQ) [18] because they address both the team development and maturity. The TWQ approach focuses on quality indicators of team work. The TCI approach developed and established over years, evaluates teams indicators related to the teams' working structures for innovation. The GDQ approach focuses on evaluating the teams' alignment with stages of group development.

In particular, key conclusions of [20] are: "Even if agility is a concept that is hard to define, the agile practices can be seen as enablers for group maturity" and "we believe that a culture that is responsive to change in absolute terms … does need to build a mature team (with productive groups norms, good conflict resolution and decision-making techniques etc.) in order to be able to work in such a flexible way." This empirical observation leads us to setting up the following approach as a basis of our aTWQ approach:

a) Team Performance is based on TWQ.
b) TWQ and the TCI have similar "content".
c) TCI works well with GDQ.
d) For specific lean and agile approaches, their aspects are added to the aTWQ approach.

Depending on the (enterprise) culture, there may be some design constraints to be taken into account, such as:

– The approach shall not use specific roles that are typically fulfilled by a particular person to avoid individual performance measures to be aligned with employee representatives mindset in enterprises.
– the approach shall be like a self-service-kit including self-assessment for fostering "team-building" in a personal or team-private fashion for higher acceptance: Many teams do not like it to be assessed and pushed by external parties (need for team autonomy).

Based on [16] and [21], we derived the initial Team-Level approach with the six aspects communication, coordination, balance of contribution, mutual support, effort, cohesion. These six quality aspects also lead to team performance [6]. This legitimates economically the effort required for the actions of measurement and further TWQ improvement.

The TCI with its four key aspects and the short-set of 19 questions extends the TWQ with its six key aspects to a holistic team evaluation questionnaire for aTWQ. The objective of the questionnaire shown in Table 1 has been to create a concise single questionnaire by merging TCI and TWQ aspects. To achieve this, the TCI questions have been extended by the missing TWQ aspects. TWQ aspects not explicitly covered by the TCI questionnaire have been added and printed in italics. Terms printed in bold letters signify the most important aspects of the respective question. Column 3 and 4 show the mapping of the questions to Scrum and SAFe®, respectively, based on the specific approach's elements covering the aspects addressed by the questions. Hence, the TCI/TWQ questions represent generic practices, while the associated elements from Scrum or SAFe® represent specific practices of either approach. Both combined constitute the practice set of aTWQ in a specific team environment. The sparsely populated columns 3 and 4 indicate that neither Scrum nor SAFe® cover aTWQ aspects well. The indicators of the approaches are based on the current versions of SAFe 5.0 [22] and for Scrum based on the Scrum Guide version of November 2017 [23].

For the integration into the project reviews [24] evaluating individual product teams, a group of teams (like programs), as well as entire organizational units, an extension beyond a typical team size is needed. For the context of aTWQ, a team is constituted by people who have common goals within a purpose. The team size is aligned with the agile definition of 7–9 individuals [25]. A group is a collection of people or teams coordinating outcomes and efforts.

In the aTWQ approach, the extension to groups larger than one team is realized with the Group Development Questionnaire (GDQ) because in scaling agile approaches there is no “one big team”. In SAFe®, for example, there exist different types of teams like the technical and business teams sharing a common basic approach. “Both types of teams strive for fast learning by performing work in small batches, assessing the results, and adjusting accordingly” [26]. This leads us to deriving that in SAFe®, a group of different types of teams is managed. To handle this appropriately, something beyond TWQ is needed to show that the group which forms a SAFe® environment works fine. Furthermore, it is not assured by default that teams of the same type share a common purpose or challenging goals within an Agile Releases Train (ART). Perfectly independent feature teams share only the same time-line for their deliveries during the ART. For the mapping, aTWQ uses the following observation of [27]: “How are team climate theory (measured by TCI) and the integrated model of group development (measured by the GDQ) connected? … the two models are closely related”. Based on this, the mapping has been made in column 5 (Level) of Table 1.

The evaluation of the readiness of organizations is based on the spiral dynamics approach, which is usable in larger social systems like the GDQ. These two models provide the basis for using the aTWQ approach from individual teams to larger organizational units including many teams that work for some shared objectives.

Based on the ISO/IEC 33001:2015 [28], a maturity rating of each question can be made in the four categories: non (0–15%), partly (16–50%), largely (51–85%) and fully (86–100%). If needed, each topic (column 1 in Table 1) can be evaluated based on the rating of its questions. The rating is needed at least in the context of UC3 to be able to compare different teams. To have some specific indicators for the rating, column 3 and 4 can be used. Furthermore, column 5 is an indicator for the maturity of teams based on the TCI/GDQ approach. The levels represent the following GDQ approach stages: (I) Dependency and inclusion, (II) Counter-dependency and fight, (III) Trust and structure, and (IV) Work and productivity. In Table 1, column 5, the numbers in parentheses indicate the rating aligned without the mindset objective primarily based on the formal application of the respective agile aspects only. For example, in the *Scrum theater,* people apply some Scrum methods “mechanically” without actually forming a Scrum team with an agile mindset – this Scrum theater have to be rated with the parentheses level.

The four maturity levels can also be easily mapped to the NPLF SPICE ratings.

Table 1. aTWQ questionnaire with specific indicators for Scrum and SAFe® and team development level.

Topic	Question (Base practices)	Scrum	SAFe®	Level
Participative safety	Do we have a “**we** are in it together” **attitude** *driven by the* ***ability and willingness*** *to help and support each other in carrying out their tasks?*			IV
	Do people keep **each other informed** about work-related issues in the team *supported by a frequent communication?*	Daily Scrum	Program, team backlog,	I
	Do people **feel understood** and accepted by one another?			III
	Are there real attempts to **share information** throughout the team driven *by openness of the information exchange?*	Daily Scrum, Retrospective	Portfolio Kanban, Inspect & Adapt	(I) III
	Is there a lot of **give and take** *by the team members’ motivation to maintain the team?*		Innovation and Planning Iteration	IV
	Do we **keep in touch** with one another as a team *by accepting that team goals are more important than individual goals?*		Pairing/frequent review	III
Support for innovation	Is this team always moving towards the development of new answers?			IV
	Is this team open and responsive to change?	Inspect & Adaptation	Innovation and Planning Iteration	III
	Do people in this team always search for fresh, new ways of looking at problems?	Retrospective	Innovation and Planning Iteration, PI Planning	III
	Do members of the team provide and share resources to help in the application of new ideas driven *by team members’ ability and willingness to share workload?*	Inspect & Adaptation	Innovation and Planning Iteration	III
	Do team members provide practical support for new ideas and their application by *prioritize the teams’ task over other obligations*?	self-organizing	Innovation and Planning Iteration	IV
Vision	How clear are you about what your team’s objectives are?	(Product) Vision, Sprint Goal	Vision	I
	To what extent do you agree with these objectives?	Sprint commitment	PI planning	I

(*continued*)

Table 1. (*continued*)

Topic	Question (Base practices)	Scrum	SAFe®	Level
	To what extent do you think other team members agree with these objectives?	Refinement	ART commitment	I
	To what extent do you think members of your team are committed to these objectives?	Sprint commitment, DoD	ART commitment	I
Task orientation	Do your team colleagues provide useful ideas and practical help to enable you to do the job to the best of your abilities?		Pairing/frequent review	IV
	Are team members prepared to question the basis of what the team is doing?	Daily Scrum, Refinement		IV
	Does the team critically appraise potential weaknesses in what it is doing in order to achieve the best possible outcome?	Refinement, Retrospective		II
	Do members of the team build on one another's ideas in order to achieve the highest possible standards of performance?	Refinement, Retrospective		(I) IV
Coordination	Is there a **common understanding** when working on parallel subtasks, and agreement on common work breakdown structures, schedules, budgets and deliverables?	Backlog, Stories	Roadmap, Portfolio, ART, Iteration plan, Stories	III

3 Evaluation and Improvement Iterations

The evaluation, rollout and establishment of aTWQ will follow these steps:

– Simulation based on historical experience of coached agile teams.
– Proof-of-Concept (PoC) with coaches in projects/programs with SAFe® and Scrum.
– Agile Community feedback about the questionnaire.
– Self-Service-Kit.
– "Coaching offer".
– Integration into project-reviews.

3.1 Simulation

In the first step, the initially designed approach was simulated with the coaches of the Agile Center of Excellence (ACE) [3] and quality experts from the Quality Innovation Network (QiNET) [29]. The simulation was realized by virtual application of the aTWO questionnaire to teams coached in the past. For each simulation a point in the past was used as timestamp for answering the aTWO questions based on the situation around the timestamp. During the simulation the answers of the teams were simulated by the coaches/experts based on their knowledge about the team. Based on the answers potential chances and risks for the team development were derived. Then the timestamp was move ahead to check if the chances or risks identified by the aTWO approach are realistic to validate the questionnaire as a starting point for team improvements.

3.2 Proof of Concept

An initial Proof of Concept (PoC) was done in the Scrumban aligned product team of TaaS [30]. Additional two teams followed in sequence to improve the question-naire and the self-service kit iteratively. The self-assessments taken ca. 1.5 h.

Some facts about the TaaS PoC: The concerned service was introduced in 2016 and has been offered in the Volkswagen Group since 2017. Over the years, evolving the team constellations have led to an established devops team with end-to-end responsibly for the service delivery. In April 2020, the team included an internal product owner, two internal software engineers and one external software engineer with a primary focus on product development and third-level ops-support, as well as one external part-time devops engineer with primary focus on first and second-level support and some third-level support activities. The team members' experience levels covers a wide range from junior developer to senior engineer. The teams constellation was changed a few weeks ago and the team was in a re-balancing phase. The application of the aTWQ questionnaire worked fine and was conducted as a dedicated task of a team retrospective. The identified enhancement potentials were used like retrospective outcomes and lead to actions for team improvement. Some small improvements based on the feedbacks and observations were made about aTWQ and are reflected in the version of Table 1.

3.3 Feedback, Self-service Kit and Coaching

In a second step, two other applications were conducted in similar team sizes and the self-assessment was conducted in a similar way to get more feedbacks from application and see how the self-service maturity is. A further reflection was provided by the feedback of the Agile Community (AC) of the Volkswagen AG to get a more holistic feedback about the understanding of the questions from different people to see what is additionally needed as explanation and supporting material in a self-service kit. To support the scaling approach of the Group IT [24], a self-service kit is designed with the questionnaire, additional information about "how to use" the kit to establish a continuous improvement of the team work quality.

For teams with more support demand, the ACE will offer coaching packages for the adoption of the aTWQ approach.

3.4 Integration into Agile Project Review

Integration into Agile Project Review [31] was the next step for the sustainable establishment of the aTWQ approach at the Volkswagen Group IT. To make this step, the aTWQ approach needed commitment in the ACE and AC. It is expected that later in 2020 this usage leads to more data about enhancements of the approach.

4 Conclusion and Outlook

Good QM/QA (Quality Management/Quality Assurance) is integrated into the products, processes and the teams by design. It is also integrated into trainings and coaching, as well as intuitively applicable and iteratively improvable. It is systematically safeguarded by design or with other measures. With the aTWQ, a series of QM and QA work within the Volkswagen Group IT from the last years is extended and systematically enhanced to more completeness:

- In 2016, the topic agile QM was identified as a strategic topic.
- In 2017, the work was systematically addressed with the QiNET and its work on self-service kits with the PQR approach for an agile product quality.
- In 2018, the agile project review and Level of Done (LoD) approach addressed the agile process quality.
- In 2019, the reflections on the optimization of scaling quality in agile transitions started.
- In 2020, the team maturity and aTWQ addressed the team-work quality.

We believe that our journey is on the right track, but we have to strive for further improvements and the ongoing establishment of a truly agile QM/QA mindset.

4.1 Contributions and Added Value

The key contributions *to practice* can be summarized by the following aspects:

- aTWQ defines a model for the awareness of the team-dimension of the three quality dimensions product-, process- and team-quality.
- an explicit indicator set has been defined for the most popular agile approaches Scrum and SAFe®; it can be used during agile transformation.
- first evidences for relevance and added-value for effective transformations in Scrumban environments have been given.

The key contributions *to theory* can be summarized by the following aspects:

- identification of the gap between the current quality-models to the real world in industrial settings which emphasize agile team work which is not explicitly addressed and covered by the established product and process quality models and approaches.
- identification of possible approaches to reduce this gap by integration of the TCI, TWQ and DGQ approach to the aTWQ approach with focus on application in real world product teams.

- initial analysis about the state-of-the-art as basis for more sophisticated research about the added value created by the aTWQ approach with some interesting challenges such as:
  - Evaluate and study the team-, multi-team- and organizational-level effects of aTWQ.
  - Extension of the view from team concepts to the spiral dynamics approach with GDQ for larger groups or organizations.
  - Extension to a more fine-grained approach-set like for example the filling of Scrum-of-Scrum in between the existing two approaches Scrum and SAFe.

4.2 Limitations

The context of the development and initial evaluation of aTWQ is a large enterprise setting with a European culture and mindset. The scope is focused to the described set of constraints given in this setting. This narrows the possibilities and degrees of freedom by design and during evaluation too. Furthermore the trade-of between fine grained evaluation models like the current Automotive SPICE approach and less fine grained models is a topic here too. Because it lets room for interpretation of what is adequate and what is not if no explicit evidences are expected and no indicators are given by the evaluation model. Currently aTWQ has an open design to leave the decision by the teams in case of self-application and by the reviewer from the governance in case of "external" team evaluations. The interpretation by a more or less constant governance reviewer team will give sufficient comparability between the teams within an organization. Real mature agile teams will actively request for external "feedbacks" to get the ranking to other teams and learn from external inspiration for their improvement journey. This kind of limitation is a chance by design to ensure continuous improvement within the teams and organizations because they have not static target like an evidence or indicator list which have to be fulfilled and the *"aTWQ story is done"*.

4.3 Contribution to the SPI Manifesto

The current version of the SPI Manifesto [32] addresses "people" as one core pillar. However, people have to work together as a team to drive the SPI initiative to a success. This goes further than the current Manifesto element "Successful SPI is based on actively involved people having sufficient information and training" [32]. This implicitly aligned team is not a sufficiently explicit quality characteristic and not addressed in the current SPI Manifesto. The presented aTWQ approach can be provide a basis for checking the "peoples' team capabilities" for the SPI initiative. Referencing or integrating the aTWQ in a future version of the SPI Manifesto is therefore strongly encouraged. The aTWQ questionnaire could be provided as a hands-on appendix like a self-service kit to check the readiness of people and teams for the SPI initiative.

4.4 Outlook

Future investigations include spreading the approach to non-IT projects, as well as other brands [33] and markets with different cultures in order to show how generic the approach is. Furthermore, some of the research questions from Sect. 1 are not (sufficiently) answered at the current state of evaluation of the aTWQ approach. Additional insights about the initial analysis aspects of the contribution listed in Sect. 4.1 are also envisaged.

References

1. Chen, R.R., Ravichandar, R., Proctor, D.: Managing the transition to the new agile business and product development model: lessons from Cisco Systems. Bus. Horizons **59**(6), 635–644 (2016)
2. Paasivaara, M., Lassenius, C., Heikkilä, V.T., Dikert, K., Engblom, C.: Inte-grating global sites into the lean and agile transformation at ericsson. In: 2013 IEEE 8th International Conference on Global Software Engineering, Bari, pp. 134–143 (2013)
3. Poth, A.: Effectivity and economical aspects for agile quality assurance in large enterprises. J. Softw.: Evol. Process **28**(11), 1000–1004 (2016)
4. Karvonen, T., Sharp, H., Barroca, L.: Enterprise agility: why is transformation so hard? In: Garbajosa, J., Wang, X., Aguiar, A. (eds.) XP 2018. LNBIP, vol. 314, pp. 131–145. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-91602-6_9
5. Poth, A., Kottke, M., Riel, A.: Scaling agile – a large enterprise view on delivering and ensuring sustainable transitions. In: Przybyłek, A., Morales-Trujillo, M.E. (eds.) LASD/MIDI -2019. LNBIP, vol. 376, pp. 1–18. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-37534-8_1
6. Lindsjørn, Y., Sjøberg, D.I., Dingsøyr, T., Bergersen, G.R., Dybå, T.: Teamwork quality and project success in software development: a survey of agile development teams. J. Syst. Softw. **122**, 274–286 (2016)
7. SAFE® Core Values. https://www.scaledagileframework.com/safe-core-values/. Accessed 03 June 2020
8. SAFE® Built-In Quality. https://www.scaledagileframework.com/built-in-quality/
9. ISO/IEC 9126:1991 Software engineering—Product quality. https://www.iso.org/standard/16722.html. Accessed 03 June 2020
10. ISO/IEC 25010:2011 Systems and software engineering—Systems and software Quality Requirements and Evaluation (SQuaRE)—System and software quality models. https://www.iso.org/standard/35733.html. Accessed 03 June 2020
11. Hevner, A., Chatterjee, S.: Design science research in information systems. In: Hevner, A., Chatterjee, S. (eds.) Design Research in Information Systems. Integrated Series in Information Systems, vol. 22, pp. 9–22. Springer, Boston (2010). https://doi.org/10.1007/978-1-4419-5653-8_2
12. Strom, P.S., Strom, R.D.: Teamwork skills assessment for cooperative learning. Educ. Res. Eval. **17**(4), 233–251 (2011)
13. Lingard, R.W.: Teaching and assessing teamwork skills in engineering and computer science. J. Syst. Cybern. Inf. **18**(1), 34–37 (2010)
14. Willey, K., Freeman, M.: Completing the learning cycle: The role of formative feedback when using self and peer assessment to improve teamwork and engagement. In AAEE-Annual Conference of Australasian Association for Engineering Education. School of Engineering, Auckland University of Technology, Auckland, New Zealand (2006)

15. Ramírez-Mora, S.L., Oktaba, H.: Team maturity in agile software development: the impact on productivity. In: 2018 IEEE International Conference on Software Maintenance and Evolution (ICSME), Madrid, pp. 732–736 (2018)
16. Hoegl, M., Gemuenden, H.G.: Teamwork quality and the success of innovative projects: a theoretical concept and empirical evidence. Organ. Sci. **12**(4), 435–449 (2001)
17. Beck, D.E., Cowan, C.C.: Spiral Dynamics: Mastering Values, Leadership and Change. Wiley, Hoboken (2014)
18. Wheelan, S.A., Hochberger, J.M.: Validation studies of the group development questionnaire. Small Group Res. **27**(1), 143–170 (1996)
19. Anderson, N., West, M.A.: The team climate inventory: development of the TCI and its applications in teambuilding for innovativeness. Eur. J. Work Organ. Psychol. **5**(1), 53–66 (1996)
20. Gren, L., Torkar, R., Feldt, R.: Group maturity and agility, are they connected? – A survey study. In: 2015 41st Euromicro Conference on Software Engineering and Advanced Applications, Funchal, pp. 1–8 (2015)
21. Dikert, K., Paasivaara, M., Lassenius, C.: Challenges and success factors for large-scale agile transformations: a systematic literature review. J. Syst. Softw. **119**, 87–108 (2016)
22. SAFe® 5.0. https://www.scaledagile.com/safe-50/. Accessed 27 May 2020
23. The Scrum Guide™ https://www.scrumguides.org/docs/scrumguide/v2017/2017-Scrum-Guide-US.pdf#zoom=100. Accessed 03 June 2020
24. Poth, A., Kottke, M., Riel, A.: Scaling agile on large enterprise level – systematic bundling and application of state of the art approaches for lasting agile transitions. In: 2019 Federated Conference on Computer Science and Information Systems (FedCSIS), Leipzig, Germany, pp. 851–860 (2019)
25. Rodríguez, D., Sicilia, M.A., García, E., Harrison, R.: Empirical findings on team size and productivity in software development. J. Syst. Softw. **85**(3), 562–570 (2012)
26. SAFE® Agile Teams. https://www.scaledagileframework.com/agile-teams/. Accessed 03 June 2020
27. Jacobsson, C., Wilmar, M.: Group processes–the links between team climate inventory and group development questionnaire. Clin. Exp. Psychol. **5**, 1–4 (2019)
28. ISO/IEC 33001:2015 Information technology—Process assessment—Concepts and terminology. https://www.iso.org/standard/54175.html. Accessed 03 June 2020
29. Poth, A., Heimann, C.: How to innovate software quality assurance and testing in large enterprises? In: Larrucea, X., Santamaria, I., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2018. CCIS, vol. 896, pp. 437–442. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-97925-0_37
30. Poth, A., Werner, M., Lei, X.: How to deliver faster with CI/CD integrated testing services? In: Larrucea, X., Santamaria, I., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2018. CCIS, vol. 896, pp. 401–409. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-97925-0_33
31. Poth, A., Kottke, M.: How to assure agile method and process alignment in an organization? In: Larrucea, X., Santamaria, I., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2018. CCIS, vol. 896, pp. 421–425. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-97925-0_35
32. SPI MANIFESTO. https://2020.eurospi.net/images/eurospi/spi_manifesto.pdf. Accessed 03 June 2020
33. Poth, A., Wolf, F.: Agile procedures of an automotive OEM – views from different business areas. In: Stolfa, J., Stolfa, S., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2017. CCIS, vol. 748, pp. 513–522. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64218-5_42

Emerging Software Engineering Paradigms

The SMARTSEA Education Approach to Leveraging the Internet of Things in the Maritime Industry

George Katranas[1], Andreas Riel[2,3](✉),
Juan Manuel Corchado-Rodríguez[4], and Marta Plaza-Hernández[4]

[1] Cerca Trova Ltd., Sofia 1000, Bulgaria
katranasg@ctrova.eu
[2] Grenoble Alps University, Grenoble INP, G-SCOP, CNRS,
38031 Grenoble, France
andreas.riel@grenoble-inp.fr
[3] ECQA GmbH, 3500 Krems, Austria
[4] BISITE Research Group, University of Salamanca, 37007 Salamanca, Spain
{corchado,martaplaza}@usal.es

Abstract. The Internet of Things technology solutions are gaining importance in shipbuilding industry to improve transparency, safety, and cost efficiency by optimising procedures, maintenance and energy efficiency. The marine surveying infrastructure transformation through IoT technologies is expected to enable the shipping industry, port authorities or environmental agencies, to inspect shipping assets, offshore structures, waterways and ensure compliance with various standards or specifications. The SMARTSEA approach presented in this paper aims at qualifying students and professionals for working in modern smart maritime and surveying industry. A further objective is stimulating transversal competences such as the increased sense of initiative and entrepreneurship. An original programme for piloting and validating the programme across Europe is also presented.

Keywords: Internet of Things · Shipbuilding · Marine industry · Education · Training

1 Introduction

The Internet of Things (IoT) is a network of physical "smart" devices (applicable in vessels, vehicles, buildings, factories, etc.) embedded with electronics, software, sensors, actuators, that allow interconnectivity between these devices and data exchange. Over the past 5 years this new technology has grown rapidly and has found applications ranging from people whose devices monitor health and wellness to manufacturers that utilize sensors to optimize the maintenance of equipment and protect the safety of workers. It is expected that by 2025, IoT may reach a total potential market impact of up to $11.1 trillion [1]. Large shipping corporations are already investing heavily in enabling IoT technology solutions in their fleet to improve transparency, safety, and cost efficiency by optimising procedures, maintenance and energy efficiency. Shipowners are

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 247–258, 2020.
https://doi.org/10.1007/978-3-030-56441-4_18

set to spend an average of $2.5 m, each, on IoT solutions over the next three years [2]. In addition, the marine surveying infrastructure transformation through IoT technologies is expected to enable the shipping industry, port authorities or environmental agencies to inspect shipping assets, offshore structures, waterways and ensure compliance with various standards or specifications. In contrast, formal Maritime & Surveying IoT training on development, installation, service, maintenance and sustainability awareness is at its infancy, especially in the European Union. After meticulous research it was made evident that currently there are no hands-on IoT educational and training programs dedicated to the application of these systems to the Shipping industry in all 90 maritime educational institutions inside the EU, neither are there in the 27 European (non-EU) educational establishments [3].

This paper presents the Surveying & MARiTime internet of thingS EducAtion (SMARTSEA) approach and project aiming at developing an advanced interactive certified Master of Science program related to Maritime & Surveyor IoT applications that will train individuals with the necessary skills and knowledge to work in the rising "Smart Maritime & Surveying" industry. The program/postgraduate is also formulated to stimulate transversal competences, such as the increased sense of initiative and entrepreneurship. It is designed to follow the European Credit Transfer and Accumulation System (ECTS) credit standards for certification recognition across the EU.

2 The IoT Context in the Maritime Sector

Digitizing key elements of our physical world is a powerful idea that is changing how goods are made and distributed, how products are serviced and refined, and how health and wellness is managed. Since the introduction of the IoT concept five years ago, it managed to revolutionise the opportunities for businesses to take their operational efficiency to a new level. Combining a large range of smart interconnected devices created new dimensions of opportunities in this digital era. This included the development of new networking capabilities and the design of analytical tools to perform the data fusion. These newly created Machine to Machine (M2M) direct intercommunication services comprise all technologies that allow network devices to exchange information and perform an action without any physical assistance of human beings. This is now used across different industries, such as in retail, banking and financial institutions, telecom and IT industry, healthcare, automotive, oil and gas, and transportation to enhance the safety, productivity, and efficiency of the establishments.

Because of the IoT's penetration to all major industries, the economic impact is expected to reach, by 2025, between $3.9 trillion to $11.1 trillion a year. This is translated to an equivalent of about 11% of the world economy [4].

For most market sectors the potential and the transition towards data-driven technology is a hard to comprehend and cumbersome process. This is in contrast to the transportation and logistics sectors that, due to their nature, have always relied on exchanging decision-making data and information. Hence, this market segment quickly adopted the new sensing and data exchange technology, placing them ahead in the transition towards IoT adaptation [5].

Especially in the Maritime sector, the implementation of IoT technology allows the shipping companies to connect the vessels in one platform for data sharing with the entire corporate ecosystem. Now asset management is possible, to receive data from all types of sensors and monitor the vessel status, location and cargo condition. Shipping companies can finally receive live information from engine and hull measurement systems and perform condition monitoring and predictive maintenance and optimise service time and energy efficiency. In addition, cargo owners can use these systems to receive live information on the status and location of their transferred goods. Overall IoT systems are set to reduce inefficiencies, risks and overall cost and provide a live sea-to-shore connectivity that stakeholders can exploit for decision-making.

Moreover, a major part in the maritime sector is the asset inspection for insurance and quality purposes. This task is handled by certified professionals called Maritime Surveyors, who are employed by the shipping industry, port authorities or environmental agencies, to inspect shipping assets, offshore structures, waterways and ensure compliance with various standards or specifications. Maritime surveys typically include the structure, machinery and equipment (navigational, safety, radio, etc.) and general condition of a vessel and/or cargo or ports and waterways. Maritime Surveying requires talented individuals that combine the impressive understanding of ships and boats with insurance and finance industries. Their knowledge targets the evaluation of a vessel's condition in a case of damage or of a request for appraisal value. In other cases, environmental surveying is mandatory if an incident occurs or if it is instructed by local authorities. In all the above cases the vessels' hull, plumping and electronics are evaluated including the surrounding waters for the presence of pollutants. Maritime Surveyors are individuals, usually with interdisciplinary studies, certified by international bodies [6, 7] to collect maritime related asset information for quality, legal and insurance needs. Until recently, a maritime survey was limited in taking photos at the site of interest and filing a report with respect to the regulations that affect the vessels operations. With the introduction of interconnected digital devices, surveyors are now equipped with analytical equipment such as ultrasonics, infrared cameras and in some cases with Remotely Operated Underwater Vehicles (ROUV) together with the support of specialized software for underwater inspection [8, 9].

The measured information can now be gathered, processed and visualised through web-based or mobile application user interfaces directly to the stakeholders for decision-making, even across the globe and in real-time. Overall, Information and Communication Technology (ICT)/and IoT systems are set to reduce inefficiencies, risks and overall cost and provide an unparallel ease of connectivity that stakeholders can exploit for decision-making.

Large shipping industries are already investing heavily on IoT and ICT techniques to improve and optimise transparency, safety and costs of their transports [1, 2]. Latest reports show that ship owners are set to spend an average of \$2.5 m, each, on IoT solutions over the next three years [10].

The European Union, through its Horizon2020 program, is providing significant funding of up to 6.3 billion Euros on the research and development of ICT and IoT technologies, showing the technology roadmap of the future [11, 12]. This clearly states that there is a need for a transition to an era of smart systems interconnectivity and procedure optimisation. The market has shown a rapid increase in Maritime &

Surveying IoT systems and their infrastructure and corporations globally are interested in investing and promoting their growth. These systems that have and will enter the market are currently designed, developed, installed serviced and maintained by people with no educational background specific to the needs of this market. In other words, there is a global shortage of trained engineers and technicians to cater safely handle this new technology.

This new technology will influence onshore and offshore personnel as they will need to have either an insight of the systems capabilities for corporate decisions, design/service this ecosystem or process/analyse the acquired data. The SMARTSEA project team detected at least 23 positions within the maritime/surveying job sector that are expected to require knowledge on IoT solutions, at various levels [13]. These positions are listed in Table 1.

Table 1. IoT influence on maritime positions.

No.	Position	Main duties	IoT influence
1	Cargo surveyor	Inspects vessel cargo and certifies compliance with health and safety regulations	Consults the IoT systems database to monitor the status and location of goods (stored or in transit)
2	Computer engineer	Develops and maintains systems software and hardware for industrial applications, as well as performing data analysis	Develops the IoT software
3	Electro-technical officer	Deals with maintaining electrical equipment and systems on vessels	IoT system maintenance
4	Fleet manager	Plans, directs and coordinates the transportation operations within the maritime company	Receives vital information on vessel and cargo status
5	IT specialist	Provides technical assistance to equipment users, maintenance of intranet and data flow	Maintains the IoT data flow through the corporate intranet and satellite communication systems
6	Marine architect	Designs and oversees construction and repair of vessels	Incorporates the IoT systems in new-build vessels or retrofits older ones
7	Marine engineer	Deals with maintenance and repair of ship's machinery	Uses the IoT systems to get centralised information on vessel status
8	Marine superintendent	Responsible for the supervision of operations, monitoring vessel's condition and compliance with policies and safety	Receives vital information on ship and cargo status
9	Marine & coastal surveyor	Collects scientific data and maps the underwater geophysics and environmental quality of oceanic and inland water bodies	Operates ICT/IoT enabled equipment and ROUVs to collect information

(*continued*)

Table 1. (*continued*)

No.	Position	Main duties	IoT influence
10	Maritime instructor	Teaches vocational or occupational subjects at the postsecondary level and prepare persons to operate and install industrial equipment	Responsible for teaching maritime professionals with the IoT systems ecosystem
11	Naval architect	Involves planning and designing of vessels	Incorporates the IoT systems in new-build military vessels or retrofits older ones
12	Maritime engineer	Designs, develops, builds, installs, inspects and maintains vessel systems and equipment	Responsible for installing the IoT systems on vessels to support the data measurement
13	Maritime surveyor	Conducts inspections, surveys or examinations of marine assets	Receives detailed and vital information from the inspection
14	Ocean engineer	Deals with the aspects of constructing offshore structures (planning, designing and construction)	Designs IoT solutions for offshore applications
15	Operations manager	In charge of the vessels' operation	Receives vital information on vessel and cargo status
16	Port engineer	Responsible for estimating, planning, and performing short and long-term maintenance, repairs, and modifications on vessels	IoT system maintenance, condition monitoring and predictive maintenance of maritime assets
17	Ship fitter	Involves repairing and maintenance of the ship under the guidance of the ship officer	Receives information on maritime system condition from the IoT systems and plans maintenance activities accordingly
18	Ship operator	In charge of the vessels' operation	Receives vital information on vessel and cargo status
19	Ship superintendent	Ensures that all the ship repairing work is performed as planned	Uses the IoT systems to get centralised information on vessel status
20	Shipbuilding engineer	Deals with the engineering aspect of design and construction of ships and marine vessels	Responsible for installing the IoT systems on new-build vessels
21	Structural engineer	Involves designing machinery and offshore structures	Receives information on machinery condition from the IoT system
22	Vessel electrician	Installs, maintains and repairs electrical wiring, equipment, control systems and fixtures	Installs, maintains and repairs the IoT system
23	Vessel mechanic	Involves repairing, installation of machinery and carrying out periodic upkeep or restoration of the vessel	Receives information on machinery condition from the IoT system and installs the sensors on the vessels equipment and hull

The positions mentioned in Table 1 are well known neuralgic positions within the maritime industry and are required to get up to date with the new advances on vessel related technology and benefits as part of their job descriptions. These are positions based either on land and/or at sea. Positions affected by the IoT technology phenomenon will be requested by surveyors, authorities, vessel operators, shipyards as well as thousands of private companies that deal with the new-builds, retrofit, maintenance, repair of vessels.

After meticulous research it was made evident that currently there are no hands-on IoT programs dedicated to the application of these systems to the Shipping industry in all 90 maritime educational institutions inside the EU, neither are there in the 27 European (non-EU) educational establishments [3]. Some engineering educational institutions, mainly in Northern Europe, have begun offering general ICT/IoT courses that are unable to convey solution to practical and technical problems that are faced within the shipping industry and it does not offer a hands-on training onboard vessels. Currently, the only ICT course globally in the market is the “ICT for Maritime Education and Training” by the World Maritime University, that focuses on software development introduction; its duration is five days, corresponding to 2.5 ECTS credits [14].

Our mission is to provide an innovative curriculum in the theme of Maritime & Surveying IoT systems and land-based infrastructure that includes interactive teaching methods and partnerships with major educational and industrial organizations, giving students a solid grounding for starting a fruitful career in the corresponding industry or enable professionals to gain extra skills knowledge and at the same time prompting local communities and authorities to embrace the new technologies and their benefits. Hence, the output of the partnership is to design a MSc program that will give additional skills to those interested in a maritime career and allow the partners to create an ongoing degree that will go well beyond the duration of this project.

For this purpose, seven academic institutions, two SMEs, one research centre and three shipping/environmental surveying and shipping companies joined in November 2019 to create the joint program SMARTSEA. More precisely, the consortium members are as follows (in alphabetic order, clustered by academic, research, SME, company):

- Infante D. Henrique Nautical School, Portugal;
- International Hellenic University, Greece;
- Maritime University of Szczecin, Poland;
- Tallinn University of Technology, Estonia;
- University of Ljubljana, Slovenia;
- University of Salamanca, Spain;
- National Institute for Marine Research and Development “Grigore Antipa”; Romania;
- Cerca Trova Ltd, Bulgaria;
- ECQA GmbH, Austria;
- RINA Hellas SA, Greece;
- Creocean; France;
- Danaos Shipping Company Ltd.; Greece.

3 Curriculum Design

The aim is to provide an innovative and eclectic mix of teaching methodologies that include state of the art equipment combined with work experience for students in real industrial environment and the globally acknowledged ECTS certification.

The interdisciplinary curriculum was meticulously designed with the participation of all consortium partners who brought their expertise from their academic/industrial background on ICT, Maritime, Marine & Business.

The syllabus will include reconfigurable laboratory apparatus, innovative demonstrator group assignment and an industrial experience to equip the participants with the necessary experience to enter the emerging market of IoT systems.

The program focuses in the following objectives:

- To develop an innovative program on Maritime & Surveyor ICT/IoTs to complete the market void in technical and maintenance specialists created by the rapid expansion of this industry;
- To develop and adapt a joint curriculum between participating HEIs, designed on an exhaustive needs analysis and focusing on a "real-life" transnational approach.
- To create a program recognized by academia and industry throughout the EU by offering a recognized and comparable educational path by using the ECTS credit system;
- To address participants from backgrounds with fewer opportunities;
- To enhance opportunities for cooperation and mobility between partner countries;
- To modernize the didactic techniques of higher education institutes by creating interactive learning methods and industrial tools;
- To create cost effective reconfigurable tools used across industries and online platform for global access to reduce the learning cost in academia and empower distant learners;
- To increase competence in new languages by interconnecting participants of different ethnic and linguistic background;
- To ignite entrepreneurship by using interactive teaching and participation methods that boost innovative thinking and to emerge them to this industry sector;

Fig. 1. Rolls Royce DART 510 engine. A view through the head-worn display predicts information provided using Augmented Reality (AR) to assist the marine mechanic [15].

Fig. 2. A Remotely Operated Underwater Vehicle (ROUV) used to underwater inspections, by the International Hellenic University [15].

Following the steps of the Fourth Industrial Revolution, a high-tech teaching set of experiments including computer sensory data analysis and ROUVs [15], along with computer guidance and Augmented Reality (AR) [16] will be an important asset of the curriculum and will match the emerging trend among surveying and shipping companies and researchers [17, 18] (Figs. 1 and 2). The surveying & maritime industry, educations and local economies are expected to benefit from this MSc program. Furthermore, a set of interactive reconfigurable laboratory apparatus will be developed to bring to life the course lessons and equip the participants with the latest industrial tools to enter this emerging market. The innovative teaching methods included are designed to stimulate entrepreneurship and enhance the participant's employability. The main modules/subjects presented in the curriculum will cover analysis and implementation of tools for advanced engineering problems, real life technical hands-on problems and maintenance techniques designed for Maritime & Surveyor IoTs, as well as market insight, business and innovation. Among the tools that the consortium will develop is the advanced e-Learning platform, that will be designed to offer teaching material, online experimentation and networking between the students.

The project aims towards the vocational education, certification and imminent professional rehabilitation of the students as also to improve the skills of the participating bodies and provide opportunities for entrepreneurship. The project also aims to sensitize and familiarize the local communities in order to accept and prepare for the future of smart sensing by presenting the benefits of this transition to Maritime & Surveyor IoTs in special seminars held at the locations of each partnership. The program is also addressed to participants from backgrounds with fewer life opportunities (immigrants, economically disadvantaged) to foster social integration and enhance intercultural understanding.

The modules'/courses' lessons that will cover the necessary background information and skills for teaching all major ICT/IoT components, technologies and surveying knowledge to the students as well as an insight on the market trends. Emphasis will be given to design a program with applied engineering concepts as well as industry insights to emerge the students into Maritime & Surveyor IoT core components, functionality, maintenance, safety and sustainability. This program is also designed to cater for the mobility of the students and educators. It is split into two time slots (TS) (Table 2), comprising 24 modules.

Table 2. Proposed SMARTSEA curriculum.

Code	Module title	ECTS	Lab	Work hours	Lecture		Lab	Exams	H/W	Practice
					Class	Web				
TS1.1	Maritime control systems	2	Y	50		24	10	4	12	
TS1.2	Marine surveying	2	–	50	30			5	15	
TS1.3	NI LabVIEW training	2	Y	50	10	8	20	3	9	
TS1.4	Data acquisition and sensors	2	Y	50		24	10	4	12	
TS1.5	IoT platforms & systems	2	Y	50		18	20	3	9	
TS1.6	ROUV Electric System	2	Y	50	24		10	4	12	
TS1.7	Artificial Intelligence	2	–	50	30			5	15	
TS1.8	Maritime environment	2	–	50	15	15		5	15	
TS1.9	Safety at work on the sea	2	–	50	20	10		5	15	
TS1.10	Diving observation techniques	2	–	50		30		5	15	
TS1.11	Language lessons	2	–	50	30			5	15	
TS1.12	Intermediate project	2.4	Y	60	20		40			
TS2.1	Remote sensing and positioning	2	–	50	10	20		5	15	
TS2.2	Distributed ledgers - blockchain	2	–	50		30		5	15	
TS2.3	Lightweight materials	2	Y	50		24	10	4	12	
TS2.4	Underwater comms & navigation	2	Y	50	16	8	10	4	12	
TS2.5	Data processing	2	Y	50	24		10	4	12	
TS2.6	Geographic information systems	2	–	50	30			5	15	
TS2.7	Underwater physics	2	Y	50		24	10	4	12	
TS2.8	Innovation & entrepreneurship mgmt.	2	–	50		30		5	15	
TS2.9	Business administration	2	–	50		30		5	15	
TS2.10	Maritime legal arrangement	2	–	50	20	10		5	15	

(*continued*)

Table 2. (*continued*)

Code	Module title	ECTS	Lab	Work hours	Lecture		Lab	Exams	H/W	Practice
					Class	Web				
TS2.11	Language lessons	2	–	50	30			5	15	
TS2.12	Developing tool demonstrator	2.4	Y	60	20		40			
–	Industrial practice	11.2	–	280					120	160
	Total		**8**	**1500**	**329**	**305**	**190**	**99**	**417**	**160**

The SMARTSEA curriculum is innovative also in the fact that it comprises interactive teaching methods and partnerships with expert academic and maritime organizations to give to the students a solid background for starting a fruitful career in the industry. The duration of the MSc program is nine months. During the execution of the program, three mobility periods are programmed. For the first two periods (duration: 14 days) the students and two educators from one university will travel to the other, and vice-versa, to participate in large-scale laboratories. The third period (duration: 1 month) is reserved for an industrial placement, a maritime on-the-job experience. Students will produce a master's thesis during the placement.

Although the postgraduate will be taught in English, local language lessons will be provided to enable the participants to immerge in the local culture during the exchange periods. The master's degree will be open for participation for anyone with a basic marine, electrical or mechanical technical background. Priority will be given according to their academic performance.

4 Conclusion and Outlook

The Industry 4.0 revolution is bringing to our everyday lives the concept of creating networks of smart interconnected objects with direct application in almost all major industries, including the maritime & surveying industry. The transportation & logistics sector has rapidly adopted the new sensing and data exchange technology, placing them ahead in the transition towards ICT and IoT adaptation. This market will rapidly require skilled individuals to design, develop, install, service and maintain these new systems.

The objective of the SMARTSEA Joint Master's Degree in Maritime & Surveying ICT/IoT systems presented here is to train engineers in the skills and knowledge they need to work in the maritime/surveying industry of the future. Hence, the rapid market growth will demand that this postgraduate program will become an integral part of academic institutions and maritime academies in Europe and beyond.

The first edition of the program (pilot program) will be financed by the European Commission. Financial sustainability of subsequent editions will be ensured by collecting student fees and through the scholarships that industrial companies will offer to train their personnel and stay competitive in this emerging field. The fees will cover the participant's training, laboratory apparatus and demonstrator equipment maintenance.

It will also cover the student's mobility between academic institutions and to the industrial partners for their practice. Furthermore, the objective is to create a synergetic partnership with maritime industry key players for co-financing the education program. Their return on invest will be clearly visible and measurable in terms of highly qualified students and professionals that can significantly help drive digitalization forward in the maritime sector.

Acknowledgements. This project is a highly collaborative endeavour requiring intense contributions of a huge number of individuals. We regret that we cannot cite all their names here and want to express our thanks to them in this way. Some of them, however, gave particularly valuable input to the work programme published in this article without being cited as co-authors. Special thanks is therefore due to Javier Prieto-Tejedor, Sara Rodríguez-González, Ana B. Gil-González from University of Salamanca; Aleksander Grm from University of Ljubljana; Theodoros Kosmanis, Dimitrios Tziourtzioumis, George Minos, Kalliopi Kravari, Fotis Stergiopoulos from International Hellenic University; Panagiotis Maroulas from Cerca Trova Ltd, Mariana Golumbeanu from National Institute for Marine Research and Development "Grigore Antipa"; Kinga Lazuga, Marcin Maka, Mariusz Dramski, Maciej Gucma from Maritime University of Szczecin; Kadi Kasepõld, Raivo Sell and Priit Ruberg from Tallinn University of Technology; Olga Delgado Ortega, Miguel Silva from Infante D. Henrique Nautical School; Michael Reiner from ECQA GmbH; Stefanos Chatzinikolaou from RINA Hellas S.A.; Olivier le Brun, Olivier Herlory, Elina Delord from Creocean; Fotios Oikonomou, Panagiota Arampatzi from Danaos Shipping Company.

The SMARTSEA project [19] is financially supported by the European Commission in the Erasmus+ Programme under the project number 612198-EPP-1-2019-1-ES-EPPKA2-KA. This publication reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

References

1. Accenture: Hyundai heavy industries and accenture to build connected smart ships (2015). https://newsroom.accenture.com/news/hyundai-heavy-industries-and-accenture-to-build-connected-smart-ships.htm. Accessed 12 Apr 2020
2. Harvard Business School: Maersk – reinventing the shipping industry using IoT and blockchain (2018). https://digital.hbs.edu/industry-4-0/maersk-reinventing-shipping-industry-using-iot-blockchain. Accessed 12 Apr 2020
3. List of maritime colleges: https://en.wikipedia.org/wiki/List_of_maritime_colleges#Europe. Accessed 12 Apr 2020
4. McKinsey: Unlocking the potential of the Internet of Things (2015). https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/the-internet-of-things-the-value-of-digitizing-the-physical-world. Accessed 12 Apr 2020
5. International Institute of Marine Surveying (iims.org.uk)
6. Lloyds maritime academy. https://lloydsmaritimeacademy.com/event/marine-surveying-distance-learning-course. Accessed 25 Mar 2020
7. The application of technology in marine surveying. The report No. 74, pp. 44 (2015). https://iims.org.uk/wp-content/uploads/2014/03/The-Report-December-2015.pdf
8. Kao, C.C., et al.: A comprehensive study on the internet of underwater things: applications, challenges, and channel models. **17**(7) (2017). https://ncbi.nlm.nih.gov/pubmed/28640220. Accessed 03 Oct 2019

9. IIMS Baltimore 2017 conference proceedings (2017). https://iims.org.uk/media/marine-surveying-videos/iims-baltimore-2017-conference. Accessed 12 Apr 2020
10. Inmarsat: Industrial IoT maritime (2018). https://inmarsat.com/press-release/inmarsat-research-programme-offers-unique-maritime-iot-insights. Accessed 12 Apr 2020
11. European Commission Press Release: EU leads the way with ambitious action for cleaner and safer seas (2017). https://europa.eu/rapid/press-release_MEMO-17-3588_en.htm. Accessed 12 Apr 2020
12. Horizon2020: Smart, green and integrated transport (2020). https://ec.europa.eu/programmes/horizon2020/en/h2020-section/smart-green-and-integrated-transport. Accessed 12 Apr 2020
13. O*Net Online: Maritime job descriptions. https://onetonline.org. Accessed 12 Apr 2020
14. World Maritime University. ICT for Maritime Education and Training (2018). https://wmu.se/professional/ict-maritime-education-and-training. Accessed 19 Jan 2019
15. Kosmanis, T., et al.: Building a cost-effective research ROUV (2017). https://researchgate.net/project/Building-a-cost-effective-research-ROUV-Remotely-Operated-Underwater-Vehicle. Accessed 12 Apr 2020
16. Columbia University: Marine augmented reality systems. https://nextbigfuture.com/2009/10/augmented-reality-helps-marine.html. Accessed 12 Apr 2020
17. Geo Oceans: Marine Survey Solutions. https://geooceans.com. Accessed 12 Apr 2020
18. Teng, L., et al.: Automated water quality survey and evaluation using an iot platform with mobile sensor nodes. Sensors **17**, 1735–1748 (2017)
19. SMARTSEA – MSc on smart maritime & surveying systems (2019). http://smart-sea.eu. Accessed 16 Apr 2020

Is Artificial Intelligence Ready for Standardization?

Thomas Zielke(✉)

Düsseldorf University of Applied Sciences, Düsseldorf, Germany
thomas.zielke@hs-duesseldorf.de

Abstract. Many standards development organizations worldwide work on norms for Artificial Intelligence (AI) technologies and AI related processes. At the same time, many governments and companies massively invest in research on AI. It may be asked if AI research has already produced mature technologies and if this field is ready for standardization. This article looks at today's situation of AI in the context of needs for standardization. The International Organization for Standardization (ISO) runs a standardization project on AI since 2018. We give an up-to-date overview of the status of this work. While a fully comprehensive survey is not the objective, we describe a number of important aspects of the standardization work in AI. In addition, concrete examples for possible items of AI standards are described and discussed. From a scientific point of view, there are many open research questions that make AI standardization appear to be premature. However, our analysis shows that there is a sound basis for starting to work on AI standardization as being undertaken by ISO and other organizations.

Keywords: Artificial Intelligence · Standardization · ISO SC 42 · Trustworthiness · AI robustness · Machine Learning

1 Introduction

The research field *Artificial Intelligence* (AI) can be traced back to the 1950s. The famous monograph "Computing Machinery and Intelligence" by Alan Turing [60], the Dartmouth Summer Research Project of 1956 [37], and the invention of the *Perceptron* [43] are distinctive examples of inspirations for a new research field that today seems omnipresent. Only in resent years it can be claimed that AI as a technology fulfills many expectations that had already been stated more than 60 years ago. However, the term *artificial intelligence* is still being controversially debated and to date no common understanding exists as to methods and technologies that make a system or a software solution *intelligent* [56].

AI as a research field has not a history of steady progress. The field regularly experienced "AI winters", *stages when technology, business, and the media get out of their warm and comfortable bubble, cool down, temper their sci-fi speculations and unreasonable hypes, and come to terms with what AI can or cannot*

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 259–274, 2020.
https://doi.org/10.1007/978-3-030-56441-4_19

really do as a technology [21]. Currently we observe the opposite of an AI winter and this time, even the automotive industry seems to be convinced of the technological and commercial potentials of AI [27]. However, existing standards for the regulation of functional safety, in particular ISO 26262, are not compatible with typical AI methods, e.g. methods for machine learning [45]. This is no surprise as a common characteristic of AI systems is dealing with uncertain data in a context that bears uncertainty, while the results may be associated with some degree of uncertainty too [24].

Despite its long history, its impressive progress in recent years, and many existing real-world applications, AI still is an emergent technology. The general economic benefits from international standards are well investigated [10]. In emerging technologies, the benefits of standardization are less obvious and there is a risk of hindering innovation by inflexible and/or quickly outdated standards. There are partially conflicting interests of the stakeholders. Startups want to fully exploit their technological head start, rapidly create new products, and gain market shares. Established companies need new standards for investment decisions, as best practice guidelines for their development departments, and as an orientation for hesitant customers. Researchers have their own culture of defining the state of the art and sometimes regard early industry standards on their subject of research as a restriction of freedom. Policymakers ask for respective technical standards when regulations become an issue. For the standards development organizations, like ISO and DIN (German Institute for Standardization) for example, initiatives for new standards are a core business in accordance with their mission, also securing their financing.

In contrast to the scientific discourse, standardization seeks consensus. This is actually a strong reason why the work on standards has a positive impact on emerging technologies. Standards establish common vocabularies and agreed definitions of terms. Standards also contribute to a more effective dissemination of innovation and they increase the confidence of investors and customers [40].

2 Objectives and Context of This Research

This article looks at current work on the creation of international standards for AI. In 2018, the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) started a project on AI standardization by founding the subcommittee ISO/IEC JTC 1/SC 42 Artificial intelligence[1]. The author is a founding member of the interdisciplinary DIN Working Committee "Artificial Intelligence" [15] which represents Germany in the ISO/IEC JTC 1/SC 42 . He is also an active member of several SC 42 working groups.

Although the foundation of the SC 42 and many associated national committees seems to indicate that AI is ready for standardization, it can be argued that past attempts at AI standardization were unsuccessful and that AI technology

[1] https://www.iso.org/committee/6794475.html, https://www.iec.ch/dyn/www/f?p=103:7:0::::FSP_ORG_ID:21538.

still lacks the level of trust needed for widely agreed standards [31]. What is different today, compared to the situation twenty-five years ago when the first efforts to create ISO standards for AI were made (see e.g. [47])? It may also be asked if a technology is ready for standardization at a stage of development where massive investments in research are needed and actually being announced by many governments. The European Union alone wants to spend €20 billion per year by the end of 2020 [18].

Besides analyzing the situation with respect to questions like the ones given above, this article provides firsthand information on the current international activities on AI standardization. It is not the intention to provide a comprehensive overview of the standardization work of the SC 42 , nor would this be possible within the scope of a conference paper. However, the general goals of the main working groups are briefly described. In addition, concrete examples are given for topics that are likely to be covered by future standards in AI.

One objective of this survey on AI standardization work is to prepare the ground for answering the question in the title "Is Artificial Intelligence Ready for Standardization?". There may be several different valid answers to this question depending on the expectations for the standardization outcomes. Therefore this article also investigates on some crucial technical issues that differentiates AI standardization from other standards.

AI receives more public and political attention than most other technologies because it is expected to have an impact on everyone's life in the long run. Floridi et al. [22] put it this way: *AI is not another utility that needs to be regulated once it is mature. It is a powerful force, a new form of smart agency, which is already reshaping our lives, our interactions, and our environments.*

This has consequences for the standardization work in AI. Even more than in other areas of information and communication technology, the compatibility of technology and the values of a democratic society has to be taken into account [32], at least from a European perspective. This article focusses on the technical aspects related to AI standardization. The reader should be aware that ethical and societal concerns are an important part of the SC 42 work too.

3 ISO/IEC JTC 1/SC 42 Artificial Intelligence

In November 2017, the Technical Management Board (TMB) of ISO decided that the Joint Technical Committee "Information Technology" (JTC 1) should found a subcommittee (SC) on Artificial Intelligence. The inaugural plenary meeting of the new SC 42 took place in Beijing, China, in April 2018. The scope of work of SC 42 is "Standardization in the area of Artificial Intelligence", specifically:

- Serve as the focus and proponent for JTC 1's standardization program on Artificial Intelligence
- Provide guidance to JTC 1, IEC, and ISO committees developing Artificial Intelligence applications

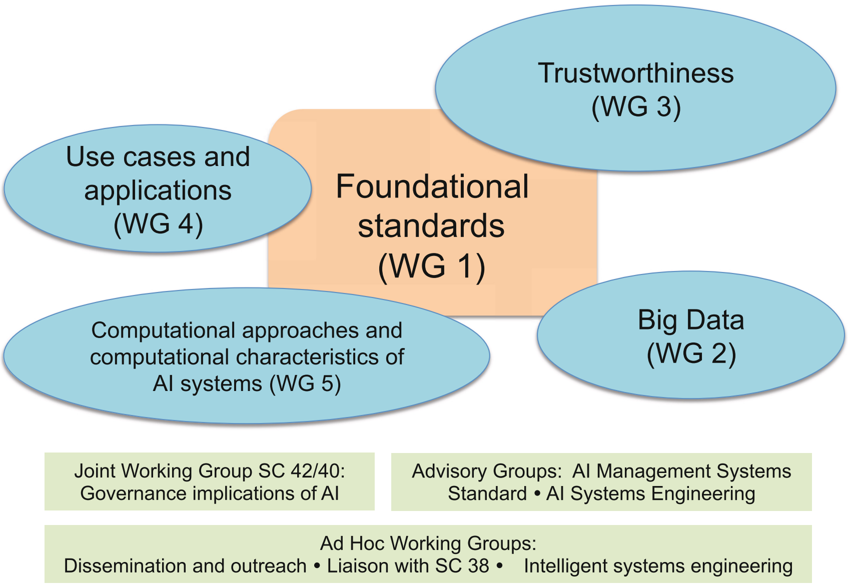


Fig. 1. Structure of the SC 42 as of March 2020. The illustration shows the main working groups (WG). There are also a joint working group (JWG) with SC 40 "IT Service Management and IT Governance", two advisory groups (AG), and three ad hoc working groups (AHG). SC 38 is on "Cloud Computing and Distributed Platforms".

Originally JTC 1 recommended that SC 42 should cover the main topics foundational standards, computational methods, trustworthiness, and societal concerns. The structure of the SC 42 as of March 2020 is shown in Fig. 1. The main working groups are on foundational standards (WG 1), trustworthiness (WG 3), use cases and applications (WG 4), computational approaches and computational characteristics of AI systems (WG 5), and big data (WG 2), which used to be covered by a separate working group under JTC 1. Societal concerns has become a subtopic of WG 3.

3.1 Foundational Standards

A basic objective of standardization is the definition of common terms. When looking at terms relating to AI, the term artificial intelligence itself is a primary subject of discussion. The Merriam-Webster dictionary offers these definitions: 1) a branch of computer science dealing with the simulation of intelligent behavior in computers 2) the capability of a machine to imitate intelligent human behavior[2]. From a technical point of view there are two problems with definitions like that. Firstly, it does not explain what "intelligent" is. Secondly, it refers to capabilities of humans that are neither defined nor objectively measurable.

[2] https://www.merriam-webster.com/dictionary/artificial%20intelligence.

A useful reflection on definitions of AI can be found in [56]. WG 1 attempts to find a workable definition by consensus. Although the concrete wording of the AI definition may not be highly crucial for the quality of the future SC 42 standards, there is a definite need for an AI definition in industry.

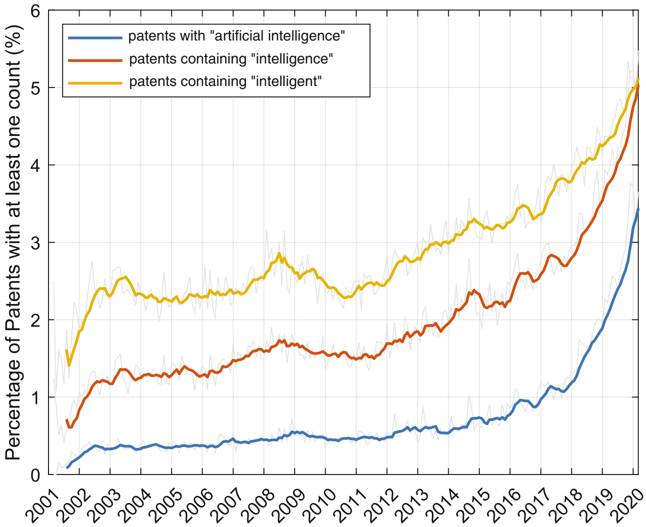


Fig. 2. Relative monthly numbers of US patent applications that contain the term "intelligent", "intelligence", or "artificial intelligence" respectively. The data analyzed cover all applications since 2001. The colored curves show moving averages taken over periods of 6 months. The grey curves show the monthly raw values.

In recent years a steep increase in the usage of the term artificial intelligence can be observed, in the media, in research work, in marketing material, and in industrial publications. As an example, we looked at the US patent applications since 2001. All text is available from the US patent office[3]. We counted all patent applications that mention the terms "intelligent", "intelligence", or "artificial intelligence" respectively at least once. The graphs in Fig. 2 show the respective monthly percentages of all patent applications in the period between March 2001 and March 2020. There is a remarkable exponential increase in recent years.

More than the wording of an AI definition, the description of the concepts, methods, and best practices for AI are important. The chapter titles in the 1982 textbook "Principles of Artificial Intelligence" by Nils J. Nilson [38] contain the following terms: production systems, search strategies, predicate calculus, resolution refutation, plan-generation systems, structured object representation. As far as these topics are still regarded AI, they belong to a category called *Symbolic AI* [55]. In symbolic AI, goals, beliefs, knowledge, and so on, as well as their

[3] https://bulkdata.uspto.gov.

interrelationships, are all formalized as symbolic structures. The report "Artificial Intelligence Concepts and Terminology" by WG 1 [48], mentions symbolic AI briefly. It is referred to as "classical AI". Forty years ago, classical AI was AI mainstream, focussing on very different methods and techniques than todays's AI which is dominated by machine learning. Modern AI predominantly is *Connectionist AI*, a term coined in the 1980s [19]. In the technical literature of the last two decades it has not much been used any more. Therefore Flasinski [20] is probably right in categorizing connectionist AI under *Computational Intelligence* (CI). He states the following common features of CI methods:

– numeric information is basic in a knowledge representation,
– knowledge processing is based mainly on numeric computation,
– *usually* knowledge is not represented in an *explicit* way.

Cognitive Computing (CC) is another relevant term here. It is sometimes used interchangeable with AI and CI. CC provides an interesting example of possible conflicts when defining a terminology in standardization. The term has been taken over by IBM (www.ibm.com) as an umbrella term for the marketing of all their products and services that somehow use AI technologies [63]. Originally, CC was meant as a notion for engineering complex intelligent systems that incorporate many features and technologies of AI [44]. The definition given by [48] covers technologies that uses natural language processing and machine learning to enable people and machines to interact more naturally to extend and magnify human expertise and cognition.

Machine learning (ML) and related topics are the current main focus of the SC 42 standardization work, specifically of the work of WG 1 on foundational standards. We deal with ML in the following subsection. The overview report [48] on foundational standards is structured by the following topics: functional view of AI systems, applications of AI, AI ecosystems, AI concepts, and AI systems categories.

Machine Learning (ML). Early AI was *knowledge-based* [28]. Today's AI is *data-driven* [64]. ML is the discipline that provides models and methods for the transformation of data into task-specific knowledge [42]. Most of the success of AI in recent years can be attributed to ML. Face recognition is an example for a prominent AI task that has a long research history. The best recognition rates on a popular benchmark test went up from ca. 70% when using methods without ML to more than 97% when *deep learning* was applied [62]. The report "Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML)" by WG 1 [49] intends to establish a framework for describing a generic AI system using ML technology. The framework describes the system components and their functions in the AI ecosystem. Under this scope, the report deals with the ML terminology, subsystems, approaches and training data, pipeline, and the ML process.

As an example of the work on the definition and classification of ML approaches, concepts, and methods, we briefly look at the taxonomy for ML

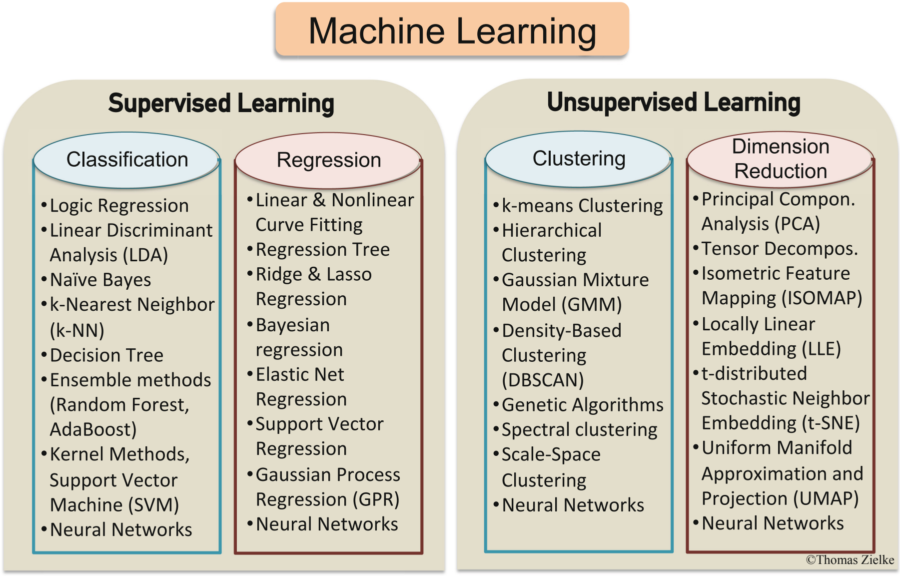


Fig. 3. Taxonomy of ML methods inspired by [36] and matched with many other ML resources. This may not be the taxonomy that WG 1 eventually adopts. The categories *supervised* and *unsupervised* are generally accepted. There is also the category of *semi-supervised* methods. Not all of the established methods could be listed in the diagram.

methods. Figure 3 shows four different groups of ML methods. The main categories are *supervised* and *unsupervised*, i.e. methods that need labelled data for training and methods that work with unlabelled data [6]. Not shown in the figure is *reinforcement learning*, a third major category of ML methods. In reinforcement learning there are model-free and model-based methods, where model refers to a possible predictive model of an unknown environment that a learning agent interacts with. Learning works through trial-and-error interactions [33] or by maximizing a numerical reward signal in doing so [58]. To be mentioned here are also *semi-supervised* methods [57], *ensemble learning* [16] and *transfer learning* [35]. The field of ML research has produced much more methods than could be listed in the figure, which shows a selection that reflects the popularity and the distinctiveness of the respective methods. Figure 3 also classifies the methods according to their respective suitability for specific tasks: classification, regression, clustering, dimension reduction. One may argue that anomaly detection is an important additional task category. However, it may also come under classification or clustering. For all categories in Fig. 3 there are methods based on *artificial neural networks* (ANN). ANNs, in particular *deep neural networks* (DNNs) provide a generic architecture for the data-driven approach to AI.

3.2 Working Groups 2–5

Big Data. A few years ago, JTC 1 established a program of work on "big data" through its working group WG 9. This work has been transferred to SC 42 and assigned to WG 2. Due to the history of big data within JTC 1, WG 2 is the only working group of SC 42 that already has published ISO standards, e.g. [4].

Trustworthiness. WG 3 on trustworthiness has the following main tasks: a) investigate approaches to establish trust in AI systems through transparency, verifiability, explainability, controllability b) investigate engineering pitfalls and assess typical associated threats and risks to AI systems with their mitigation techniques and methods c) investigate approaches to achieve AI systems' robustness, resiliency, reliability, accuracy, safety, security, privacy.

Trustworthiness may be defined as the degree to which a user or other stakeholder has confidence that a product or system will behave as intended. From a perspective that not only considers technical aspects, trustworthiness can be described following [53]:

- Ability is the capability the AI system to do a specific task (robustness, safety, reliability, etc.).
- Integrity may be viewed as the insurance that information will not be manipulated in a malicious way by the AI system (completeness, accuracy, certainty, consistency, etc.).
- Benevolence is the extent to which the AI system is believed to do good, or in other terms, to what extent the "Do No Harm" principle is respected.

The first publications of WG 3 are the reports"Overview of trustworthiness in artificial intelligence" [53], "Assessment of the robustness of neural networks" [50], "Bias in AI systems and AI aided decision making" [51], and "Overview of ethical and societal concerns" [52]. Robustness is a topic of particular high concern. Section 4 deals with that in more detail.

Use Cases and Applications. WG 4 has the following main tasks: a) identify different AI application domains and the different context of their use b) describe applications and use cases using the terminology and concepts defined in ISO/IEC AWI 22989 and ISO/IEC AWI 23053 and extend the terms as necessary c) collect and identify societal concerns related to the collected use cases.

The first publication of WG 4 is the report "Use cases and applications" [54].

Computational Approaches and Computational Characteristics of AI Systems. The initial task of WG 5 has been to develop a technical report with the title "Overview of computational approaches for AI systems". Its scope is the state of the art of computational approaches for AI systems, describing: a) main computational characteristics of AI systems b) main algorithms and approaches used in AI systems, referencing use cases contained in the report of WG 4 [54].

4 Robustness of AI

The integration of AI components into products and industrial processes is currently limited to industries that do not have requirements for rigorous software verification. Software verification is an essential part of many industrial processes. The objective is to ensure both safety and performance of the software in all parts of the system. In some domains, the software verification process is also an important part of system certification, e.g. ISO 26262 in the automotive industry [3]. While many methods exist for validating non-AI systems, they are mostly not directly applicable to AI systems, and neural networks in particular.

The problem is widely referred to as robustness of AI. Robustness is used as a general term for describing properties that are required for the acceptance of new high-stakes AI applications [14]. Many recent publications on this problem deal with so-called adversarial examples that cause malfunctions of a deep neural network model although the respective input patterns are very similar to valid data [7]. In practice, robustness has to be defined goal-oriented in the context of the respective application domain. Typical examples of robustness goals in machine learning applications are:

– Adherence to certain thresholds on a set of statistical metrics that need to hold on the validation data.
– Invariance of the functional performance w.r.t. certain types of data perturbations [29].
– Invariance of the functional performance w.r.t. systematic variations in the input data, e.g. measurement drifts [23] or operating conditions [17].
– Stability of training outcomes under small variations of the training data and with different training runs under stochastic influences [8].
– Consistency of the model output for similar input data (resistance to adversarial examples) [7].

Traditional machine learning models with few parameters (shallow models) are better suited to meet robustness goals than complex (deep) models. But DNNs are heavily contributing to the success of AI. Deep models can be effectively trained while yielding superior generalization [9]. The complexity of deep models poses a risk in terms of robustness. Standardization is a way to manage that risk and to enable industry to use deep models without compromising on safety or other aspects related to robustness. The report "Assessment of the Robustness of Neural Networks" by WG 3 [50] suggests that the state-of-the-art in statistical and empirical methods for the assessment of robustness is sufficient for the development of standards. There are also formal methods for the assessment of robustness which potentially are most suitable for safety critical systems.

Neural network architectures, in particular DNN, represent a specific challenge as they are both hard to explain and sometimes have unexpected behavior due to their nonlinear nature and the large number of model parameters. For this, formal methods for the verification of neural networks do not play a significant role in practice yet, as stated by Guidotti [26]: *In spite of the extensive research* *done on NNs verification the state-of-the-art methods and tools are still far from being able to successfully verify the corresponding state-of-the-art NNs.* Because of the potential importance of formal methods for robustness verification, WG 3 will work on this topic with a dedicated project in the future.

4.1 Example for an Empirical Approach to Testing: Field Trials

Establishing trust in AI systems requires many aspects to be studied, but the number of feasible approaches for analyzing a black box system's behavior and performance are limited. AI systems consist of software to a large extent. Much of it is not a black box and therefore software testing standards can be applied. ISO 29119 [2] describes the primary goals of software tests: *Provide information about the quality of the test item and any residual risk in relation to how much the test item has been tested; to find defects in the test item prior to its release for use; and to mitigate the risks to the stakeholders of poor product quality.* These goals are very difficult to achieve for all parts of a typical AI system, and only as of late, testing of AI systems is being researched on [39]. While AI is being considered as a tool for Software Process Improvement (SPI), e.g. as a support tool for software test management [41], new approaches will have to be developed to test AI software itself and software testers need new tools and practices for validating that an AI software complies with the requirements [13].

Defects and poor product quality are concerns when testing AI systems as much as with conventional systems. However, the failure of an AI system in a functional test may not be related to a "software bug" or an erroneous design, AI systems showing occasional malfunctions may be regarded useful for their intended purpose, and the efficacy of an AI system may not be measurable by conventional approaches to software testing. Another fundamental difference between many AI systems and conventional systems is, that the latter are developed, produced, and quality controlled to strictly meet certain specifications. AI systems, in contrast, may reveal their degree of efficacy during deployment only, as is the case with systems like Amazon's Alexa and Apple's HomePod, for example. This often applies to AI systems that operate in interaction with or dependency of natural environments and humans.

How to deal with the uncertainty of a product's efficacy and the risks of its deployment is subject to regulations in the medical domain. Medical AI systems have to comply with DIN/EN/ISO 14155 [1]. They undergo "clinical investigations", a procedure that resembles "clinical trials" [46]. For non-medical AI systems, field trials have for long been a recognized means of comparing and proving the performance of solutions. Some prominent examples are: facial recognition trials [11], tests of decision support systems [12], testing driverless cars [61], and tests of speech and voice recognition systems [34].

Currently, field trials for AI systems greatly vary w.r.t. methodology, number of users or use samples involved, status of the responsible organization/persons, and documentation of the results. A good practice guideline for field trials is a concrete example for a desirable AI standard. In analogy to clinical investigations of medical devices, it should specify general requirements intended to

– protect the rights, safety and well-being of human participants,
– ensure the scientific conduct of the field trial and the credibility of the results,
– define the responsibilities of the sponsor and principal investigator, and
– assist sponsors, investigators, regulatory authorities and other bodies involved in the conformity assessment of AI systems.

5 Discussion

The results of this work can be structured according to the four basic functions of technology standards described by Tassey [59]:

– Quality/Reliability
– Information
– Compatibility/Interoperability
– Variety Reduction

The current work on AI standards mainly addresses **quality/reliability** and **information**. The WG 1 report [48] shows that there is a rich set of terms and definitions that are specific for AI technologies and applications. Section 3.1 describes examples for definitions and taxonomy. The WG 3 report [50] is a promising basis for the development of standards for measuring the quality and reliability of AI systems. WG 3 also has a project on AI risk management, aiming at a standard that could pave the way for certification processes. Section 4 describes concrete examples for dealing with quality and reliability in AI.

Compatibility/interoperability has not much been in the focus of AI standardization yet. Progress in this area is mainly be driven by the open source community. The ONNX initiative (Open Neural Network Exchange), for example, tries to create an open standard for machine learning interoperability. NNEF (Neural Network Exchange Format) is another example[4]. However, these exchange formats do not yet address features such as scalable and incremental updates and compression. There is a standard under development on compression of neural networks for multimedia content description and analysis [5]. The responsible SC 29 has a liaison with SC 42 for a joined project on neural network compression. **Variety reduction**, the fourth function of technology standards, may not be a realistic goal in the foreseeable future, mainly because of the vitality of AI research. It is important though that developers in industry get some guidance on the choice of models, methods, and algorithms in AI.

For many experts, the work on standards for AI is not only about the four functions or objectives discussed above. AI technologies have the potential of reshaping our lives, our interactions, and our environments [22]. There is the expectation that international AI standards also address ethical and societal issues. The way this can be done is limited by the nature of international technical standards: Any bias toward value-sets that are specific for certain cultures or countries has to be avoided. However, there is an official ISO document named

[4] https://onnx.ai, https://www.khronos.org/nnef.

"Guidance on social responsibility" [52] which *is not intended to be interpreted as an "international standard", "guideline" or "recommendation"*. SC 42 WG 3 is going to publish a document on"ethical and societal concerns" [52].

6 Conclusion

This article describes results and related background information from research into the current state of the international standardization of artificial intelligence. It is an up-to-date overview of the current work at the International Organization for Standardization (ISO) on the development of standards for AI. Exemplarily, several important topics of AI standardization are elaborated on in detail, e.g. the definition of the terms, the taxonomy of machine learning, and the assessment of robustness of AI systems. Observing an exponential increase in the usage of the term artificial intelligence in the patent literature, it can be concluded that the market for technical solutions based on AI is no longer a niche. Consequently, technical standards for AI are needed. Given the long development time needed for ISO standards, it seems acceptable that certain important topics, e.g. assessment of trustworthiness of AI, still have a shallow scientific basis. The motivating question for this article "Is Artificial Intelligence Ready for Standardization?" may best be answered by "AI standardization is ready for takeoff".

It has been shown in this article that certain concepts, definitions, methods, and procedures of AI are ready for standardization. For several essential topics, the scientific basis is not sufficiently solid yet. Future AI research has to address a number of crucial application requirements for which standards currently could not describe a solution. The following examples are given with literature references that provide an understanding of the relevant issues.

- Formal methods for the verification of deep neural networks or for the assessment of their robustness [30].
- Architectures and training methods for robust solutions based on deep neural networks [8].
- Methods and tools for generating comprehensible explanations for AI-based decision processes [25].

Acknowledgements. We would like to thank Dominic Kalkbrenner and Jens Lippel for programming the data analytics on the bulk data from the US patent office. Dr. Andreas Riel provided valuable feedback and discussions on the structure of the paper and the relevance for the automotive industry.

References

1. ISO 14155:2011: clinical investigation of medical devices for human subjects - good clinical practice (2011)
2. ISO/IEC/IEEE 29119–1:2013: software and systems engineering - software testing - part 1:concepts and definitions (2013)

3. ISO 26262–1:2018: road vehicles - functional safety - part 1: vocabulary (2018)
4. ISO/IEC 20546:2019: information technology - big data - overview and vocabulary (2019)
5. ISO/IEC WD 15938–17: multimedia content description interface - part 17: compression of neural networks for multimedia content description and analysis (2020)
6. Ang, J.C., Mirzal, A., Haron, H., Hamed, H.N.A.: Supervised, unsupervised, and semi-supervised feature selection: a review on gene selection. IEEE/ACM Trans. Comput. Biol. Bioinf. **13**(5), 971–989 (2016)
7. Bastani, O., Ioannou, Y., Lampropoulos, L., Vytiniotis, D., Nori, A.V., Criminisi, A.: Measuring neural net robustness with constraints. In: Proceedings of 30th International Conference on on Neural Information Processing Systems, pp. 2621–2629. Curran Associates Inc. (2016)
8. Becker, M., Lippel, J., Stuhlsatz, A., Zielke, T.: Robust dimensionality reduction for data visualization with deep neural networks. Graph. Models **108**, 101060 (2020). https://doi.org/10.1016/j.gmod.2020.101060
9. Belkin, M., Hsu, D., Ma, S., Mandal, S.: Reconciling modern machine-learning practice and the classical bias-variance trade-off. Proc. Natl. Acad. Sci. **116**(32), 15849–15854 (2019)
10. Blind, K., Jungmittag, A., Mangelsdorf, A.: The economic benefits of standardisation. An update of the study carried out by DIN in 2000. DIN Berlin, January 2012
11. BSI: An investigation into the performance of facial recognition systems relative to their planned use in photo identification documents - BioP I. Technical report, Bundesamt für Sicherheit in der Informationstechnik (BSI), Bundeskriminalamt (BKA), secunet AG, April 2004. https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Publications/Studies/BioP/BioPfinalreport_pdf.pdf
12. Burke, J., Dunne, B.: Field testing of six decision support systems for scheduling fungicide applications to control mycosphaerella graminicola on winter wheat crops in Ireland. J. Agric. Sci. **146**(04), 415 (2008)
13. Cunningham, S., et al.: Software testing: a changing career. In: Walker, A., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 731–742. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_57
14. Dietterich, T.G.: Steps toward robust artificial intelligence. AI Mag. **38**(3), 3–24 (2017)
15. DIN: Interdisciplinary DIN working committee "artificial intelligence" (2018). https://www.din.de/en/innovation-and-research/artificial-intelligence/ai-working-committee
16. Dong, X., Yu, Z., Cao, W., Shi, Y., Ma, Q.: A survey on ensemble learning. Front. Comput. Sci. **14**(2), 241–258 (2019). https://doi.org/10.1007/s11704-019-8208-z
17. Duthon, P., Bernardin, F., Chausse, F., Colomb, M.: Benchmark for the robustness of image features in rainy conditions. Mach. Vis. Appl. **29**(5), 915–927 (2018). https://doi.org/10.1007/s00138-018-0945-8
18. EU: Funding for AI. https://ec.europa.eu/info/research-and-innovation/research-area/industrial-research-and-innovation/key-enabling-technologies/artificial-intelligence-ai_en
19. Feldman, J.A., Ballard, D.H.: Connectionist models and their properties. Cogn. Sci. **6**(3), 205–254 (1982). https://doi.org/10.1207/s15516709cog0603_1
20. Flasinski, M.: Introduction to Artificial Intelligence. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-319-40022-8
21. Floridi, L.: AI and its new winter: from myths to realities. Philos. Technol. **33**(1), 1–3 (2020). https://doi.org/10.1007/s13347-020-00396-6

22. Floridi, L., et al.: AI4people–an ethical framework for a good AI society: opportunities, risks, principles, and recommendations. Minds Mach. **28**(4), 689–707 (2018)
23. Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., Bouchachia, A.: A survey on concept drift adaptation. ACM Comput. Surv. **46**(4), 1–37 (2014)
24. Ghahramani, Z.: Probabilistic machine learning and artificial intelligence. Nature **521**(7553), 452–459 (2015). https://doi.org/10.1038/nature14541
25. Goebel, R., et al.: Explainable AI: the new 42? In: Holzinger, A., Kieseberg, P., Tjoa, A.M., Weippl, E. (eds.) CD-MAKE 2018. LNCS, vol. 11015, pp. 295–303. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-99740-7_21
26. Guidotti, D.: Enhancing neural networks through formal verification. In: Alviano, M., Greco, G., Maratea, M., Scarcello, F. (eds.) Discussion and Doctoral Consortium papers of AI*IA 2019–18th International Conference of the Italian Association for Artificial Intelligence, Rende, Italy, 19–22 November 2019. CEUR Workshop Proceedings, vol. 2495, pp. 107–112. CEUR-WS.org (2019)
27. Hatani, F.: Artificial intelligence in Japan: policy, prospects, and obstacles in the automotive industry. In: Khare, A., Ishikura, H., Baber, W.W. (eds.) Transforming Japanese Business. FBF, pp. 211–226. Springer, Singapore (2020). https://doi.org/10.1007/978-981-15-0327-6_15
28. Hayes-Roth, F., Jacobstein, N.: The state of knowledge-based systems. Commun. ACM **37**(3), 26–39 (1994)
29. Hendrycks, D., Dietterich, T.G.: Benchmarking neural network robustness to common corruptions and perturbations. In: 7th International Conference on Learning Representations, ICLR 2019. OpenReview.net, New Orleans, May 2019
30. Huang, X., et al.: A survey of safety and trustworthiness of deep neural networks. arXiv preprint arXiv:1812.08342 (2018)
31. Hurlburt, G.: How much to trust artificial intelligence? IT Prof. **19**(4), 7–11 (2017). https://doi.org/10.1109/MITP.2017.3051326
32. Iversen, E.J., Vedel, T., Werle, R.: Standardization and the democratic design of information and communication technology. Knowl. Technol. Policy **17**(2), 104–126 (2004). https://doi.org/10.1007/s12130-004-1027-y
33. Kaelbling, L.P., Littman, M.L., Moore, A.W.: Reinforcement learning: a survey. J. Artif. Intell. Res. **4**, 237–285 (1996)
34. Lamel, L., et al.: Field trials of a telephone service for rail travel information. In: Proceedings of of IVTTA 1996. Workshop on Interactive Voice Technology for Telecommunications Applications, pp. 111–116. IEEE (1996)
35. Liang, H., Fu, W., Yi, F.: A survey of recent advances in transfer learning. In: 19th IEEE International Conference on Communication Technology, ICCT 2019. pp. 1516–1523. IEEE, Xi'an, October 2019
36. Louridas, P., Ebert, C.: Machine learning. IEEE Softw. **33**(5), 110–115 (2016)
37. Moor, J.: The dartmouth college artificial intelligence conference: the next fifty years. AI Mag. **27**(4), 87–87 (2006)
38. Nilsson, N.J.: Principles of Artificial Intelligence. Symbolic Computation. Springer, Heidelberg (1982)
39. Numan, G.: Testing artificial intelligence. In: Goericke, S., et al. (eds.) The Future of Software Quality Assurance, pp. 123–136. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-29509-7_10
40. O'Sullivan, E., Brévignon-Dodin, L.: Role of standardisation in support of emerging technologies. Technical report, Institute for Manufacturing, University of Cambridge, June 2012

41. Poth, A., Beck, Q., Riel, A.: Artificial intelligence helps making quality assurance processes leaner. In: Walker, A., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 722–730. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_56
42. Rao, V.R.: How data becomes knowledge, part 1: from data to knowledge, March 2018. https://www.ibm.com/developerworks/library/ba-data-becomes-knowledge-1/index.html
43. Rosenblatt, F.: The perceptron: A probabilistic model for information storage and organization in the brain. Psychol. Rev. **65**, 386 (1958)
44. Rozenblit, J.W.: Cognitive computing: principles, architectures, and applications. In: Proceedings of 19th European Conference on Modelling and Simulation (ECMS) (2005)
45. Salay, R., Queiroz, R., Czarnecki, K.: An analysis of ISO 26262: using machine learning safely in automotive software. CoRR abs/1709.02435 (2017)
46. Santos, I.C., Gazelle, G.S., Rocha, L.A., Tavares, J.M.R.: Medical device specificities: opportunities for a dedicated product development methodology. Expert Rev. Med. Devices **9**(3), 299–311 (2012)
47. SC1: ISO/IEC 2382–31:1997(en) information technology - vocabulary - part 31: Artificial intelligence - machine learning (1997)
48. SC42 WG1: Artificial intelligence concepts and terminology. Technical report CD 22989, ISO/IEC JTC 1/SC 42 Artificial Intelligence (2019)
49. SC42 WG1: Framework for artificial intelligence (AI) systems using machine learning (ML). Technical report CD 23053, ISO/IEC JTC 1/SC 42 Artificial Intelligence (2019)
50. SC42 WG3: Assessment of the robustness of neural networks - part 1: overview. Technical report CD TR 24029–1, ISO/IEC JTC 1/SC 42 Artificial Intelligence (2019)
51. SC42 WG3: Bias in AI systems and AI aided decision making. Technical report AWI TR 24027, ISO/IEC JTC 1/SC 42 Artificial Intelligence (2020)
52. SC42 WG3: Overview of ethical and societal concerns. Technical report AWI TR 24368, ISO/IEC JTC 1/SC 42 Artificial Intelligence (2020)
53. SC42 WG3: Overview of trustworthiness in artificial intelligence. Technical report PRF TR 24028, ISO/IEC JTC 1/SC 42 Artificial Intelligence (2020)
54. SC42 WG4: Use cases and applications. Technical report CD TR 24030, ISO/IEC JTC 1/SC 42 Artificial Intelligence (2019)
55. Smolensky, P.: Connectionist AI, symbolic AI, and the brain. Artif. Intell. Rev. **1**(2), 95–109 (1987). https://doi.org/10.1007/BF00130011
56. Stone, P., et al.: Artificial intelligence and life in 2030. Technical report, Stanford University, September 2016
57. Stuhlsatz, A., Lippel, J., Zielke, T.: Feature extraction with deep neural networks by a generalized discriminant analysis. IEEE Trans. Neural Netw. Learn. Syst. **23**(4), 596–608 (2012)
58. Sutton, R.S., Barto, A.G.: Reinforcement Learning: An Introduction, 2nd edn. The MIT Press, Cambridge (2018)
59. Tassey, G.: Standardization in technology-based markets. Res. Policy **29**(4–5), 587–602 (2000). https://doi.org/10.1016/s0048-7333(99)00091-8
60. Turing, A.M.: Computing machinery and intelligence. Mind LIX **59**(236), 433–460 (1950). https://doi.org/10.1093/mind/lix.236.433

61. UK-Government: The pathway to driverless cars: a code of practice for testing. Technical report, Department for Transport, Great Minster House, 33 Horseferry Road, London (2015). https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/446316/pathway-driverless-cars.pdf
62. Wang, M., Deng, W.: Deep face recognition: a survey. CoRR abs/1804.06655 (2018)
63. Weiss, H.: IBM verwettet seine zukunft auf cognitive computing. Computerwoche, October 2015. https://www.computerwoche.de/a/ibm-verwettet-seine-zukunft-auf-cognitive-computing,3218187
64. Yu, B., Kumbier, K.: Artificial intelligence and statistics. Front. Inf. Technol. Electron. Eng. **19**(1), 6–9 (2018)

A Framework for Automated Testing

Thomas Fehlmann[1(✉)] and Eberhard Kranich[2]

[1] Euro Project Office, 8032 Zurich, Switzerland
thomas.fehlmann@e-p-o.com
[2] Euro Project Office, 47051 Duisburg, Germany
info@e-p-o.com

Abstract. Autonomous Real-time Testing requires test automation. Test automation is closely related to Continuous Integration/ Continuous Delivery (CI/CD).

However, test automation is a difficult undertaking. While many tools exist that automate the execution of tests, the generation of tests remains manual even though complex systems require a high number of test cases.

This paper explains how to generate new test cases by recombination, distinguish relevant test cases from redundant test cases and proposes a framework for how to automate test generation and execution.

Keywords: Software metrics · Test metrics · Test coverage · Continuous integration · Continuous delivery · Continuous deployment · Continuous testing · DevOps · COSMIC size measurement

1 Introduction

Continuous Integration/Continuous Delivery (CI/CD), together with *Continuous Deployment* and *DevOps*, have become mainstream in the software industry [1]. CI is the practice of merging all developers' working copies to a shared mainline several times a day. Grady Booch first proposed the term CI in his 1991 method [2], although he did not advocate integrating several times a day. Extreme programming (XP) adopted the concept of CI and did advocate integrating more than once per day – perhaps as many as tens of times per day [3]. Today, CI/CD is supposed to be fully automated by automating steps in the software delivery process, such as initiating code builds, running automated tests, and deploying to a staging or production environment.

Continuous deployment goes a step further than continuous delivery as it includes continuous release and acceptance testing in the delivery pipeline. This allows programmers to deploy their code immediately to users and customers, without waiting for a release date. DevOps in turn addresses the organizational and cultural changes around CI/CD, and especially continuous deployment.

The crucial step in CI/CD is testing. Neither test automation nor creating meaningful test cases is easy, and while automation in test execution is often possible, creating test cases is still a manual task. Moreover, meaningful tests for complex systems that count several thousands of function points involve test sizes that are significantly larger than what can be handled manually.

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 275–288, 2020.
https://doi.org/10.1007/978-3-030-56441-4_20

Functional sizing and test sizing are related. Depending upon the level of security and safety risk exposure, a factor of ten to fifty seems reasonable. For functional size and test size, consult Fehlmann [4, p. 33 ff].

1.1 The SPI Manifesto [5]

Automated testing is a prerequisite for software addressing safety and privacy (Value B). Today's software development lifecycles make testing explicit, involving the user, or customer, in the process. Otherwise, software is not suitable for critical applications.

2 A Testing Tool Overview

2.1 Tools for Continuous Integration/Continuous Delivery

One of the standard tools used in software development is the open-source software Jenkins. The default interaction model with Jenkins, historically, has been web UI driven, requiring users to manually create jobs, then manually fill in the details through a web browser. This requires effort to create and manage jobs to test and build multiple projects, it also keeps the configuration of a job to build/test/deploy separate from the actual code being built/tested/deployed. This prevents users from applying their existing CI/CD best practices to the job configurations themselves.

The new approach is creating pipelines. Users can implement a project's entire build/test/deploy pipeline in a "Jenkinsfile" and store that alongside their code, treating their pipeline as another piece of code checked into source control [6].

Similar approaches are advertised by Atlassian, the provider of Jira, the software development tool of choice used by agile teams [7].

2.2 Test Automation Tools

Test automation is the use of bespoke software separate from the software under test to control the execution of tests and the comparison of actual outcomes with predicted or expected outcomes. Test automation can automate some repetitive but necessary tasks in a formalized testing process already in place or perform additional testing that would be difficult to do manually. Test automation is critical for continuous delivery and continuous testing. There are two approaches to automatic testing:

- *Graphical user interface testing* is a testing framework that generates user interface events such as keystrokes and mouse clicks and observes the changes that result in the user interface, to validate that the observable behavior of the program is correct.
- *API driven testing* is a testing framework that uses a programming interface to the application to validate the behavior under test. Typically, API driven testing bypasses application user interface altogether, and thus fails to test the latter. It can also be testing public – or sometimes also private – interfaces to classes, modules, or libraries with a variety of input arguments to validate that the results which are returned are correct.

A standard in software development is the use of unit testing frameworks such as, for example, JUnit [8]. It allows the execution of unit tests to determine whether various sections of the code are acting as expected under specific circumstances. Test cases on code level verify that the program runs as expected.

2.3 Tools for Continuous Testing

While it is standard functionality of a CI/CD pipeline to include testing, if possible automated, sometimes just manual, testing is understood as a one-time gateway activity that should happen prior to delivery, or at least before deployment. Shift-left is a buzzword within the testing community, meaning that tests should start as early as possible, should be included in sprints and, ideally, with *Test-Driven Development* (TDD) approach [9], started even before coding. Once a software is delivered, testing is stopped.

This does not reflect today's product deployment strategies. While software that is delivered continuously is supposed to be delivered in well-tested state, the reality is that such software will be updated often, after initial delivery, and will cooperate with other software from other vendors. By the CI/CD delivery pipeline, such software should arrive in the same well-tested state as the initial deployment; however, what if such a software does interact with other services from different suppliers? Or if the software contains *Artificial Intelligence* (AI) and adapts behavior to new learnings?

What is required, is *Autonomous Real-time Testing* (ART). It means that devices, such as smartphones or *Advanced Driving Assistance Systems* (ADAS), should be continuously tested while in operation, to continuously assess a few critical characteristics such as safety, security, and privacy protection levels. This considers not only functionality modified by updates, but also deep learnings that have the ability of changing safety- or security-relevant behavior.

3 Representing Tests

3.1 What Is Software Testing?

Software Testing means the process of defining *Test Stories* (or *Test Scenarios*), each containing *Test Cases*, and execute them with the aim of detecting unexpected behavior. A *Test Case* is a structure consisting of *Test Data* $x_1, x_2, \ldots, x_n$ and a *Test Response* y, where each test data item x_i as well as the test response is an *Assertion*. The assertion describes the state of the program under execution. Formally, a test case is expressed by the following *Arrow Term*:

$$\{x_1, x_2, \ldots, x_n\} \rightarrow y \tag{1}$$

For the origins of arrow terms see Engeler [10]. For a more recent application, how arrow terms define a neural algebra on "how does the brain think?", see Engeler [11].

In our case, the assertions describe the status of the software-intense system under test. A simple assertion describes the value, or value range, of a software variable; it can also describe a certain status of the system, such as listening to some device, waiting for confirmation or executing a database search, or simply identify the starting point for some test case. Related to Six Sigma, the left-hand finite set of an arrow term is referenced as *Controls*, the right-hand singleton is the *Response*.

Assertions use the basic numerical operations between variables and constants such as equality, greater than, or inequality. It is not necessary to combine assertions using logical operations AND, OR, and NOT. The test data sequence within the left-hand side acts as an AND. Instead of an OR, use two arrow terms. NOT is more complicated to substitute by arrow terms: sometimes, negation is immediately available as with equality, sometimes, negated assertions split into two or more positive assertions. The test response y is not necessarily unique; several assertions might become true under identical test data assertions $x_1, x_2, \ldots, x_n$, for instance depending where the system under test is investigated for the test result.

A test case passes successfully if we can execute the software with valid initial test data assertions, and after execution of the test, the assertion y for the test response is also valid. A test story passes if all its test cases pass.

Assertions may include stronger assertions. For instance, the assertion $a = 20$ is more restrictive than $a \leq 20$. Test cases always contain weakest assertions; thus, inequalities or range specification rather than sample numbers; a set of program states that yield equal results rather than an extra assertion per equivalent state.

Tests may execute manually, following a test script. A tester compares the result of an executed test with the assertion y. Automatic tests can replace manual execution and result checking. However, the time needed for executing tests is of essence. *Autonomous Real-time Testing* (ART) means executing automated tests in a limited time frame. For doing this, we need arrow terms as a testing framework.

3.2 A Standard for Representing Test Assertions

Since test cases are possibly something that shall be exchanged between different systems, even from different manufacturers, standardization is needed. If software from different suppliers shall cooperate, standards must be agreed and implemented that enable communication and cooperation. In the IoT and automotive area, such standards exist. For real-time testing, with focus on communication, an international standard for specifying test cases exists: *Testing and Test Control Notation* [12], now in its version 3 (TTCN-3). According Ebner [13], the test notation is useful for automatically generating test cases from UML sequence diagrams, covering the base system. In our context, TTCN-3 is suitable for stating assertions. However, TTCN-3 is much more than simply a framework for stating test assertions such as fixing test data and test responses. It also contains the necessary instructions for test instantiation and test automation.

Thus, using TTCN-3 for test assertions, software tests can be described by a standard that is independent from the programming environment and from the supplier. Tests can be interchanged between different actors related to software testing.

3.3 A Representation for Tests

However, modern software is dynamic. In a learning system, its behavior changes, unpredictbly. Trying to model software by static assertions is missing the dynamic behavior of a system. For this reason, we extend our definition of a *Test Case* to include not only basic assertions but recursively other test cases as well.

Let $\mathcal{L}$ be the set of all assertions over a given domain. Examples include statements about customer's needs, solution characteristics, methods used, etc. These statements are assertions about the business domain we are going to model. A sample language $\mathcal{L}$ is the TTCN-3 mentioned before.

Denote by $\mathcal{G}(\mathcal{L})$ the power set containing all *Arrow Terms* of the form (2). The left-hand side is a finite set of arrow terms and the right-hand side is a single arrow term. This definition is recursive; thus, it is necessary to establish a base definition saying that every assertion itself is considered an arrow term. The arrows of the arrow terms are distinct from the logical imply that some authors also denote by an arrow. The arrows denote cause-effect, not a logical implication.

The recursive formal definition, in set-theoretical language, is given in Eq. (2) and (3) below:

$$\mathcal{G}_0(\mathcal{L}) = \mathcal{L}$$

$$\mathcal{G}_{n+1}(\mathcal{L}) = \mathcal{G}_n(\mathcal{L}) \cup \{\{a_1, \ldots, a_m\} \rightarrow b | a_1, \ldots a_m, b \in \mathcal{G}_n(\mathcal{L}), m = 0, 1, 2, 3\ldots\} \quad (2)$$

where $n \in \mathbb{N}$. $\mathcal{G}(\mathcal{L})$ is the set of all (finite and infinite) subsets of the union of all $\mathcal{G}_n(\mathcal{L})$:

$$\mathcal{G}(\mathcal{L}) = \bigcup_{n \in \mathbb{N}} \mathcal{G}_n(\mathcal{L}) \quad (3)$$

The elements of $\mathcal{G}_n(\mathcal{L})$ are arrow terms of level n. Terms of level 0 are *Assertions*, terms of level 1 *Test Cases*. Sets of test cases are called *Rule Set* [14]. When a rule set is finite, it is called *Test Schema*; thus, a finite set of arrow terms. We call infinite rule sets a *Knowledge Base*. Hence, knowledge is a potentially unlimited set of test cases.

A *Test Story* is a

- finite rule set
- whose level 1 elements, those from $\mathcal{G}_1(\mathcal{L})$, relate to some specific business domain.

3.4 Combining Tests

Let M, N be two rule sets, consisting of test cases. N is a rule set consisting of arrow terms of the form $b_i = (\{x_1, x_2, \ldots, x_n\} \rightarrow y)_i$. Then, application of M to N is defined by

$$M \bullet N = \{c | \exists \{b_1, b_2, \ldots, b_m\} \rightarrow c \in M; b_i \in N\} \quad (4)$$

In other words, if all b_i executed in N with pass, the test story M can be applied to a rule set N as a set of test cases. This represents the selection operation that chooses those rules $\{b_1, b_2, \ldots, b_m\} \rightarrow c$ from test story M which are applicable to the rule set N. Combining tests is a strong means to create new test stories from existing ones.

Combinatory Algebra [15] is the mathematical theory of choice for automatically extending test cases from a simpler, restricted system, to test stories that fully cover a larger, expanded system. The extension works only if software testing is not only automated but also measured. Metrics must be independent from current implementation and from actual system boundaries.

The algebra of test cases (arrow terms) is a combinatory algebra that is a model of Combinatory Logic. This model is Turing-complete: i.e., all programs that are executable by computers can be simulated. This guarantees the best achievable test coverage.

Definition (4) looks somewhat counter-intuitive. To apply one test case to another, it is required that the result of application contains all the full test cases providing the response sought.

Intuition would expect writing test cases as $a \rightarrow b$, where a describes the program state prior to execution, and state b holds after executing the test. It can be shown that combining such terms would lead to a contradiction to Turing's halting problem [16]. Such terms can be formally combined in such a way that the result yields a contradiction, i.e., there is no model for such a logic.

3.5 Arrow Term Notation

To avoid the many set-theoretical parenthesis, the following notations are applied:

- a_i for a finite set of arrow terms, i denoting some finite indexing function for arrow terms.
- a_1 for a singleton set of arrow terms, i.e. $a_1 = \{a\}$ where a is an arrow term.
- $\emptyset$ for the empty set, such as in the arrow term $\emptyset \rightarrow a$.
- (a) for an (potentially) infinite set of arrow terms, where a is an arrow term.

The indexing function cascades; thus, $a_{i,j}$ denotes the union of a finite number of arrow term sets, selected by the indexing function j:

$$a_{i,j} = \bigcup_j (a_i)_j \tag{5}$$

With these writing conventions, $(x_i \rightarrow y)_j$ denotes a rule set, i.e., a finite set of arrow terms having at least one arrow. Thus, they are level 1 or higher.

Using this notation, the application rule (4) for $M = (b_i \rightarrow a)$ and $N = (b_i)$ now reads

$$M \bullet N = ((b_i \rightarrow a) \bullet (b_i)) = (a) = \{a | \exists b_i \rightarrow a \in M, b_i \subset N\} \tag{6}$$

The index i thus is a selection function, constructively selecting finitely many arrow terms from a rule set. Parenthesis around an arrow term without index indicates any set, finite or infinite, of arrow terms.

3.6 Arrow Terms – A Model of Combinatory Logic

Combinatory algebras are models of *Combinatory Logic* [17] and [18]. Such algebras are combinatory complete. This means that there is a combination operation $M \bullet N$ for all elements M, N in the combinatory algebra and the following two *Combinators* **S** and **K** can be defined with the following properties

$$\mathbf{K} \bullet M \bullet N = M \tag{7}$$

and

$$\mathbf{S} \bullet M \bullet N \bullet L = M \bullet L \bullet (N \bullet L) \tag{8}$$

where M, N, L are elements in the combinatory algebra.

If the algebra of arrow terms is a combinatory algebra and thus a model of combinatory logic, it means that tests can cover every conceivable computation, and because of recursivity, even testing tests.

The following definitions demonstrate how arrow terms implement the combinators **S** and **K** fulfilling Eqs. (7) and (8).

- $\mathbf{I} = (a_1 \to a)$ is the *Identification*, i.e. $(a_1 \to a) \bullet (b) = (b)$
- $\mathbf{K} = (a_1 \to \emptyset \to a)$ selects the 1st argument:
  $\mathbf{K} \bullet (b) \bullet (c) = ((b_1 \to \emptyset \to b) \bullet (b)) \bullet (c) = (\emptyset \to b) \bullet (c) = (b)$
- $\mathbf{KI} = (\emptyset \to a_1 \to a)$ selects the 2nd argument:
  $\mathbf{KI} \bullet (b) \bullet (c) = ((\emptyset \to c_1 \to c) \bullet (b)) \bullet (c) = (c_1 \to c) \bullet (c) = (c)$
- $\mathbf{S} = \left(a_i \to \left(b_j \to c\right)\right)_1 \to \left((d_k \to b)_i \to \left(b_{j,i} \to c\right)\right)$

These equations proof that the algebra of arrow terms is a model of combinatory logic. Only the proof that the arrow terms' definition of **S** fulfils Eq. (8) is somewhat more complex. The interested reader can find it in Engeler [10, p. 389]. With **S** and **K**, an abstraction operator can be constructed that builds new knowledge bases. This is the *Lambda Theorem*; it is proved along the same way as Barendregt's Lambda combinator (Barendregt [19] and Fehlmann [20, p. 37]). The Lambda theorem shows that this combinatory algebra allows constructing combinators for everything that is computable. The combinatory algebra is Turing-complete, i.e., it is computationally universal [21].

However, knowledge bases are quite unhandy since they are infinite; much more attractive seem finite sets of test cases (arrow terms), especially test stories. As said before, a test story is a homogeneous set of test cases in the sense that a test story addresses the behavior of the software under test that needs to be ascertained, because it matters for the user or customer. Thus, a test story is a set of arrow terms that share certain characteristics of interest for the software product. Security, safety, or privacy protection are among the properties that matter most; see the authors' previous work

[22] and [23]. If test stories in turn approximate knowledge bases, then it becomes interesting. However, what means ‘approximate”? In mathematics, it means that there is measure for arrow terms that can be used to assess distance between test stories.

The combinators allow combining, and extracting, test cases from test stories and thus generating additional test cases in an automated way.

4 Selecting Relevant Tests

Generating test cases automatically matters because the number of tests needed for complex systems grow with its functional size. But when a few hundred thousand of test cases are needed to test a complex systems – such as an autonomous vehicle, an autopilot or an *Advanced Driving Assistant System* (ADAS), then there is no other choice than selecting test cases automatically, and execute them using ART.

4.1 Software Under Test

Not every piece of coded functionality is relevant for ART. We need a model for the functionality that we want to test, the *Software under Test*. For instance, when testing an ADAS, the software built into a permanent storage device, e.g., an SSD device, is not relevant, while all functional processes belonging to the car driving functions are. Camera, visual recognition engine and driving recommendation engines provide relevant functionality whose cooperation needs being tested.

The *Data Movement Map* [14, p. 157] in Fig. 1 identifies the COSMIC data movements between objects of interest such as functional processes, persistent storage, devices and other applications. Other application engines need separate module tests.

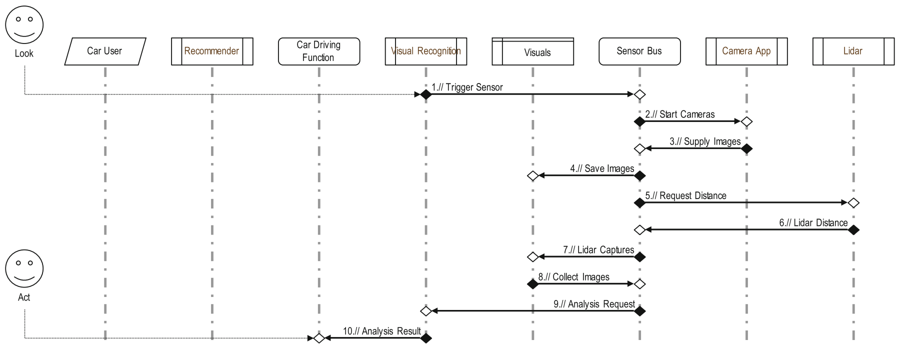


Fig. 1. ADAS look for obstacles – high-level extract

The size of software for an ADAS is huge. For testing, we need to divide its functionality into layers, preferably using the definitions from ISO/IEC 19761 [24] and the COSMIC Measurement Manual [25]. Figure 1 is a simplified excerpt of the top layer for this kind of software, modeling the ADAS Look & Act functionalities on a high, user-recognizable way. ISO/IEC 14143 [26] defines the term *Functional User Requirements* (FUR) for identifying the relevant user expectations against software.

FUR transform into *User Stories* in an agile product development environment. We use statistical methods (ISO 16355, [27]) for mapping test stories into user stories, based on *Six Sigma Transfer Functions*. For details how to set up Six Sigma transfer functions and how to solve them in a software product development environment, see Fehlmann [14, p. 21 ff].

4.2 The Size of Tests

For a testing framework, we need to be able to measure the size of tests. Since we already use ISO/IEC 19761 COSMIC for measuring functional size, it is obvious that the identical measuring method serves for tests as well.

Each test case executes a piece of code that can be measured using COSMIC. This is the *Size* of a test case. For arrow terms, we need to expand this definition recursively. Let $a_i \in \mathcal{G}_0(\mathcal{L})$ and $b \in \mathcal{G}_0(\mathcal{L})$ be arrow terms of level 0, i.e., assertions. Set $\|b\| = 0$ for every arrow term of level 0. Let $\|a_i \to b\|$ denote the COSMIC size $Cfp(a_i \to b)$ of the functionality executed when running the test case $a_i \to b$, i.e., the number of unique data movements touched when executing $a_i \to b$. This defines the recursion base for level 1 arrow terms.

Then the following equations recursively define the size of tests:

$$\begin{aligned} \|a\| &= 0 \text{ for } a \in \mathcal{G}_0(\mathcal{L}) \\ \|a_i \to b\| &= Cfp(a_i \to b) \text{ for } a_i \in \mathcal{G}_0(\mathcal{L}) \text{ and } b \in \mathcal{G}_0(\mathcal{L}) \\ \|c_i \to d\| &= \sum_i \|c_i\| + \|d\| \text{ for all arrow terms } c_i \text{ and } d \end{aligned} \tag{9}$$

When sizing test stories, or knowledge bases, the addition does not look whether data movements are unique; thus, the size of two test cases is always the sum of the sizes.

4.3 Metrics for Test Coverage

Test Coverage is still understood as code coverage by tests. However, when using cloud services, or AI [28], there is no code that can be tested. Moreover, not all tests are equally relevant for the user or customer. Since tests are a scarce resource, we need metrics that allow identifying the relevance of tests.

The method of choice is *Quality Function Deployment* (QFD) according ISO 16355 [27]. QFD uses standardized *Voice of the Customer* methods to identify a *Goal Profile* for product development, most often named “Customer Needs”. Various techniques are listed in ISO 16355 for uncovering customer needs, including Saaty’s *Analytic Hierarchy Process* (AHP) [29] and Reichheld’s *Net Promoter Score* (NPS) [30].

A goal profile for the expectation of users, or customers (Fig. 2), allows to use transfer functions to derive a profile for the respective User Stories (Fig. 3). For the full details, consult Fehlmann and Kranich [31] and Fehlmann [14, p. 273 ff].

Customer Needs

	Customer Needs Topics	Attributes			AHP Priorities Weight	Profile
Y.a Drive Fast	y1 Agile Driving	Arrive safe	Do not block c	Have fun	16%	0.36
	y2 Smooth Driving	Drive predictib	Do not break u		15%	0.32
	y3 Arrive in Time	Arrive predicti	Avoid obstacle		23%	0.50
Y.b Drive Safe	y4 Avoid Incidences	Drive foresight	Know what's a	Know my way	27%	0.58
	y5 No Surprises	Communicate	Never surprise	Give signs	19%	0.42

Fig. 2. The expectations of an ADAS user

The *Test Coverage Matrix* (Fig. 4) maps test stories onto user stories. The sample shown only takes very few test stories against the also very few user stories. Real samples are significantly larger. However, the important facts can be shown anyway.

4.4 What Numbers to Put into the Test Coverage Matrix Cells?

Let $\mathfrak{S}$ be a finite set of data movement maps. A level 1 test case $a_i \rightarrow b$ can be executed in $\mathfrak{S}$ if a data movement map in $\mathfrak{S}$ exists that transforms the program states a_i into b, both expressed as assertions.

Denote by $\cup\ \mathfrak{S}$ the union of all data movement maps in $\mathfrak{S}$. The *union* is defined in the straightforward manner by identifying all identical objects of interest within all data movement maps in $\mathfrak{S}$. Obviously, $\cup\ \mathfrak{S}$ is itself a data movement map. It represents the program under test, or more exactly, the part of the program that is covered by test cases, executable in $\mathfrak{S}$. Note that when combining executable test cases from program $\cup\ \mathfrak{S}$ using Eq. (6), the result is also executable in $\cup\ \mathfrak{S}$.

User Stories

	User Stories Topics	As a user...	I want to ...	such that ...	yielding benefit...	Priority Weight	Profile
1)	Q001 Populated Area	Car User	let my car reduce speed	my car can safely stop	my car is not causing delays by an incidence	19%	0.46
2)	Q002 Obstacle	Car User	let my car avoid obstacles	my car can drive around	my car is not stopping unnecessarily	13%	0.30
3)	Q003 Know my Way	Car User	let my car take appropriate routes	my car avoids blocked routes and traffic jams	I know when I'll arrive	14%	0.33
4)	Q004 Amend my Way	Car User	optimize my route when needed	no incidence blocks my way	I still can predict when I'll arrive	23%	0.54
5)	Q005 Check my Way	Car User	approve or disapprove the car's choice for routing	I can take my preferred route	I feel in control	14%	0.33
6)	Q006 Able to Stop	Car User	have my car break soon enough	it can avoid dangerous situations	It recognizes obstacles ahead	18%	0.43

Fig. 3. User stories (FUR) for ADAS

The number in the matrix cells represent the total test size that correlate between the respective user story and test story. This correlation cell test size is the number of data movements within all test cases in the specific test story that pertain to the respective user story. It can be calculated automatically once it is known which data movements support some specific user story.

User Stories	Goal Test Coverage	1) A.1 People around	2) B.1 Obstacle ahead	3) C.1 Get route	4) C.2 Change route	5) C.3 Update position	6) D.1 Approval	7) E.1 Arrival time	8) E.2 Learnings	9) F.1 Keep under control	10) F.2 Brake action	11) F.3 Avoid stops	Achieved Coverage
Q001 Populated Area	0.46	25	22	9	7	11	9	10	8	12		14	0.42
Q002 Obstacle	0.30	10	15	13	5	15	7	11	9	13	16	10	0.36
Q003 Know my Way	0.33	2	5	17	6	15	12	9	6	7	9	9	0.27
Q004 Amend my Way	0.54	24	19	14	19	21	9	25	9	17	15	21	0.59
Q005 Check my Way	0.33	16	13	6	5	7	23	12	8			20	0.35
Q006 Able to Stop	0.43	26	25	5	2	10	4	6	8	10	8	13	0.39
	Ideal Profile for Test Stories:	0.44	0.41	0.25	0.20	0.32	0.24	0.32	0.19	0.25	0.20	0.36	Convergence Gap 0.11

768 Total Test Size
0.15 Convergence Range
0.20 Convergence Limit

Fig. 4. Baseline test coverage matrix

Following Fehlmann and Kranich [32], the cell test size is calculated as follows:

- Test stories are a collection of rule sets (test cases) that share a common purpose. Let $\boldsymbol{t} \in \mathcal{S}$ be a test story, member of some rule set $\mathcal{S}$.
- For every test story $\boldsymbol{t}$, there is a mapping (10)

$$map(\boldsymbol{t}) \in \bigcup \mathfrak{S} \tag{10}$$

where $\cup \ \mathfrak{S}$ is the set of all data movement maps within a software program.

- Furthermore, there is a choice function $\|map(\boldsymbol{t})\|_f$ identifying which data movements pertain to some specific user story $\boldsymbol{f}$ and counting them.
- For each cell, we start with a rule set of test cases $\boldsymbol{t}_{i,j} \in \mathcal{S}_j$, where i, j are the respective cell indices of the matrix $\boldsymbol{A}$ and $\mathcal{S}_j$ is the respective test story in that matrix. Then the Eq. (11)

$$\sum \|map(\boldsymbol{t}_{i,j})\|_f \tag{11}$$

counts for each test case how many data movements pertain to the respective user story. The summation runs over all test cases $\boldsymbol{t}_{i,j} \in \mathcal{S}_j$ for each matrix cell that has index i, j.

A data movement may appear in many test cases and pertain to more than a single user story. We count the total amount of times that a data movement is used, not the data movements as for test size.

The *Ideal Profile for the Test Stories* at the bottom of the matrix (Fig. 4) is the solution profile for the Test Coverage transfer function. The importance profile for test stories is the solution of the Test Coverage transfer function for the importance profile for user stories.

4.5 The Convergence Gap

Let $\boldsymbol{A}$ be test coverage matrix; $\boldsymbol{x}$ its solution profile and $\boldsymbol{y}$ the importance profile of the user stories. The *Convergence Gap* is the Euclidian difference between the importance profile for user stories and the matrix $\boldsymbol{A}$ applied to the importance profile of the test stories. Thus, Eq. (12)

$$|\boldsymbol{Ax} - \boldsymbol{y}| \tag{12}$$

defines the convergence gap, when $|\ldots|$ denotes the Euclidean length of a vector. Since the profiles $\boldsymbol{Ax}$ and $\boldsymbol{y}$ are unit length vectors, the convergence gap does not depend on the length of either $\boldsymbol{Ax}$ or $\boldsymbol{y}$. Obviously, the convergence gap is a metric for how well the tests that fill the test coverage matrix cells cover the functionality expressed with the user stories.

With this remark, we can distinguish test stories and – especially – test cases that are relevant for the customer. This relevance allows for automatic generation and selection of test cases.

4.6 Optimizing the Test Coverage Matrix

With this state of the framework, it is possible to generate as many test cases as needed. This yields a search tree. With standard AI search techniques, it is now relatively easy to add new test cases, using the convergence gap as hash function. Minimizing the convergence gap when adding test cases means improving the customer relevance of the tests.

Thus, the framework is complete once a *Support Vector Machine* (SVM) exists that completes such searching within reasonable time. Pupale [33] might be helpful for advice. However, solving a transfer function is a standard task in many scientific domains, and it looks promising to consult work done for QFD in that respect. A *Sensitivity Analysis* [34] identifies those matrix cells whose value needs being increased or lowered. This is an interesting start; however, test cases most often consist of more than one data movement, and the same data movement occurring in a specific test case might be used in other ***map***-mappings for other test cases as well, see Eq. (10).

The optimization problem thus is not easy, it might be in a higher computational complexity class. This is still unknown. Experience shows that humans are at least as good, using contextual understanding, to invent new test cases or expand existing test stories when the convergence gap fails to close. The limits of human activity are reached when complexity, or functional size, growths high. For instance, highly complex systems such as ADAS, autonomous cars, planes or railway stock simply have too many interdependent functional processes to be addressed by human testers [28].

5 Conclusion

Modern software testing is moving from a human-driven discipline into an automated and artificially intelligent environment. The combinatory algebra provides a mathematical framework for such testing.

However, testing is not on an urgent agenda point for many suppliers even of high-risk equipment. Usually, delivering value in time is much more important. Nevertheless, a few recent examples with planes that crash, railway stock that cannot get commissioned, autonomous cars that fail to hit the streets because of liability issues, and smartphones viruses that infect whole ICT operations suggest that software testing will become a critical asset for bringing new products to market.

References

1. Erich, F., Amrit, C., Daneva, M.: A qualitative study of DevOps usage in practice. J. Softw. Evol. Process **29**(6), June 2017
2. Booch, G., et al.: Object-Oriented Analysis and Design with Applications. The Addison-Wesley Object Technology Series, vol. 3. Addison-Wesley, Upper Saddle River (2007)
3. Beck, K.: eXtreme Programming Explained. Addison-Wesley, Boston (2000)
4. Fehlmann, T.M.: Autonomous Real-time Testing - Testing Artificial Intelligence and Other Complex Systems. Logos Press, Berlin (2020)
5. Korsaa, M., et al.: The SPI manifesto and the ECQA SPI manager certification scheme. J. Softw. Evol. Process **24**(5), 525–540 (2012)
6. CloudBees, Inc. Jenkins - Build great things at any scale. https://jenkins.io. Accessed 31 Mar 2020
7. Atlassian. Inc. Atlassian Software Development. https://www.atlassian.com/?tab=code-build-and-ship. Accessed 31 Mar 2020
8. JUnit Team. The new major version of the programmer-friendly testing framework for Java, Open Source Development. https://junit.org/junit5/. Accessed 8 Apr 2020
9. Poppendieck, M., Poppendieck, T.: Implementing Lean Software Development. Addison-Wesley, New York (2007)
10. Engeler, E.: Algebras and combinators. Algebra Universalis **13**, 389–392 (1981). https://doi.org/10.1007/BF02483849
11. Engeler, E.: Neural algebra on "how does the brain think?". Theor. Comput. Sci. **777**, 296–307 (2019)
12. ETSI European Telecoms Standards Institute. TTCN-3 Standards, ETSI, Sophia-Antipolis Cedex, France (2018). http://www.ttcn-3.org/index.php/downloads/standards. Accessed 11 Dec 2018
13. Ebner, M.: TTCN-3 test case generation from message sequence charts, Göttingen, Germany. In: Workshop on Integrated-reliability with Telecommunications and UML Languages (ISSRE04:WITUL) (2004)
14. Fehlmann, T.M.: Managing Complexity - Uncover the Mysteries with Six Sigma Transfer Functions. Logos Press, Berlin (2016)
15. Engeler, E.: The Combinatory Programme. Birkhäuser, Basel (1995)
16. Turing, A.: On computable numbers, with an application to the Entscheidungsproblem. In: Proceedings of the London Mathematical Society, vol. 42, no. 2, pp 230–265 (1937)
17. Curry, H., Feys, R.: Combinatory Logic, vol. I. North-Holland, Amsterdam (1958)

18. Curry, H., Hindley, J., Seldin, J.: Combinatory Logic, vol. II. North-Holland, Amsterdam (1972)
19. Barendregt, H.P.: The type-free lambda-calculus. In: Barwise, J. (ed.) Handbook of Math Logic, vol. 90, Amsterdam, North Holland, pp. 1091–1132 (1977)
20. Fehlmann, T.M.: Theorie und Anwendung des Graphmodells der Kombinatorischen Logik, ETH Dissertation 3140-01, Zürich, CH (1981)
21. Barwise, J., et al.: Handbook of mathematical logic. In: Barwise, J. (ed.) Studies in Logic and the Foundations of Mathematics, vol. 90. North-Holland Publishing Company, Amsterdam (1977)
22. Fehlmann, T., Kranich, E.: Theoretical aspects of consumer metrics for safety & privacy. In: Larrucea, X., Santamaria, I., O'Connor, Rory V., Messnarz, R. (eds.) EuroSPI 2018. CCIS, vol. 896, pp. 640–653. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-97925-0_54
23. Fehlmann, T.M., Kranich, E.: Practical aspects of consumer metrics for safety & privacy. In: Proceedings der GI-Informatik 2018 (proposed), Zurich, Switzerland (2018)
24. ISO/IEC 19761. Software engineering - COSMIC: a functional size measurement method, ISO/IEC JTC 1/SC 7, Geneva, Switzerland (2011)
25. COSMIC Measurement Practices Committee. The COSMIC Measurement Manual for ISO 19761 – Version 5.0, Part 1–3, The COSMIC Consortium, Montréal (2020)
26. ISO/IEC 14143-1. Information technology - Software measurement - Functional size measurement - Part 1: Definition of concepts, ISO/IEC JTC 1/SC 7, Geneva, Switzerland (2007)
27. ISO 16355-1:2015. ISO 16355-1:2015, 2015. Applications of Statistical and Related Methods to New Technology and Product Development Process - Part 1: General Principles and Perspectives of Quality Function Deployment (QFD), Geneva, Switzerland: ISO TC 69/SC 8/WG 2 N 14, ISO TC 69/SC 8/WG 2 N 14, Geneva, Switzerland (2015)
28. Fehlmann, T.: Testing artificial intelligence. In: Walker, A., O'Connor, Rory V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 709–721. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_55
29. Saaty, T.L.: Decision-making with the AHP: Why is the principal eigenvector necessary? Eur. J. Oper. Res. **145**, 85–91 (2003)
30. Fehlmann, T.M., Kranich, E.: Uncovering Customer Needs from Net Promoter Scores, Istanbul, Turkey (2014)
31. Fehlmann, T.M., Kranich, E.: Testing artificial intelligence by customers' needs. Athens J. Sci. **6**(4), 265–286 (2019)
32. Fehlmann, T.M., Kranich, E.: Intuitionism and computer science – why computer scientists do not like the axiom of choice. Athens J. Sci., Submitted
33. Pupale, R.: Support Vector Machines (SVM) - An Overview, 16 June 2018. https://towardsdatascience.com/https-medium-com-pupalerushikesh-svm-f4b42800e989. Accessed 28 Mar 2019
34. Fehlmann, T.M., Kranich, E.: A Sensitivity Analysis Procedure for QFD, Duisburg (2020, to appear)

Sustainability Efficiency Challenges of Modern IT Architectures – A Quality Model for Serverless Energy Footprint

Alexander Poth[1(✉)], Niklas Schubert[1(✉)], and Andreas Riel[2(✉)]

[1] Volkswagen AG, Berliner Ring 2, 38436 Wolfsburg, Germany
{alexander.poth,niklas.schubert}@volkswagen.de
[2] Grenoble Alps University – Grenoble INP – G-SCOP, 38031 Grenoble, France
andreas.riel@grenoble-inp.fr

Abstract. Cloud technology is transforming IT architecture. Serverless is a new paradigm driven by cloud technology that offers chances for scalable and reliable IT services. All bigger cloud service providers (CSP) offer a set of serverless services to build entire business processes on it. The services are mostly based on efficient containerization of the workload. However, the service offers are driven by parameters like scaling, availability, ease of use and costs. So far, no cloud provider has been providing information about service efficiency, in particular energy efficiency. However, breaking down complex services into micro, nano or pico services as serverless comes with an energy overhead. This energy footprint of serverless architectures demands a specific quality model to refine the ISO 25010 for the serverless domain. The model presented in this article is provided as a self-service to the Volkswagen Group IT to enable all architects and software engineers to evaluate the trade-offs with respect to sustainability aspects.

Keywords: Serverless architecture · Efficiency evaluation · Energy footprint · Sustainable service model · Function as a Service (FaaS)

1 Motivation and Context for the Demand

Enterprises are running different initiatives in parallel like the agile transformation, which is an organizational change, and the cloud transformation, which is a technology- driven change. Both transformations have the objective of improving time to market and reacting fast to changes in the world. The cloud native paradigm can be realized with serverless technology. There exist many use cases for using the serverless approach like websites, analytics with thousands of requests per second or establishing reliable workflows [1]. Many other use cases with event-driven, short running and state-less scenarios [2] are also candidates for serverless. As an unbroken trend, the data centers consume more and more energy [3], in particular those hosting serverless IT services.

Often, serverless computation is called Function as a Service (FaaS). However, it is actually more than that because it is also based on event-driven workflows [4]. In an abstract view, serverless computing is a kind of containerized workload. The serverless

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 289–301, 2020.
https://doi.org/10.1007/978-3-030-56441-4_21

service is optimized to run many small containers or virtual micro-machines. The approach uses code as a function which is deployed within a container that includes an adequate runtime environment for the code [5]. The serverless paradigm is spread over the entire value stream of an IT system by establishing a complete serverless stack which includes serverless data base, object store and API gateway to feed the serverless compute service. Serverless stacks can be used as a service like the offers from big cloud service providers AWS, Microsoft Azure and Google Cloud. They all provide similar functionality in their serverless stacks. As serverless is not really served without servers, a serverless stack can be deployed also on premise in private data centers. Furthermore, there are open source serverless projects like knative [6] or openWhisk [7]. Currently there is no project providing the entire serverless stack. The PLONK (Prometheus, Linerd, OpenFaaS, NATS and Kubernets) stack is an approach for establishing an open source based serverless stack. However, there existing other open source projects like FONK (FaaS, Object store, NoSQL and Kubernetes) for establishing an open source serverless stack [8]. Also, knative with its deep integration into Kubernetes [9] is close to filling the gap with for example kong [10] or traeffik [11] as API-gateway and minio [12] as object store and one of the many open source databases like MongoDB as scalable NoSQL data base. Also nuclio [13] is an open and integrated serverless approach for efficient event processing.

The abstraction of serverless applications from the server infrastructure leads to service offers that are based on costs correlated with function utilization. From the application point of view, this is a new cost model because it comes without the fixed infrastructure costs and scales linearly with the function calls. Currently no serverless platform makes transparent how energy efficiently the service is provided. This work presents a model to evaluate serverless, especially FaaS, in the context of efficiency to make early sustainable architecture decisions.

2 Efficiency of Serverless in Related Work

The current trend to serverless system and application architectures raises the need to have a deeper look into the ecological impact of the serverless approach. We identified four Research Questions (RQ) to get the base for a systematic approach to balance the cost and energy efficiency of serverless:

1. RQ1: How can we measure the energy efficiency of the abstraction layers by serverless computing?
2. RQ2: How does a generic model to measure the energy efficiency trends of serverless computing look like?
3. RQ3: How many resources are consumed by the abstraction layers of serverless computing?
4. RQ4: How to design economically efficient serverless systems by keeping ecological aspects in mind?

Furthermore, we will investigate current serverless computing offers in terms of their environmental footprints and how they are reflected in the pricing model. This scope is necessary because the pricing model is a key business driver for serverless decisions.

To design a generic evaluation approach, some assumptions about serverless approaches have been made, aligned with the current technology typically used to implement serverless platforms:

- Function calls get committed resources allocated at least during execution time.
- Function calls can run in parallel to scale.
- Function calls have a delay between invocation and execution start.
- Management services runs to orchestrate function calls and resource allocations.

These assumptions are based on the infrastructure to execute serverless functions. Note that they do not require a scheduler integrated in the service, which make solutions using the kubernetes scheduler eligible. These assumptions do not focus on the infrastructure needed to develop and deploy serverless functions. On the usage level, serverless can be defined by a more rigor view like shipping only code artifacts with library packages – like openWhisk – or less rigor by allowing also containers shipped for execution – like knative. Depending on the design of the serverless platform, there is a static allocation of CPU and RAM for each function invocation. Demanded data transfer and storage connections are established for each invocation. Serialization/deserialization of data as an overhead and sometimes more inefficient with blocking I/O. Based on the specific implementation, the efficiency is limited and reduced by the overhead of the serverless service management.

In [14], the generic overhead observation is defined by measurement of some time slots to define efficiency: Mean Time Between Calls (MTBC), Mean Time To Setup (MTTS) and Mean Time To Execute (MTTE). The Efficiency ε is defined by:

$$\varepsilon = \mathrm{MTBC} - (\mathrm{MTTS} + \mathrm{MTTE})$$

The resource saving will be better with $\varepsilon \gg 0$. The most problematic area is a small positive ε in which the MTTS has a significant part in ε.

In practice, the MTTS is defined by the cold and warm start of the function invocation. In our context, we define MTTS by the following three generic sub-stages which are happen in each serverless platform by design:

- *Container Setup* – establish a secure and resource allocated environment for the execution of the function (serverless platform defined).
- *Runtime Setup* – establish the function specific runtime environment with all its configurations to serve the function (depending on the service model defined by the serverless platform or the workload owner with like AWS lambda layers).
- *Workload Setup (Dependency Resolving)* – deploys the workload and ensure that all dependencies of the function are resolved to run the workload function (workload owner defined).

3 Generic Model for Evaluation of Serverless Consumption

3.1 Execution Life-Cycle Stage Model

The missing part of the model in [14] is the sub-stage *Clean Up* for the “garbage-collection” overhead after the function has terminated. Furthermore, the overhead for the management of the serverless function during runtime, as well as the service API that is invoked to set up a function are missing in [14]. Figure 1 shows the stage model in detail. This model is based on a sequence which is mapped to a timeline (x-axis) and the energy consumption (y-axis). In more detail, the stages are as follows:

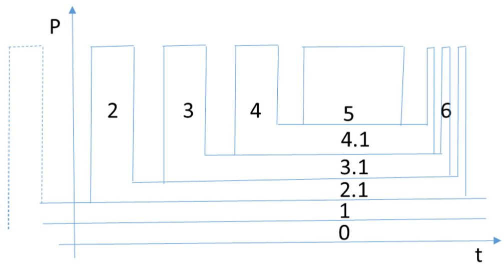


Fig. 1. Schematic correlation of function related execution stages.

0. *Core:* Includes all that is needed to run on a specific server hardware, i.e. hardware drivers, Hypervisor and the basic operating system (the peaks of the initial start and final stop are not relevant in the holistic view).
1. *Serverless:* service environment with the API and container management (the peaks of the initial start and final stop are not relevant in the holistic view).
2. *Container Setup:* workload to initialize the container environment based on the demand CPU, RAM and local storage quota and the establishment of the network.
   2.1. *Container Idle:* workload that is caused by the administration for maintain an idling container.
3. *Runtime Setup:* workload to load and initialize the language runtime with the right settings and parameters to serve the function.
   3.1. *Runtime Idle:* workload which is caused by the running language runtime in idle.
4. *Workload setup:* deploys the demanded function and resolves the dependencies of its libraries – includes everything before the payload function handler is invoked.
   4.1. *Workload Idle:* workload that is caused by the dependencies and connections.
5. *Run workload*: this is the demanded workload of the serverless function.
6. *Clean Up*: steps 2 to 4 have to be cleaned up after function termination, these short peaks are wrapped up in this stage for simplification.

One can distinguish between the temporary energy consumption of 2, 3, 4 and 6 with their intensive spikes and the overhead consumption of 0, 1, 2.1 and 3.1, that is constant over the life-cycle of the function. Furthermore, it is useful to distinguish between the stages 0, 3.1 and 5 to run a workload, which will happen also in a non-serverless environment with a similar amount, and the other stages, which will happen much more often in a serverless architecture. An additional deeper look into 5 can be interesting to optimize the function code to deal with the specific constraints of the serverless platform like establishment of network or storage access. This aspect is not in the scope of this generic model but can lead to a significantly more inefficient function execution.

The execution frequency of the steps 2, 3 and 4 in correlation to the workload execution of step 5 is important for the energy efficiency. A *cold start* signifies the execution of all the steps from 2 to 6. To optimize utilization of the hardware, the serverless service (1) tries to reuse containers for a recall of the same function. This reduces the steps 2 to 4 into a cache load which is much faster and more energy efficient, and is called *warm start*. The CPU utilization directly correlates with energy consumption [15] and blocks or allocates many other components of the hardware during CPU utilization. This leads to the following efficiency view – and answer to RQ1 – for the time to serve a function call based on the defined six steps:

$$\begin{aligned} &\text{Efficiency (function\_serving)} = \text{run\_workload}/(\text{core} + \text{serverless} \\ &+ \text{container\_setup} + \text{container\_idle} + \text{runtime\_setup} + \text{runtime\_idle} \\ &+ \text{workload\_setup} + \text{workload\_idle} + \text{clean\_up}) \end{aligned}$$

Figure 2 shows the efficiency of a function call in relation to the runtime of the workload (run_workload). RQ2 can be answered with this calculation and visualization. This calculation assumes that the workload does everything efficiently. Depending on some serverless platforms, the workload is blocked if there is some waiting for data (i.e. blocking I/O). This will negatively impact utilization and energy efficiency and is to be mitigated by writing “clever” workload code, which avoids waiting in a function. Furthermore, the aspect of short time running code in step 5 cannot get the full advantage of runtime optimizations like branch predictions [16]. To compensate this topic, more resources and energy are needed which is not addressed in the calculation

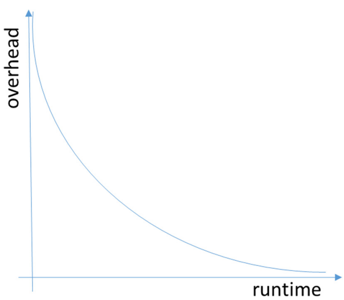


Fig. 2. Schematic correlation between function overhead and workload execution runtime.

because it strongly depends on the specific workload and their dependencies [17]. RQ3 looks in our setup as negligible and a systematic measurement was not conducted.

3.2 Quality Model Aligned with ISO/IEC 25010

Based on the stage model and the calculations, a generic quality model for serverless workloads can be derived. An established quality model for IT systems, services and products is the ISO/IEC 25010 which will be applied to the serverless technology.

The ISO/IEC 25010 defines eight quality characteristics and some sub-characteristics for products or IT systems that are instantiated to serverless:

- Functional suitability: completeness, correctness and appropriateness – depends mostly on the use case of the workload.
- Performance efficiency: time behavior (the topic warm and cold starts are a relevant), resource utilization (this leads to our investigation focus: energy efficiency) and capacity (in most cases, scaling is an inherent feature of serverless).
- Compatibility: co-existence (inherently given based on the API-approach of serverless) and interoperability (in many cases a login to the specific serverless platform, provider or service vendor).
- Usability: depends on the serverless platform and workload.
- Reliability: maturity, availability, fault tolerance and recoverability – given for many established serverless platforms given; availability and fault tolerance are inherent in the serverless platforms.
- Security: a least privileges approach can be defined on function level; a large number of roles leads to a complex configuration management for fine grained least privileges.
- Maintainability: high modularization with small code “pieces” are easy to maintain, however serverless leads to many small pieces which can lead to a configuration management issue etc.
- Portability: is limited by the login to a specific (proprietary) serverless platform.

The generic model for serverless technologies shows that many topic are defined by the used serverless platform. The limitation of the workload influences many of these quality characteristics. Based on this observation, it is important to make the decision for serverless consciously and select the most appropriate platform to avoid lock-in decisions.

3.3 Application of the Life-Cycle and Quality Model for Serverless Workloads

For the application of the derived model, a workload use case specific evaluation is needed. The three generic scenarios for serverless computing workloads are:

a) *Infrequent calls:* the function is called in a frequency which hits cold starts. A mitigation action can be setting up a warming activity. The warming activity triggers a cyclic call of the function to avoid their garbage collection. Depending on the serverless platform this trigger frequency varies.

b) *Frequent calls:* the function is called in a constant frequency to run as warm start. In this case serverless is running well out of the box.
c) *Spiky calls:* the bursts are demanding scaling of function instances and leads to cold starts of the additional demanded instances. It is difficult to handle this from the service consumption side. Many serverless platforms offer pre-provisioning of instances to handle this. In this scenario, the art is to activate pre-provisioning just in time for handling the spikes which are typically driven by the business demands.

For handling a specific workload use case, the case has to be mapped to one of these scenarios. For real-world applications, the calculation of workloads has be applied to each scenario. This scenario is a linear consumption. Scenario b) is covered by using the present calculation for each call. Scenario a) needs additional calls to keep the function warm. This warming function is a practical "hack" to fix a design issue of the serverless approach. Here the trade-off has to be monitored because the energy consumption for the keep the function warm grows over time, and at some point the cold start is more efficient. This point depends on the time during which a specific serverless platform holds functions in the caches and how efficient the warming is realized. This scenario is the basic consumption of serving a function. Scenario c) is by design the anti-pattern of serverless because the platform offers an over-provisioning to serve function call demands. This approach adds a scaling factor to scenario b) as an offset to scenario a).

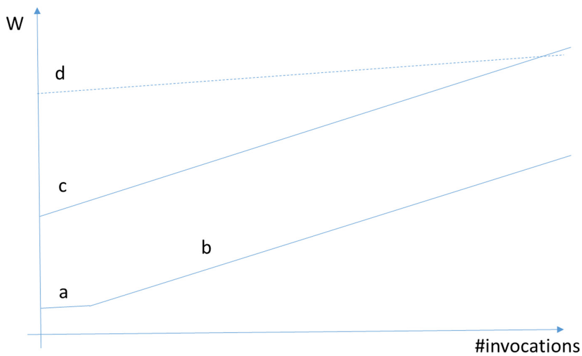


Fig. 3. Schematic correlation of different serverless usage scenarios.

Figure 3 shows the energy consumption of the three scenarios and the break-even point to classical (auto-)scaling of instances (d). The gradient of the instances is given by the growing workload. All gradients and offsets are schematic in this figure.

4 Model Instantiation and Evaluation

4.1 Setting

For the instantiation and validation of the model, a transparent open source solution has been chosen wit the following selection: use as many components as possible from typical stacks and solutions having a historical track record with many active contributes. This has led us to openFaaS as an established FaaS solution based on kubernetes and knative with many active developers and frequent release updates.

Our experiment setup is aligned with [18] for transparency. The tests have been made on an AMD Ryzen 3700x CPU based machine with 16 GB RAM running minikube 1.19 with kubernetes 1.18 and openFaaS 0.18. For the runtime, node.js in the current LTS version 12.13 has been selected from the official template store as one of the most used runtimes in the serverless context [20]. The “hello world” function runs in a container with openFaaS default settings without performance tuning or additional resource assignments/limits. Minikube resources are limited to 4 vCPU and 8 GB RAM. Tests are executed 20 times.

4.2 Evaluation

Table 1 shows measures to support the stages defined in chapter 3.1. To validate that the added measurement instrumentation does not significantly impact the performance and behavior of openFaaS, the tests were run on an out-of-the-box installation too. The validation tests confirm that no measurable derivations exists. Aligned with [19], different cold start types were investigated: to address platform specific architecture aspects, the Instance Cold Start (first run of the function on the node) has been extended with a Cache Cold Start (function is scaled to zero, but in the node caches populated) and the Warm Start based on a running instance of the function. Provider cold starts (deploy) were not in scope of the measures.

Table 1. Measures of the execution life-cycle stages of a serverless function in openFaaS.

Stage	Comment	Min [ms]	Max [ms]	Avr [ms]
0	5% CPU and 1,5 GB RAM in idle	–	–	–
5	Cache Cold Start	3844	6121	4951
	Warm Start	6	13	9

The measures show that for short running functions with less than 10 ms, the Warm Start overhead is more than 100% to the workload exaction time and leads to a node utilization lower than 50%, as well as existing additional overhead of the stages 0 and 1). Furthermore, no node is used at 100% to avoid an instable system. For infrequently used functions, there are 5 s of overhead for each invocation. The Instance Cold Start invocation happens rarely (only if the caches size forces displacement or setup on an other instances), and is only relevant for the user waiting for the function execution. For the scenario pre-warming the instance the trade-off are 550 warming invocations.

The overhead of the serverless platform can be significant in case of short running functions. For them, a trade-off point from the energy performance perspective is the point in which the average utilization of function invocations on a dedicated instance container/pod is higher than the FaaS function container with its overheads can serve. In case of infrequent function invocation which causes a scale to zero, there is a lot of overhead, however from the energy efficiency perspective there is no better approach.

4.3 Transfer and Generalization

Usage statistics from [20] shows that 123 lambdas are running in average on an AWS account and uses node.js with 76% as runtime. Commercial serverless services struggle with the same start types issues. Microsoft [21] and AWS [22] are working continuously on this topic to optimize the non-required workload ("launch a microVM in as little as 125 ms" [23]). Figure 4 is from the AWS blog [24] and shows the optimizations of cold starts to 167 ms and 39 ms overheads for an invocation of a lambda function. However, it is an inherent issue of the architecture of current serverless approaches and their technology base. Most of the approaches to avoid visible performance issues leads to "pre-warming". The providers [25] are pro-actively allocating resources to serve expected future function invocations. This wastes unused resources lowering the energy efficiency of the function invocation and resembling traditional auto-scaling of server instances.

Dependencies are an issue for serverless function because they have to be established often and need time e.g. for database connections [26]. For establishing the connection to write log messages, 25–50 ms are needed and at least 50 ms for setting up a database connection. In [27], an investigation of different typical dependencies for AWS are measured with a median range from 2 ms for no dependencies to 410 ms for loading the AWS and X-Ray SDK. This shows the responsibility of the developer for selecting carefully the referred libraries. E.g. a dynamoDB client dependency implies additional 100 ms.

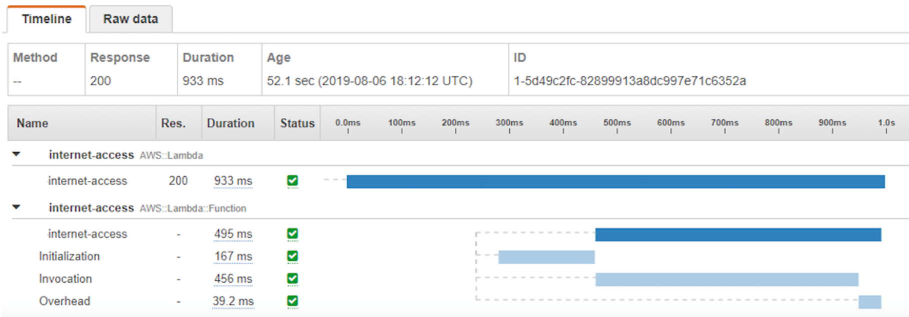


Fig. 4. Cold start of an AWS Lambda after optimizations by AWS Lambda service team.

4.4 Limitations of the Approach

Energy consumption is correlated with utilization [28]. Furthermore, performance can be correlated with utilization [29]. However, not every performance and/or utilization optimization optimizes energy efficiency [30]. This aspect has not been addressed. In general, any workload consolidation can help safe power consumption [31] and shows that a smart serverless approach helps reduce energy consumption.

5 Design of Sustainable Serverless Services

To answer RQ4, the pricing model of services has to be considered form the energy consumption perspective. To ensure green serverless computing, the sharing of resources and code quality by optimizing the runtime is identified in [32]. As shown in our model, this not automatically leads to energy efficient green computing because the pricing model of the public serverless platforms do not price the overheads directly to the function costs. The big providers like AWS, Microsoft Azure and Google Cloud have a price model based on execution time for allocated resources as key cost factor and an invocation fee for bundles as a secondary cost factor. Some providers like Google Cloud are finer grained by applying more factors in their pricing model, but these are not more directly energy focused. A more sustainable approach could be a design of the price-model based on the principle “pay what you consume”.

This leads to a metering of at least the following utilization indicators for FaaS of:

- The workload.
- The initialization of the runtime environment.
- The (cold) start of the container environment.

Finer grained metering is possible like to map overheads of virtual machine management to functions which can lead to make the RAM/CPU allocation per invocation to a non-linear function but for customers the cost transparency should also be kept simple as shown in Fig. 5. Based on the metering information, an overhead cost factor can be applied for service delivery costs driven by aspects like costs for

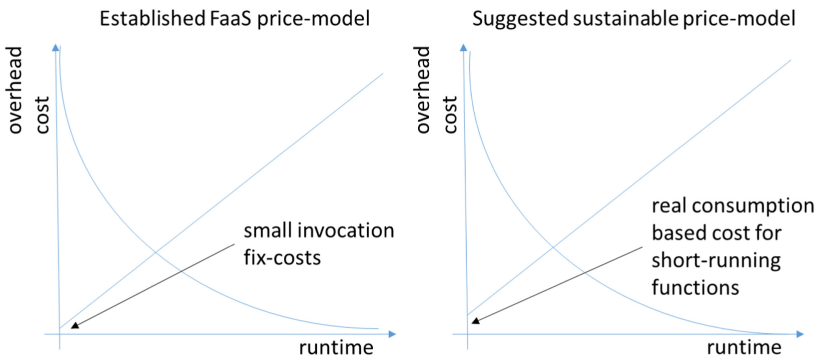


Fig. 5. Schematic correlation of price models.

operating and development. Finally, in commercial services, the targeted margin has to be added. By design of sustainable pricing models, the customers and the service provider will align their efforts on optimizing their value stream inherent on the green IT paradigm. Especially for service providers that do not run their data centers entirely with green energy [33], this can be the right step to align the fast growing energy consumption of IT with value provided to the users.

6 Conclusion, Implications and Outlook

This work shows that ensuring transparency about energy performance of serverless platforms is far from being trivial to find a balance between energy efficiency and offer reliable service [34]. This issue is driven by the specific setup of the platforms and in case of proprietary commercial services the lack of information and capability to measure this black box. However, the generic model derived from the core approach of FaaS helps to identify typical energy performance aspects to keep them in mind during serverless application design decisions. Therefore, this work has implications to different stakeholder groups:

The Serverless Service Providers
Currently public service offers are not transparent about the real energy consumption. Customers are not able to design their workloads according to the green IT paradigm without reverse engineering and un-confirmed assumptions. A potential solution keeping business internals private could be that the service providers take over responsibility for their sustainability and align the service price model with the real energy consumption to ensure that all parties optimize their workloads in a sustainable way.

The Serverless Consumers
Currently it is an explicit task to investigate and evaluate the energy performance of serverless application workloads and to argue for trade-offs between other optimization targets. A generic evaluation heuristic has been presented helping to reduce the effort for influencing and monitoring energy performance. This heuristic approach will provide a rough indication only as long as there is no detailed data from the specific serverless service available, since reverse engineering is not the right approach.

The Researchers
Currently a 3-digit number of research papers handling serverless computing are available, most of them treating timing performance topics. However, there is few published research concerning the energy performance topic. Are we developing the right IT-, service architectures and design approaches for the (near) future if we do not explicitly care about energy performance by design? Each abstraction layer has an ecological footprint. More comfort for developers by abstraction leads to a higher energy consumption as inbound effect, which the growing demand for serverless clearly shows. The presented approach shows that energy as one big position of IT's ecological costs can be aligned with economical costs.

An outlook to future work includes measurement of more stages by instrumentation of the openFaaS platform, the development of a self-service kit for different kind of cloud and IT services to evaluate the energy footprint easily and quickly to consider this aspect during architecture and design decisions as a first class quality characteristic of the IT system or service. Furthermore, a self-service kit for deriving systematically sustainable pricing-models for different types of IT services is a backlog item of the QiNET [35] to support end-to-end the green IT paradigm as a first class quality characteristic in the business.

References

1. https://aws.amazon.com/lambda/resources/customer-case-studies/. Accessed 5 June 2020
2. Fox, G.C., Ishakian, V., Muthusamy, V., Slominski, A.: Report from workshop and panel on the status of serverless computing and Function-as-a-Service (FaaS) in industry and research. In: First International Workshop on Serverless Computing (WoSC) (2017)
3. https://www.forbes.com/sites/forbestechcouncil/2017/12/15/why-energy-is-a-big-and-rapidly-growing-problem-for-data-centers/. Accessed 5 June 2020
4. Van Eyk, E., et al.: Serverless is more: from PaaS to present cloud computing. IEEE Internet Comput. **22**(5), 8–17 (2018)
5. Agache, A., et al.: Firecracker: lightweight virtualization for serverless applications. In: 17th Symposium on Networked Systems Design and Implementation, pp. 419–434 (2020)
6. https://knative.dev/. Accessed 5 June 2020
7. https://openwhisk.apache.org/. Accessed 5 June 2020
8. fonk-apps.io. Accessed 5 June 2020
9. https://kubernetes.io/. Accessed 5 June 2020
10. https://github.com/kong/kong. Accessed 5 June 2020
11. https://docs.traefik.io/. Accessed 5 June 2020
12. https://github.com/minio/minio. Accessed 5 June 2020
13. https://github.com/nuclio/nuclio. Accessed 5 June 2020
14. Kanso, A., Youssef, A.: Serverless: beyond the cloud. In: Proceedings of the 2nd International Workshop on Serverless Computing, pp. 6–10 (2017)
15. Hsu, C.H., Slagter, K.D., Chen, S.C., Chung, Y.C.: Optimizing energy consumption with task consolidation in clouds. Inf. Sci. **258**, 452–462 (2014)
16. Shahrad, M., Balkind, J., Wentzlaff, D.: Architectural implications of function-as-a-service computing. In: Proceedings of the 52nd Annual IEEE/ACM International Symposium on Microarchitecture, pp. 1063–1075 (2019)
17. https://dashbird.io/blog/optimising-your-serverless-applications/. Accessed 5 June 2020
18. Kuhlenkamp, J., Werner, S., Borges, M.C., Ernst, D., Wenzel, D.: Benchmarking elasticity of FaaS platforms as a foundation for objective-driven design of serverless applications. In: Proceedings of the 35th Annual ACM Symposium on Applied Computing. pp. 1576–1585 (2020)
19. Lloyd, W., Ramesh, S., Chinthalapati, S., Ly, L., Pallickara, S.: Serverless computing: an investigation of factors influencing microservice performance. In: IEEE International Conference on Cloud Engineering (IC2E), pp. 159–169 (2018)
20. https://dashbird.io/blog/state-of-lambda-functions-2019/. Accessed 5 June 2020
21. https://azure.microsoft.com/de-de/blog/understanding-serverless-cold-start/?ref=msdn
22. https://aws.amazon.com/about-aws/whats-new/2019/12/aws-lambda-announces-provisioned-concurrency/. Accessed 5 June 2020

23. https://aws.amazon.com/de/blogs/aws/firecracker-lightweight-virtualization-for-serverless-computing/. Accessed 5 June 2020
24. https://aws.amazon.com/de/blogs/compute/announcing-improved-vpc-networking-for-aws-lambda-functions/. Accessed 5 June 2020
25. https://azure.microsoft.com/en-us/updates/azure-functions-premium-plan-is-now-generally-available/. Accessed 5 June 2020
26. https://blog.coinbase.com/benchmarking-aws-lambda-ca3cfb3c25cd. Accessed 5 June 2020
27. https://theburningmonk.com/2019/03/just-how-expensive-is-the-full-aws-sdk/. Accessed 5 June 2020
28. Weissel, A., Bellosa, F.: Dynamic thermal management for distributed systems. In: Proceedimgs of Workshop on Temperature-Aware Computer Systems(TACS) (2004)
29. Podzimek, A., Bulej, L., Chen, L.Y., Binder, W., Tuma, P.: Analyzing the impact of CPU pinning and partial CPU loads on performance and energy efficiency. In: 15th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, Shenzhen, 2015, pp. 1–10 (2015)
30. Tsirogiannis, D., Harizopoulos, S., Shah, M.A.: Analyzing the energy efficiency of a database server. In: Proceedings of the 2010 ACM SIGMOD International Conference on Management of Data, pp. 231–242 (2010)
31. Verma, A., Dasgupta, G., Nayak, T.K., De, P., Kothari, R.: Server workload analysis for power minimization using consolidation. In: Proceedings of the Conference on USENIX Annual Technical Conference, p. 28. USENIX Association (2009)
32. Shafiei, H., Khonsari, A., Mousavi, P.: Serverless Computing: A Survey of Opportunities, Challenges and Applications (2020)
33. https://aws.amazon.com/about-aws/sustainability/. Accessed 5 June 2020
34. Beloglazov, A., Abawajy, J., Buyya, R.: Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing. Future Gener. Comput. Syst. **28** (5), 755–768 (2012)
35. Poth, A., Heimann, C.: How to innovate software quality assurance and testing in large enterprises? In: Larrucea, X., Santamaria, I., O'Connor, Rory V., Messnarz, R. (eds.) EuroSPI 2018. CCIS, vol. 896, pp. 437–442. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-97925-0_37

Designing a Cyber Range Exercise for Educational Purposes

Xabier Larrucea[1,2](✉) and Izaskun Santamaría[1]

[1] TECNALIA, Basque Research and Technology Alliance (BRTA), Bizkaia, Spain
{xabier.larrucea,izaskun.santamaria}@tecnalia.com
[2] University of the Basque Country, Vitoria-Gasteiz, Spain

Abstract. Cybersecurity is an emergent need in educational and industrial contexts. There is an increase need to train skilled people, students and employees on cybersecurity techniques. Training is one of the weaknesses identified within the industry especially by practitioners, and the use of cyber ranges is motivated. Cyber ranges are envisaged as infrastructures for training purposes. This paper provides an overview of cyber ranges, and a cyber range exercise design process illustrated with an example. This exercise is used in academic sector and it can be extrapolated to industrial contexts.

Keywords: Cyber range · Cyber scenario deployment

1 Introduction

Cybersecurity is playing a key role and it is having a huge impact in industry [1]. Its increasing relevance [2] and impact can be extended and applied to almost all of our lives. In fact, countries are adding cyber security within the curricula of informatics [3], or it is being considered as part of their national strategy [4]. In fact, the United States of America recognizes that they need to "*identify and bridge existing gaps in responsibilities and coordination among Federal and non-Federal incident response efforts and promote more routine training, exercises, and coordination*." [4], including the development of a robust cybersecurity workforce. Therefore, it represents an area to be strengthened by industry, academia and governments.

Cybersecurity is a broad area covering a wide set of different approaches and areas. One relevant aspect within the cybersecurity's umbrella is to analyze the applications' source code in order to mitigate malicious uses or attacks [5]. Traditionally, such activities are relying on programmers' skills, but this practice is too risky. The General Data Protection Regulation [6] released by the European Union, the security and privacy concepts have been taken seriously. In fact, the "Security by Design" and "Privacy by Design" concepts have been taking much more attention. From the "Security by Design "point of view, software developers must emphasize and integrate security practices within their developments tasks [7]. One of these tasks are the use of static analysis tools for inspecting their source code in order to discover weaknesses or flaws. In this sense, there are some research works proposing the use of static analysis

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 302–312, 2020.
https://doi.org/10.1007/978-3-030-56441-4_22

for GDPR [8]. In this sense, there are examples on the use of source code analysis tools [9] and its relationship with the GDPR [10].

Training is one of the weaknesses identified for the industry [11]. In fact, the industry, academia and governments are required to provide means for enhancing cyber skills to key personnel [12]. It is being evidenced that "*college graduates are forced to enroll in numerous certification programs in order to develop critical industry useful hands-on skills that are not taught in college*" [13]. So traditional training such as attending courses at universities [14], is not enough for training personnel in industry, and they are required to develop their skills in real scenarios [15]. In addition, "Awareness and Training" is a category of the NIST framework for improving cybersecurity [16]. From an academic point of view, teaching programmers and software engineers on secure practices is critical [17]. From an industrial point of view, they are being the targets of cyber attacks [18], and according to [19] 33% of companies do not provide awareness briefs or formal training to their employees. So the industry is facing training as a challenge [20]. In general, the most important approach is to follow principles for secure software development, and to ensure that software designers have sufficient security expertise [21].

There are several technologies for teaching cybersecurity [22], but industry requires a more specific environment simulating working conditions, and hands on experience. In this sense, cyber ranges such as Emulab [23] or DETERlab [24] are being considered as environments for testing industrial scenarios. In fact, Information Technology (IT) research and training can be emulated in different environments. But the problem arises when practitioners are emulating Operational Technology (OT) infrastructure. Hands-on labs are being used as testbeds [24], and some research works are focused on IT/OT infrastructure for dealing with the training [25]. Practitioners must evaluate the needs [26] and the associated costs for the training [27]. The effectiveness of trainings does not produce real tangible Return On Investments (ROI) [28], and this includes the definition of cyber range scenarios. The definition of cyber exercise such as Cyber Defence Exercises (CDX) is an expensive approach to train, test, and verify the professional skills of organisation workforce at the highest preparedness tier [29].

This paper deals with the following research questions:

- RQ1: Is there any research work related to cyber ranges?
- RQ2: Is it possible to reduce the cost of a cyber range? Is there any cost-effective approach or exercise for a cyber range?
- RQ3: Can we describe an industrial and academic exercise?

This paper is structured as follows. Firstly, it provides a background on systematic literature review and cyber ranges. Secondly, we described the systematic mapping approach highlighting the most relevant results. Thirdly, we describe how a tool is used for defining a cyber exercise with an example. Finally, the main conclusions are provided at the end of this paper.

2 Background

A cyber range is a virtual environment used mainly for training purposes. It is also used for research and development activities in order to simulate complex situations. In fact, there are several cyber ranges uses such as attack & defence, evaluate a novel technology, or to evaluate risks. In this sense, there are testbeds [30] such as SECUSIM, RINSE, ARENA, and LARIAT for simulating and testing networks scenarios [31]. In this sense, Emulab [23] or DETERlab [24] have been proposed as open source solutions. Emulab is a Network Lab Bed developed by the University of Utah Computer School. This network experiment bed consists of a series of computers, servers and routers, and a dedicated software system to manage operations. Some recent research in this field suggests platform independent mechanisms for partitioning experimental networks and breaking it down into smaller experiments that can be run sequentially [32].

Developing cybersecurity competencies through the use of Cyber Ranges, and their use for research topics is of increasing interest. In addition, Cyber Ranges include tools that help improve the stability, security and performance strength of IT/OT systems. From a physical point of view, these Cyber Ranges are composed of a host room (Security Operations Center (SOC)) which controls the Cyber Range and includes monitoring tools (IPS, IDS, SIEMS). Then there is a second room which is where the training and research exercises are performed. The scenarios must be defined and modeled in the Cyber Range. During training, the Cyber Range allows Capture The Flag (CTF) or Red-VS-Blue exercises. CTF is a challenge in which different people compete to attack a series of servers and if they succeed, they earn a series of points. In the case of Red-VS-Blue 2 teams are created, one to attack and the other to defend.

Cyber Ranges are having a major impact at the country or region level for national security reasons, and states are positioning themselves with the aim of improving their capabilities [33]. The industry is a key element in any country and begins to be threatened by the effects of attacks from cybercriminals. In-vestment in cybersecurity is increasing due to the growing real threat of cyber-crime, and it is a must for every company and Country. In order to mitigate the risks from cybersecurity, the need to investigate possible attack and defence scenarios has been identified, and for this, Cyber Ranges play a fundamental role in training and investigating these possible complex scenarios. In order to implement and adapt a Cyber Range environment, it is necessary to use appropriate software for the identified needs and to develop the appropriate scenarios to be evaluated. Therefore, this paper deals with the scenario's generation. Finally, it should be noted that computer security education and training activities are critical [34]. In fact, recent studies such as [34] suggest the use of Moodle e-learning systems for interactive training.

3 Systematic Mapping on Cyber Range

3.1 Research Strategy

In order to answer the first research question we carried out a systematic mapping following the guidelines provided by Kitchenham and Charters [35]. This paper is focused on providing an answer to RQ1, and we have followed the activities described in the following Fig. 1. which describes our systematic review strategy. This strategy starts by defining a research area and limiting the scope which has been described in "motivations and objectives" section.

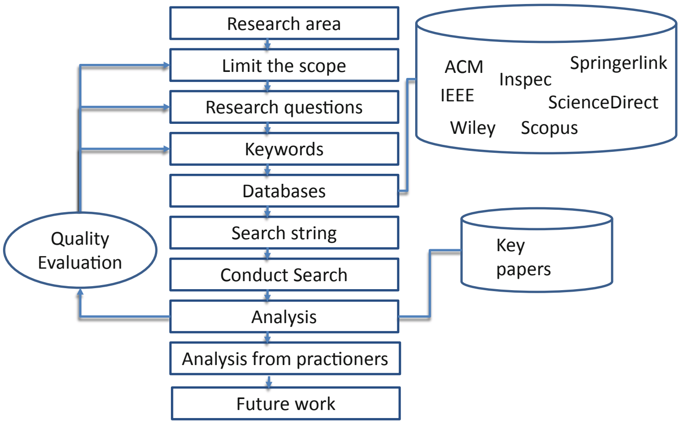


Fig. 1. Systematic approach

3.2 Data Source and Retrieval

Our strategy starts by defining a research area and limiting the scope to cyber range. This approach helps us to identify a set of research questions stated in the next section. The approach is similar to [36, 37], and a set of well-known databases are selected such as ScienceDirect (http://www.sciencedirect.com/), SpringerLink (http://link.springer.com/), IEEE Explore (http://ieeexplore.ieee.org/Xplore/home.jsp), ACM Digital Library(http://dl.acm.org/) and Wiley Online Library (http://onlinelibrary.wiley.com/).

The purpose of this RQ1 is to have an overview of the existing research works related to cyber range. Therefore, the search string is "cyber range". The search results are just limited to English papers. As result of this step we have obtained 125 papers from the data sources identified related to cyber range. The following Table 1 shows the number of papers identified per data source. We did a fist analysis and we removed the entries not related to cyber range (10). So at the end 115 papers were selected. The evolution of publications (Fig. 2) concludes that cyber ranges are having an increased interest from 2017 to nowadays.

Table 1. Number of papers per source

Source	URL	Number of papers
ScienceDirect	http://www.sciencedirect.com/	41
SpringerLink	http://link.springer.com/	5
IEEE explore	http://ieeexplore.ieee.org/Xplore/home.jsp	33
ACM digital library	http://dl.acm.org/	26
Wiley	http://onlinelibrary.wiley.com/	20

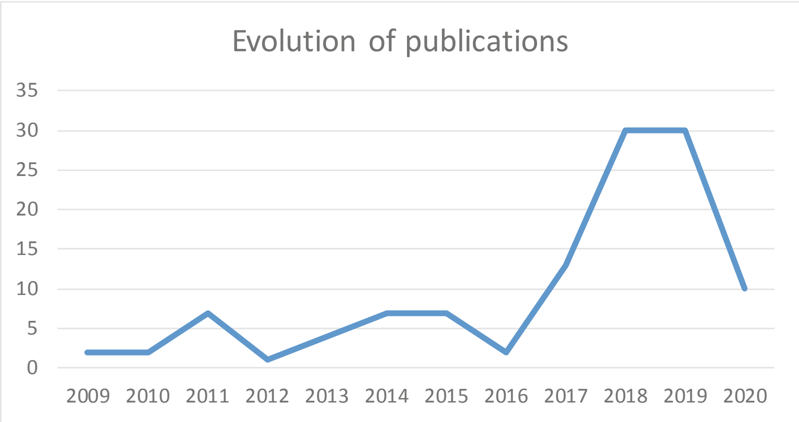


Fig. 2. Evolution of publications related to cyber range

4 Cyber Range Approach

4.1 Cyber Range Exercise Definition

According to the study carried out by [30] an updated taxonomy of a cyber range provides a meaningful description of each element involved in a cyber range. It is worthy to note that a scenario defines a storyline within an execution scenario, and the students must execute a set of steps in order to achieve training objectives. Based on the guidelines [30] we define the following scenario:

- **Purpose:** this section explains the main objectives of the execution or the experimentation.
- **Environment:** the topology of the exercise. The environment can be computer aided or can be executed without the use of any digital equipment.
- **Storyline:** this section describes the storie(s) behind the exercise. It includes the relevant actions and events within the exercise. This storyline may include a single or multiple test cases for validating new technologies.
- **Type:** This section indicates whether the scenario is static or dynamic.
- **Domain:** this aspect describes the application domain of the scenario such as network, cloud etc.
- **Tools:** This aspect describes the tools used for creating a scenario.

4.2 Case Study

The case study is chosen from the current curricula from the software quality control and assurance subject at the Bachelor of Computer Engineering in Management and Information Systems in the University of the Basque Country. However, this study can be also applied in industrial contexts in order to strengthen source code analysis capabilities. Based on the schema defined in previous section, we define the exercise:

- **Purpose:** the main purpose of this exercise is to analyse source code of a running application because it contains a set of bugs and vulnerabilities (Fig. 3).

Fig. 3. Evolution

- **Environment:** the environment is computer aided and students are asked to identify vulnerabilities.
- **Storyline:** students are required to identified hidden bugs and vulnerabilities in a running application. This application is a game and it is difficult to define Junit for testing and finding bugs because some bugs are at user interface level. Students must use Sonarqube and to define Quality Gates for identifying vulnerabilities, and bugs. Figure 4 shows a set of quality gates defined for this exercise, and students must define these rules, identify the issues and to solve them.
- **Type:** It is a static case and the source code does not change over the time.
- **Domain:** It is a stand-alone application.
- **Tools:** Basically, students are required to use their Eclipse development environment, and the Sonarqube tool as a source code static analyser.

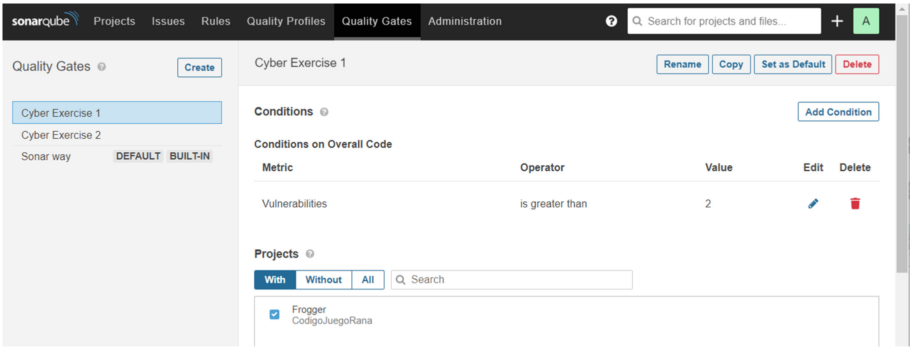


Fig. 4. Sonarqube quality gates definition

As results, Sonarqube is the cornerstone of this cyber range approach exercise because students must use and interact with this environment. In addition, Sonarqube provides a complete set of analysis, and therefore students can realize the different vulnerabilities included within this running application (Fig. 5).

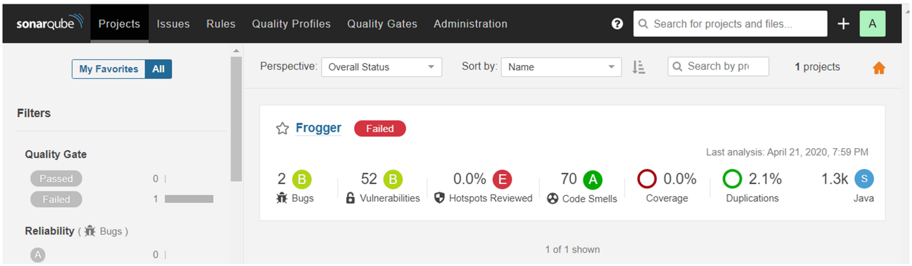


Fig. 5. Sonarqube results

Additionally, students must show the following screen (Fig. 6) and to solve the related vulnerabilities. The current Fig. 6 shows the result of the cyber exercise where the student is not able to solve the vulnerabilities included within the source code.

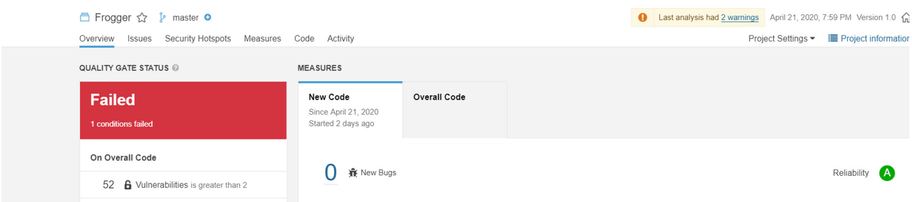


Fig. 6. Sonarqube

Additionally, for managing the exercises, we have a balanced scored card (BSC) summarizing the security status of the different projects included within the cyber range exercises (Fig. 7).

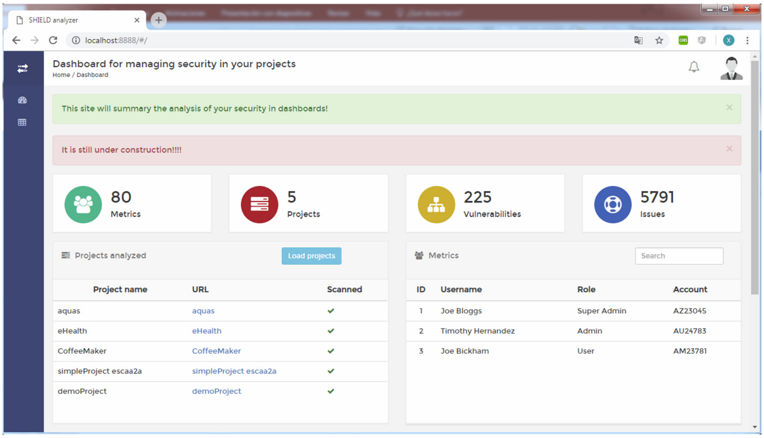


Fig. 7. Sonarqube security Balanced Scored Card

5 Discussion

This paper is structured into 2 parts. The first part (Sect. 3) deals with the study of contributions made to cyber ranges. Until now there are scarce contributions related to cyber ranges, but during these recent years, the interest on this kind of infrastructure is increasing. There are some commercial tools supporting cyber range exercises which are not described within this systematic mapping, but their cost on adoption and maintenance are too high. Therefore, we propose the use of a free tool such as Sonarqube to define and to run a cyber exercise focused on the discovery of vulnerabilities in a running application. This exercise is extracted from a University of Basque Country subject. However, this exercise can be used for training industrial workforce, because it is focused on advanced aspects such as discovery of vulnerabilities within source code.

Cybersecurity practices such as source code analysis are key elements to be considered along the deployment of any product, and they must be explicitly handled. Often, organisations do not manage security threats in a uniform and standard way. Our approach makes use of SonarQube for setting a common platform for analysing source code. Several cybersecurity related frameworks are dealing with the activities required for adding or enhancing security within applications. In this sense, the National Institute of Standards and Technology (NIST) published a Framework for Improving Critical Infrastructure Cybersecurity [16] which activities are identify, protect, detect, respond and recover, and the analysis of source code is a must.

Training is an expensive task especially in industrial environments. Companies are facing several challenges for improving the security skills of their workforce. From one side, they need a separate environment for testing and trying different solutions. They do not want to test or to inject vulnerabilities within their development environments. In this sense, Tecnalia is managing a cyber range [38] for supporting trainings. From other side, they need cost effective solutions for training people. SonarQube is envisaged as a tool for defining and running exercises related to source code analysis. We have developed a small BSC for controlling the status of the exercises.

6 Conclusions

This paper deals with the following research questions:

- RQ1: Is there any research work related to cyber ranges? We have performed a lightweight systematic mapping in order to identify research works related to cyber ranges. We were interested on knowing how exercises are defined and the current trend on this area.
- RQ2: Is it possible to reduce the cost of a cyber range? Is there any cost-effective approach or exercise for a cyber range? Basically, we are proposing the use of separate environments such as a cyber range [38], and the use of SonarQube as key tool for analysis and supervising the exercises.
- RQ3: Can we describe an industrial and academic exercise? The proposed exercise is stemming from an academic environment, but it can be used for training workforce from different companies requiring the ability to train people on source code analysis and vulnerability detection approaches.

This paper represents a first step on defining a cyber range infrastructure and a cost-effective exercise by using SonarQube. A small tool controls the exercises. As future work, we are working on the identification of new exercise to be deployed and used with a cyber range.

Acknowledgements. This work has been partially funded by the Sendai - Segurtasun Integrala Industria Adimentsurako (KK-2019/00072) del programa Elkartek del Gobierno Vasco.

References

1. Thames, Lane, Schaefer, Dirk (eds.): Cybersecurity for industry 4.0. SSAM. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-50660-9
2. Lezzi, M., Lazoi, M., Corallo, A.: Cybersecurity for industry 4.0 in the current literature: a reference framework. Comput. Ind. **103**, 97–110 (2018). https://doi.org/10.1016/j.compind.2018.09.004
3. Lorenz, B., Kikkas, K., Sõmer, T., Laugasson, E.: Cybersecurity within the curricula of informatics: the Estonian perspective. In: Pozdniakov, S.N., Dagienė, V. (eds.) ISSEP 2019. LNCS, vol. 11913, pp. 159–171. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-33759-9_13
4. The White House: National Cyber Strategy of the United States of America (2018). https://www.whitehouse.gov/wp-content/uploads/2018/09/National-Cyber-Strategy.pdf

5. MITRE: Secure Code Review. https://www.mitre.org/publications/systems-engineering-guide/enterprise-engineering/systems-engineering-for-mission-assurance/secure-code-review. Accessed 23 Apr 2020
6. The European Parliament and of the Council: Directive 95/46/EC (General Data Protection Regulation) (2016). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679
7. Larrucea, X., Santamaria, I., Fernandez-Gauna, B.: Managing security debt across PLC phases in a VSE context. J. Softw. Evol. Process (2019). https://doi.org/10.1002/smr.2214
8. Hicken, A.: Using Static Analysis to Achieve "Secure-by-Design" for GDPR. https://blog.parasoft.com/using-static-analysis-to-security-design-in-gdpr. Accessed 23 Apr 2020
9. Larrucea, X., Santamaria, I., Colomo-Palacios, R.: Assessing source code vulnerabilities in a cloud-based system for health systems: OpenNCP. IET Softw. **13**, 195–202 (2019). https://doi.org/10.1049/iet-sen.2018.5294
10. Larrucea, X., Moffie, M., Asaf, S., Santamaria, I.: Towards a GDPR compliant way to secure European cross border Healthcare Industry 4.0. Comput. Stand. Interf. **69**, 103408 (2020). https://doi.org/10.1016/j.csi.2019.103408
11. Baines, T., Lightfoot, H.W.: Servitization of the manufacturing firm. Int. J. Oper. Prod. Manage. **34**, 2–35 (2014)
12. Tropina, T.: Public–private collaboration: cybercrime, cybersecurity and national security. Self- and Co-regulation in Cybercrime, Cybersecurity and National Security. SC, pp. 1–41. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-16447-2_1
13. Fenton, D., Traylor, T., Hokanson, G., Straub, J.: Integrating cyber range technologies and certification programs to improve cybersecurity training programs. In: Auer, M.E., Tsiatsos, T. (eds.) ICL 2018. AISC, vol. 917, pp. 632–643. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-11935-5_60
14. Larrucea, X.: Modelling and certifying safety for cyber-physical systems: an educational experiment. In: 2016 42th Euromicro Conference on Software Engineering and Advanced Applications (SEAA), pp. 198–205. IEEE, Limassol (2016). https://doi.org/10.1109/SEAA.2016.28
15. Korwin, A.R., Jones, R.E., et al.: Do hands-on, technology-based activities enhance learning by reinforcing cognitive knowledge and retention? J. Technol. Educ. **1**(2), 26–33 (1990)
16. National Institute of Standards and Technology: Framework for Improving Critical Infrastructure Cybersecurity, Version 1.1. https://csrc.nist.gov/publications/detail/white-paper/2017/12/05/cybersecurity-framework-v11/draft
17. Dawson, M., Taveras, P., Taylor, D.: Applying software assurance and cybersecurity NICE job tasks through secure software engineering labs. Procedia Comput. Sci. **164**, 301–312 (2019). https://doi.org/10.1016/j.procs.2019.12.187
18. CCN-CERT: Cyber threats and Trends 2019. https://www.ccn-cert.cni.es/informes/informes-ccn-cert-publicos/4041-ccn-cert-ia-13-19-threats-and-trends-report-executive-summary/file.html. Accessed 02 July 2020
19. Make UK: Cyber security and manufacturing: a briefing manufactureres. https://www.makeuk.org/-/media/cyber-security-and-manufacturing-a-briefing-for-manufacturers.pdf
20. Thames, L., Schaefer, D.: Industry 4.0: an overview of key benefits, technologies, and challenges. In: Thames, L., Schaefer, D. (eds.) Cybersecurity for Industry 4.0. SSAM, pp. 1–33. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-50660-9_1
21. Jøsang, A., Ødegaard, M., Oftedal, E.: Cybersecurity through secure software development. In: Bishop, M., Miloslavskaya, N., Theocharidou, M. (eds.) WISE 2015. IAICT, vol. 453, pp. 53–63. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-18500-2_5
22. Salah, K., Hammoud, M., Zeadally, S.: Teaching cybersecurity using the cloud. IEEE Trans. Learn. Technol. **8**, 383–392 (2015). https://doi.org/10.1109/TLT.2015.2424692

23. Network Emulation Testbed. https://www.emulab.net/
24. Cyber-Defense Technology Experimental Research Laboratory Testbed. http://deter-project.org/
25. Morelli, U., Nicolodi, L., Ranise, S.: An Open and Flexible CyberSecurity Training Laboratory in IT/OT Infrastructures. In: Fournaris, A.P., et al. (eds.) IOSEC/MSTEC/FINSEC -2019. LNCS, vol. 11981, pp. 140–155. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-42051-2_10
26. Scaffidi, C.: What Training is Needed by Practicing Engineers Who Create Cyberphysical Systems? August 2015. https://doi.org/10.1109/SEAA.2015.19
27. Corallo, A., Lazoi, M., Lezzi, M.: Cybersecurity in the context of industry 4.0: a structured classification of critical assets and business impacts. Comput. Ind. **114**, 103165 (2020). https://doi.org/10.1016/j.compind.2019.103165
28. Kweon, E., Lee, H., Chai, S., Yoo, K.: The utility of information security training and education on cybersecurity incidents: an empirical evidence. Inf. Syst. Front. (2019). https://doi.org/10.1007/s10796-019-09977-z
29. Brilingaitė, A., Bukauskas, L., Juozapavičius, A.: A framework for competence development and assessment in hybrid cybersecurity exercises. Comput. Secur. **88**, 101607 (2020). https://doi.org/10.1016/j.cose.2019.101607
30. Yamin, M.M., Katt, B., Gkioulos, V.: Cyber ranges and security testbeds: scenarios, functions, tools and architecture. Comput. Secur. **88**, 101636 (2020). https://doi.org/10.1016/j.cose.2019.101636
31. Ferrag, M.A., Ahmim, A.: Security Solutions and Applied Cryptography in Smart Grid Communications. IGI Global (2016)
32. Yao, W.-M., Fahmy, S.: Flow-based partitioning of network testbed experiments. Comput. Netw. **58**, 141–157 (2014). https://doi.org/10.1016/j.comnet.2013.08.029
33. Fang, Binxing: Positions of states toward cyberspace and cyber-relating regulations. Cyberspace Sovereignty, pp. 243–320. Springer, Singapore (2018). https://doi.org/10.1007/978-981-13-0320-3_8
34. Beuran, R., Tang, D., Pham, C., Chinen, K., Tan, Y., Shinoda, Y.: Integrated framework for hands-on cybersecurity training: CyTrONE. Comput. Secur. **78**, 43–59 (2018). https://doi.org/10.1016/j.cose.2018.06.001
35. Kitchenham, B., Charters, S.: Guidelines for Performing Systematic Literature Reviews in Software Engineering Version 2.3. Keele University and University of Durham, Keele University (2007)
36. Nair, S., de la Vara, J.L., Sabetzadeh, M., Briand, L.: Classification, structuring, and assessment of evidence for safety – a systematic literature review. In: 2013 IEEE Sixth International Conference on Software Testing, Verification and Validation, pp. 94–103 (2013). https://doi.org/10.1109/ICST.2013.30
37. Larrucea, X., Fernandez-Gauna, B.: A mapping study about the standard ISO/IEC29110. Comput. Stand. Interf. (2019). https://doi.org/10.1016/j.csi.2019.03.005
38. Tecnalia: Tecnalia Cyber Range. https://www.cyberssbytecnalia.com/sites/cysstec.drupal.pulsartecnalia.com/files/LAB-CIBER-RANGES.pdf. Accessed 24 Apr 2020

Gamified Strategy Oriented to Decrease SPI Change Resistance: A Case Study

Gloria Piedad Gasca-Hurtado[1], María Clara Gómez-Álvarez[1](✉), Mirna Muñoz[2], and Samayda Herrera Betancur[1]

[1] Maestría en Ingeniería de Software, Facultad de Ingenierías, Universidad de Medellín, Carrera 87 No. 30-65, Medellín, Colombia
{gpgasca, mcgomez, saherrera}@udem.edu.co
[2] Centro de Investigación en Matemáticas, Avenida Lassec, Andador Galileo Galilei, Manzana, 3 Lote 7 CP, 98160 Zacatecas, ZAC, Mexico
mirna.munoz@cimat.mx

Abstract. Software development companies seek to improve their processes implementing best practices through SPI initiatives. Such initiatives look for modifying the conditions and behavior of the process stakeholders. For this reason, change resistance increases the complexity of the SPI initiatives implementation; and exist a permanent search of strategies for facing with change resistance. Gamification is an alternative for influencing employee behavior, allowing them to improve the SPI initiatives. With this motivation, this paper shows a case study of the application of a gamified strategy oriented to decrease change resistance. The methodology for the strategy definition includes a sequence of steps for selecting and associating change resistance causes, change management models, and gamification principles. The study case evidences a positive impact on causes like lack of commitment management and a general improvement in the change resistance in SPI initiatives when the gamification is the central element of the strategy.

Keywords: Change resistance · Gamification · Software process improvement

1 Introduction

Software development companies aim to improve their processes by implementing best practices to increase software product quality. Such best practices are defined by the SPI (Software Process Improvement) models like CMMI [1], Bootstrap [2], and the quality standards ISO 9000 (ISO 9001) [3], ISO 29110 [4]. These models are considered quality patterns with a vital role in identifying, integrating, measuring, and optimizing software development best practices [5].

However, software process improvement initiatives include a determinant factor associated with stakeholders for modifying their behavior conditions and generate a change in the ways of doing their daily tasks. Despite the models promote the development of high-quality software products focusing on technical, instrumental, and procedural aspects, leaving social and human factors in the background [6]. Nevertheless, social and human factors are fundamental in the implementation of SPI

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 313–325, 2020.
https://doi.org/10.1007/978-3-030-56441-4_23

strategies due to such initiatives are carried out by people constituting a key element in the software processes [7]. Such factors must be considered in the context of change resistance, influencing the complexity of software process improvement success.

Change resistance can generate unnecessary dynamics in work teams affecting this success negatively [8]. Such dynamics can cause work teams to abandon the improvement initiatives, keeping the processes or activities under the same working schemes [9]. Therefore, there is an interest in using new approaches to obtain successful SPI initiatives in software development companies. In this context, gamification is one of the suitable strategies as an interesting solution.

Gamification has the purpose of modifying and influencing people's behavior, helping with the achievement of goals previously defined. Gamification seeks to generate motivation in participants becoming in an interesting alternative to face SPI aspects included in the agile manifesto. Specially, gamification can contribute to the principle: "Motivate all people involved". Also, gamification has shown results associated with the principle: "Use dynamic and adaptable models as needed", helping to assume changing needs in organizations. Thanks to gamification benefits, this paper describes a case study using a gamified strategy oriented to decrease change resistance in SPI approaches. Such a strategy implies an association between a) change resistance causes, b) change management models and c) gamification principles. Such elements are the result of literature exploitation and the use of the transitive law.

The methodology used for strategy design is an adaptation of the approach of the School of Mines – Universidad Nacional de Colombia [10] adequate for qualitative research to solve research problems, proposing a solution based on theoretical fundamentals and effectively validating this solution.

This proposal includes the development of a software platform as a validation instrument. As an empiric validation method, a case study in an IT company is designed. Such a case study allows measuring the decrease in change resistance by actors of the improved process. Also, it is possible to highlight a positive impact on causes like lack of commitment management, among others.

The paper is structured as follows: Sect. 2 presents background about using gamification in SPI initiatives. In Sect. 3, we present the gamified strategy proposal and the methodology used to design it. In Sect. 4, we describe the case study applied in an IT enterprise. Section 5 presents the case study results associated with decrease change resistance in the improved process and the feedback of a perception questionnaire. In Sect. 6, we discuss the conclusions and future work.

2 Background

A set of recent studies, between 2018 and 2020, shows an increasing interest in solving SPI challenges using gamification as an improvement strategy. One of the most relevant is the proposal by Herranz *et al.* [11], who define a gamification framework for deployment SPI efforts to increase motivation in software employees. Such a framework was implemented in Spanish software development organization, and the results showed that this framework contributes to enhancing the SPI task performance.

However, most studies are oriented to systematic literature reviews to understand the current state of gamification as a strategy for improving processes associated with software project management (SPM). In [12], the authors present a review of using gamification in SPM and some results show: (1) the most commonly reported research method is the solution proposal and the most commonly used gamification element is the point system, and (2) a predominance of studies in project management areas related to integration, resources, and scoping. Future research could be addressed to use gamification in unexplored project management areas for facilitate the implementation of good practices in software development projects. Alhammad *et al.* [13] present other review whose results evidence that the purpose of applying gamification in software engineering (SE) is mostly related to improving student engagement and, to a lesser extent, to improving student knowledge. The authors also discussed aspects like the implementation cost of gamification, gamification elements more used in SE efforts and the complexity of deciding which gamification approach to follow for lack of information to develop this activity.

Also, we identify a similar study focused in social factors in gamification approaches presented by Herranz and Colomo-Palacios [14] who defined SPI enablers and barriers and some pro, cons of gamification, and show some of the anti-patterns but also the enablers of a set of initiatives conducted in the joint field of SPI and gamification. Finally, Muñoz *et al.* [15] present a proposal of using a contraption based on the Rube Goldberg machine to gamify the identification of software development team roles focusing on knowledge, skills, and interactive styles toward the integration of high-performance teams.

3 Gamified Strategy Proposal

Gamified strategy proposal was designed starting from the association of change resistance causes, change management models, and gamification principles. The definition of such an association is a result of the application of a methodology described below [16].

3.1 Methodology for Selection and Association of Change Resistance Causes, Change Management Models and Gamification Principles

This methodology includes three phases: 1) Contextualization, 2) Association, and 3) Synthesis. Table 1 presents a summary of such phases and activities required to achieve such an association.

Table 1. Methodology activities description

Phase/Activities	Artifacts	Techniques
Phase 1. Contextualization a) Select and unify change resistance causes b) Categorize unified causes c) Select and describe change management models d) Identify and adopt gamification principles	– Categories of unified causes – Documentation of selected models – List of gamification principles	– Systematic literature review – Systematic mapping – Terms Unification
Phase 2. Concepts association e) Associate causes categories and models f) Associate causes categories and gamification principles g) Use Transitive Law to associate models and gamification principles	– Causes categories and models – Causes categories and gamification principles – Models and gamification principles (association)	– Experts judgment – Transitive Law
Phase 3. Information synthesis h) Collect information about concepts association i) Consolidate information	– Information synthesis document	– Statistical tools – Information extraction templates

As a result of this methodology, we obtain an information extraction template as a consolidation mechanism. Table 2 shows an extract of such a template.

Table 2. Extract of information synthesis template

Causes categories	Suggested model	Suggest principles
Communication factors	Prosci change management model	6. Amusement factors 2. Persuasive elements
	Moitra model	
Organizational factors	13 steps model	8. Well-being oriented 7. Transformative
	Moitra model	
Project management factors	Prosci change management model	3. Learning orientation
	Moitra model	

3.2 Design of a Gamified Strategy Oriented to Decrease SPI Change Resistance

From the association established (See Table 2), the gamified strategy was designed. Figure 1 summarizes the steps for implementing the strategy. Such steps are defined based on the results of the methodology for selection and association of change resistance causes, change management models and gamification principles (Sect. 3.1)

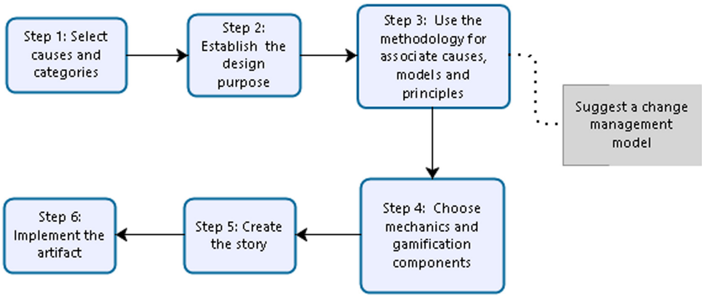


Fig. 1. Steps for implementing the gamified strategy

The gamified strategy described below is the result of the application of the steps showed in Fig. 1. The strategy goal is mitigating change resistance causes in software development processes. The implementation of this strategy requires identifying an ongoing and measurable process with tasks assigned to a responsible. Some of these processes are incident management, project management, or testing.

Below, the steps followed for the gamified strategy design are described.

Step 1: Select SPI change resistance causes and their categories.
The strategy to be developed will be as a goal mitigating change resistance causes related to commitment factors, knowledge factors, and predisposition for change factors.

Step 2: Establish a strategy purpose.
The strategy aims to strengthen the team's knowledge of the process to be intervened, to motivate them to develop the different activities of the project appropriately, and to integrate the work teams around the intervened process.

Step 3: Use the methodology for associate change resistance causes, change management models and gamification principles.
From the selected causes, the application of the methodology results in the association presented in Table 3.

Table 3. Association of change resistance causes, change management goals and gamification principles

Change resistance changes	Change management models	Gamification principles
Commitment factors Knowledge factors Predisposition for change factors	– Mathiassen model	– Orientation – Amusement factors – Persuasive elements – Learning orientation – Knowledge-based – Achievement based on rewards – Transformative – Well-being oriented

The change management model selected is the Mathiassen approach encompassing at the same time, the three factors or categories of change resistance causes. The gamification principles presented, correspond according to expert assessment, with the principles with influence equal or superior to 80% over the change resistance factors.

Step 4: Choose mechanics and game components related to gamification principles suggested as a result of methodology application.

The suggested gamification principles after the methodology application are presented in Table 3. Considering such principles, game mechanics for implemented the gamified strategy are selected. The strategy goal is to mitigate the factors or categories of change resistance causes. From a previous list of gamification elements, the authors selected the following: challenge, competency, feedback, rewards, incentives, and cooperation, for being included in the strategy design.

The implementation of game mechanics is briefly described below:

- *Challenges* were implemented through missions with an increasing complexity level.
- *Rewards* and *incentives* awarded in the strategy were points for accessing virtual goods, gifts, or badges.
- *Competency* would be motivated through position boards where the player could ascend thanks to winning points.
- *Feedback* and *cooperation* were incorporated into the strategy through the solution of missions.

Step 5: Create the story

The following is a story involving the different mechanisms suggested for the gamified strategy.

Interplanetary Exploration

Welcome to interplanetary mission Cassiopeia X64, we are over the Great Magellan Cloud, and we are going to the central station, Diorite IV. We are looking for the Kappirium mineral, the raw material for our fuel, and technology, which we will be able to extract from the Teckitz we have found scattered throughout our journey.

You are part of our scientific expedition and will be in command of a small space exploration ship, to collect as many Teckitz as possible to achieve our mission.

Our scientists have recorded the following information about the Teckitz:

It is a precious metal that can be used as fuel; it is obtained from the record of activities of a given process.

At the current speed of research, we must process the accumulated Teckitz every 15 days, which we have called the "technocratic cycle", which is why your Teckitz counter will be reset at the beginning of each cycle. Still you will retain the accumulated experience throughout the mission.

With the accumulated Experience you collect, you will be able to

- *make improvements to your interstellar ship*
- *Scale in a polytechnic grade of the fleet*
- *Accumulate badges*
- *Scale in technocratic cycle ranking (month to month)*

Step 6: Implement the artifact.
A software platform supports the gamified strategy implementation according to the previous story. The architectural design follows the Model-View-Controller Pattern (MVC) [17]. In [18], the authors present the different architectural views (logical, development, process, and scenarios view) of the software platform to face the various stakeholders' needs.

4 Study Case

The validation of the gamified strategy proposed in this paper was carried out in the IT area of the @Medellin project. @Medellín is a project born from the synergy of the Higher Education Institutions (IES) attached to the municipality of Medellín. The project aims to integrate the virtual academic offer of the different institutions. The organization has an IT area responsible for managing, developing, and supporting the project's various digital platforms.

This organization begins its improvement process in the last trimester of 2018 and has shown results in the incidents management process associated with software product development. However, there are some change resistance factors related to the following causes: (1) Lack of commitment management, (2) Lack of knowledge or confusion about the process, and (3) Inertia associated with the previous process.

The organization and the authors identified such causes and categorized them on the factors defined in this paper. Thus, the causes correspond to factors of commitment, knowledge, and predisposition to change.

From the identified causes, the implementation process of the "Interplanetary Exploration System" strategy begins. The analysis of the results obtained in the metrics of the process and the users' perception will be the strategy's validation variables. Below validation plan is presented.

To provide the necessary guidelines for implementing the gamified strategy in the @Medellín organization, the implementation of the strategy will be carried out by executing the following validation plan

1. Analyzing the process and support tools
   a. Define the organizational profile and target population (user selection)
   b. Identify the process improvement implications
   c. Define the process baseline, analyzing current process improvement status

2. Analyzing the process to be intervened
   a. Identify process variables and metrics
   b. Assess the strategy according to users' perception
3. Implementing the strategy
   a. Integrate the software platform with the support tools of the organization
   b. Present the strategy to stakeholders
   c. Define the strategy implementation parameters like duration, rewards, among others.
4. Applying the strategy
   a. Define the use period of the strategy application
   b. Monitor the process
   c. Consolidate of strategy evaluation results
   d. Present and share results

The different elements of the gamified strategy validation plan are summarized below:

4.1 Analyzing the Process and Support Tools

- *Organizational profile*: @Medellín is a project that aims to integrate the virtual academic offer of the different higher education institutions. The organization has an IT area responsible for managing, developing, and supporting the different digital platforms of the project.
- *The process to be intervened*: The process of interest is an incident management process in software product development.
- *Process improvement implications:* Some of these implications are new communication channels implementation, the publication of frequently asked questions for helpful user attendance, and platform updates.
- *Process improvement status:* The improved process of incident management was implemented according to the findings of an interview with the leader of the IT area. However, some change resistance causes were evidenced, such as lack of commitment management and lack of knowledge about the process.
- *Support tools:* The support tool for the incident management process is OsTicket. Such is a web tool under General Public License (GNU) developed in PHP language and a MySQL database.

4.2 Analyzing the Process to Be Intervened

- *Process variables and metrics:* The metrics for assessing change resistance mitigation are:
  – Number of tickets register on the OsTicket platform
  – Completeness of the registered tickets description
  – Compliance in response times according to Services Level Agreements (SLA)
  – User perception about the gamified strategy and software platform "Interplanetary exploration"

4.3 Implementing the Strategy

- Integrate the software platform with OsTicket: In this activity, the equivalence between ticket solved and Teckitz is defined through rules for Teckitz assignation to the participants according to their performance in the incident management process in OsTicket tool (solution time, priority, etc.).
- Strategy implementation parameters:
  – Target population: technical support teams' level 1, in charge of attending medium and low priority incidents, and level 3 in charge of attending high priority incidents and requests from partner institutions.
  – Period of application: The interplanetary system will be available by 15 days.
  – Reward: a gift card for the winner upon completion of the strategy

4.4 Applying the Strategy

The results associated with the gamified strategy are presented in the next section.

5 Case Study Results

The platform for implementing the gamified strategy is "Interplanetary Exploration" which is integrated with the organization incident management system (Osticket). The implementation period was from March 15 to March 30 of 2019.

The proper development of the registration and ticketing activities of the incident management process in software products and its documentation on the OsTicket tool provided the fuel (Teckitz) needed to fulfill the proposed missions, obtain points, move up the ranking table and obtain the rewards.

Also, the data analysis of the incident management process consisted in comparing the incident management reports of the two months before the implementation of the strategy and the software platform "Interplanetary system" versus the information obtained during the implementation period of the strategy. The variables analyzed correspond to the metrics defined in the previous section. Finally, a questionnaire for user perception was applied to the gamified strategy participants.

5.1 Process Metrics

Concerning the *Number of registered tickets,* we can see the results in Fig. 2. Such number had a considerable increase, showing that in the strategy implementation period was registered the 88% of the total number of tickets registered during the whole month of January and February were obtained. Such Figure evidences the number of tickets closed or solved during the strategy implementation.

About the *Completeness of the registered tickets*, a previous analysis of 50 tickets of January and February allows identifying difficulties such as the tickets only has a general categorization (two options: platform access problems and others). This fact doesn't help the process responsible; the ticket documentation is oriented to the user but does not have detailed information for monitoring the solution process. Most tickets are solved by the same responsible. After the strategy implementation, we can observe a better ticket categorization according to Fig. 3. These results also show a better distribution of tickets solution among the different participants increasing the process commitment level by the strategy participants.

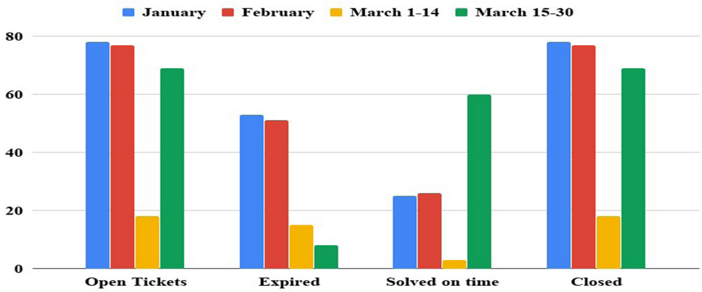
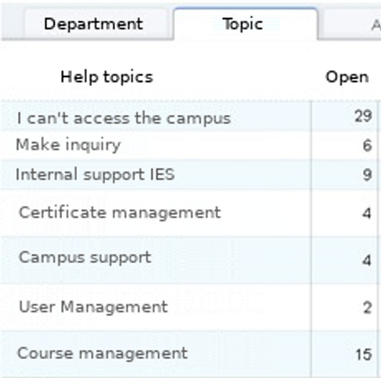


Fig. 2. State of registered tickets



Help topics	Open
I can't access the campus	29
Make inquiry	6
Internal support IES	9
Certificate management	4
Campus support	4
User Management	2
Course management	15

Fig. 3. New tickets classification

5.2 Participants Perception

Participants' satisfaction was measured using a questionnaire, which is an instrument that consists of a set of questions related to variables that are desired to be measured during the research [19]. The questionnaire's purpose is to know the users' perception of the gamified strategy and the Interplanetary system software platform.

This questionnaire was based on [19] and included three sections: 1) perception of the strategy; 2) individual perception, and 3) perception of the experience and satisfaction.

Five users who participated in the strategy and the leader of the IT area of the @Medellín project filled out the questionnaire. Some of the analyzed variables were *enjoyment level*, which obtained an acceptable and good score with percentages of

66.7% and 33.3%, respectively indicating it is necessary to make improvements on the strategy or platform, providing more emphasis on the gamification principle amusement factors.

The variable *difficulty level* is associated with process knowledge by the participants since the success or failure of the missions depends on this knowledge. Users' perception of this variable is different, with 16.7% considering it acceptable, 33.3% considering it good, 16.7% considering it very good, and 33.3% considering it excellent. These results indicate that a greater effort is needed to include gamification principles such as learning orientation and knowledge-based principles in the gamification strategy. An alternative is adding a greater variety of missions and improving the feedback provided in the different missions.

The variable *material quality* focuses mainly on knowing the perception of the users concerning the platform. In general, the users assigned the value good and excellent with a percentage of 50% in each value, this fact allowed to carry out the pilot in a stable way and to fulfill the established goals. However, the users suggested modifying the conversion scale of Teckitz and the implementation of a more attractive design with 3D objects and animated elements.

5.3 Threats to Validation

We identify two threats to validation:

(1) The study case included five participants responsible for the incidents management process in an IT enterprise. The sample correspond with the actors of such process. In future strategy applications we look for have a bigger sample for reinforce the impact of results obtained.
(2) The observation time of the strategy was 15 days by effects to present the first results. However, the gamified platform is still using in the pilot organization and we can collect new results for out coming publications.

6 Conclusions and Future Work

This paper shows a gamified strategy applied in the IT area where the improvement of the incident management process in software products was carried out. The process presented causes of resistance to change, such as lack of commitment management, lack of knowledge, and confusion about the process, inertia, and habit in previous processes. After the implementation of the strategy, a decrease in the change resistance causes was evidenced by the increase of 88% in the number of tickets resolved according to the Services Level Agreements. In causes like inertia and habits in previous processes, an increase of 88% in the number of tickets registered on the OsTicket platform in comparison with previous months could be identified, and the elimination of support through other means such as email was achieved. In causes like ignorance and confusion in the process, the completeness of the tickets was improved, achieving a better answer description and making use of internal notes for improving their traceability. Also, there is an improvement in 1) the ticket categorization, and 2) the assignment of the tickets to the professionals in a more balanced way.

The software platform “Interplanetary Exploration” performed well during the strategy implementing, allowing the interaction of the participants and correct with the OsTicket system, is obtained through the perception questionnaire and in direct communication with the participants’ feedback.

In future work, we will: 1) improve the software platform about the reward assignment system, the missions’ variety and the integration with other organizational support tools, 2) apply the strategy in other improved software process with a larger sample, and 3) increase the number of gamification principles included in the gamified strategy.

References

1. C. P. Team. CMMI for Development, Version 1.2, CMMI-DEV v1. 2, Softw. Eng. Institute, Carnegie Mellon Univ (2006)
2. Bicego, A., Khurana, M., Kuvaja, P.: Bootstrap 3.0 — software process assessment methodology. In: Hawkins, C., Ross, M., Staples, G. (eds.) Software Quality Management VI. Springer, London (1998). https://doi.org/10.1007/978-1-4471-1303-4_3
3. Hoyle, D.: ISO 9000 Quality Systems Handbook: Using the Standards as a Framework for Business Improvement. Routledge, New York (2009)
4. O’Connor, Rory V., Laporte, Claude Y.: Software project management in very small entities with ISO/IEC 29110. In: Winkler, D., O’Connor, Rory V., Messnarz, R. (eds.) EuroSPI 2012. CCIS, vol. 301, pp. 330–341. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-31199-4_29
5. De la Villa, M., Ruiz, M., Ramos, I.: Modelos de evaluación y mejora de procesos: Análisis comparativo. In: 5th ADIS Workshop (Apoyo a la Decisión en Ingeniería del Software). Málaga, España (2004)
6. Ferreira, M.G., Wazlawick, R.S.: Complementing the SEI-IDEAL Model with Deployers’ Real Experiences: The Need to Address Human Factors in SPI Initiatives (2010)
7. Guterres, M., Sidnei, R.: Software process improvement: a organizational change that need to be managed and motivated. Comput. Inf. Eng. **5**, 134–142 (2011)
8. Muñoz, M., Mejía, J., Calvo-Manzano, J., Cuevas, G., San Feliu, T.: Evaluación de los Procesos Software de una Organización Enfocando en la Disminución de la Resistencia al Cambio. Comput. Sci. Eng. **2**, 66–73 (2012)
9. Herranz, E., Colomo-Palacios, R., De Asmecua, A., Yilmaz, M.: Gamification as a disruptive factor in software process improvement initiatives. J. Univers. Comput. Sci. **20**, 885–906 (2014)
10. M. en I. Q. MIQ, Plan de trabajo de la tesis de maestría. Universidad Nacional de Colombia. Sede Medellín (2005)
11. Herranz, E., Guzmán, J.G., de Amescua-Seco, A., Larrucea, X.: Gamification for software process improvement: a practical approach. IET Softw. **13**(2), 112–121 (2018)
12. Machuca-Villegas, L., Gasca-Hurtado, G.P.: Gamification for improving software project management processes: a systematic literature review. In: Mejia, J., Muñoz, M., Rocha, Á., Peña, A., Pérez-Cisneros, M. (eds.) CIMPS 2018. AISC, vol. 865, pp. 41–54. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-01171-0_4
13. Alhammad, M.M., Moreno, A.M.: Gamification in software engineering education: a systematic mapping. J. Syst. Softw. **141**, 131–150 (2018)

14. Herranz, E., Colomo-Palacios, R.: Is gamification a way to a softer software process improvement? A preliminary study of success factors. In: Larrucea, X., Santamaria, I., O'Connor, Rory V., Messnarz, R. (eds.) EuroSPI 2018. CCIS, vol. 896, pp. 207–218. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-97925-0_17
15. Muñoz, M., et al.: Gamification to identify software development team members' profiles. In: Larrucea, X., Santamaria, I., O'Connor, Rory V., Messnarz, R. (eds.) EuroSPI 2018. CCIS, vol. 896, pp. 219–228. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-97925-0_18
16. Gasca-Hurtado, G.P., Gómez-Álvarez, M.P., Herrera, S.: Resistencia en la mejora de procesos de software: Una propuesta de clasificación de causas, modelos de resistencia y gamificación. In: 14th Iberian Conference on Information Systems and Technologies (CISTI). IEEE (2019)
17. Pourkomeylian, P.: Software process improvement strategy. In: ECIS, 2006, pp. 330–339 (2006)
18. Gasca-Hurtado, G.P., Gómez-Álvarez, M.C., Herrera, S.: Herramienta gamificada para mitigar causas de Resistencia al cambio en mejora de procesos software. In: 2020 15th Iberian Conference on Information Systems and Technologies (CISTI). IEEE (2020)
19. Gómez-Alvarez, M.C., Gasca-Hurtado, G.P., Manrique-Losada, B., Arias, D.M.: Method of pedagogic instruments design for software engineering. In: 2016 11th Iberian Conference on Information Systems and Technologies (CISTI), pp. 1–6. IEEE (2016)

Assessing Application Lifecycle Management (ALM) Potentials from an Industrial Perspective

Zuleyha Akgun[1], Murat Yilmaz[2,3](✉), and Paul Clarke[2,3]

[1] Akgun Software Inc., Ankara, Turkey
zuleyha.akgun@akgunyazilim.com.tr

[2] Lero, the Science Foundation Ireland Research Center for Software, Limerick, Ireland

[3] Dublin City University, Dublin, Ireland
{murat.yilmaz,Paul.M.Clarke}@dcu.ie

Abstract. The notion of software lifecycle management requires the incorporation of effective software engineering processes that encompass not only development but also deployment and maintenance. Although software engineering methodologies continue to emerge, software development continues to experience a significant amount of failure rates. To overcome the complexities of modern software development, we envision that software developers should familiarize themselves with emerging tools to make required changes incrementally by rigorously tracking deployment. Application Lifecycle Management (ALM) is an emerging trend that reflects such a novel attitude towards software development, which blends the tasks performed by a firm's software development and systems operations teams, monitor software artifacts from inception to completion while improvements are documented. The outcome of this study is to provide guidance for practitioners in tailoring ALM practices from the idea phase through the selection of the underlying toolset. Ultimately, an assessment at an industrial scale is conducted to manage the ALM transformation using in a large-scale corporate environment. Overall, these results indicate that participatory action research is a robust approach to investigate complex software engineering issues.

Keywords: Application Life-Cycle Management · Software life-cycle management tools · Action research

1 Introduction

Software projects are complex structures that contain many scopes and are carried out on the basis of a regular process [1]. This process starts with gathering the requirements and continues until reaching the end-user. It includes in itself the stages of planning, gathering, analysis of requirements, design, production, testing, maintenance and distribution. It is the most important and required point during the negotiations with the stakeholder to have an explicit knowledge of the project scope, quality, cost and schedule elements [2].

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 326–338, 2020.
https://doi.org/10.1007/978-3-030-56441-4_24

Application Life-Cycle Management (ALM) is an approach covering all processes of a software project output from idea to development, distribution, and maintenance [3]. This provides powerful reporting and tractability capability throughout software development period [4]. The main purpose of ALM is to improve software quality and seek ways to reduce the development costs [5]. It comprises some very important building blocks. For example; it encompasses selected methodologies that enable the management of such concepts as the management of claims, software architecture, design, coding, testing of the product, review of the live processes, product improvement as a result of the review, and ultimately releasing new versions as a result of such improvements [6].

Different tools that include different ALM processes (e.g., planning, development, design, testing, maintenance, etc.), may exchange information and interoperate with the tools that manage other processes [7]. The ALM approach covers all processes of requirements management, test management, building management, source code management, project management and risk management [3]. Carrying out all these different processes as integrated with each other is very important in the successful conclusion of a software project and in determining the quality level [8]. There are many ALM tools used in our days, such as IBM Rational Team Concert (IBM JAZZ), Microsoft Team Foundation Server (TFS), HP Quality Server, and Atlassian JIRA.

The specific objective of this study was to review the product lines of AKGUN Software Inc., (i) to ensure that a rigorous transformation roadmap was implemented in AKGUN Software Inc. and that the production lines run in an integrated way during the product lifecycle, (ii) to digitalize the processes and (iii) to enable a better ALM transformation. This study aimed to address the following research question: *Can we systematically manage the ALM transformation process of a large-size software development organization by conducting action research?*.

The overall structure of the study is structured as follows. Section 2 reviews the literature on ALM practices. Section 3 describes the actions research method that was used in this study and summarizes the main findings. Finally, the conclusions are drawn in the last section.

2 Background

Herden et al. [9] define Application Life-Cycle Management (ALM) *as the set of processes, methods and tools used by the IT departments of an organization to ensure the smooth operation of the critical business practices and adaptation to new requirements*. Chappell [10], emphasizes that the definition of Application Life-Cycle Management is not an easy task and that the ALM is of great importance in project management. Software methodologies can be thought of as a small part of ALM which supports the life of an application with tool and process.

Kravchik [11] argues that the ALM platforms i.e, IBM Jazz and Team Concert, Microsoft Team Foundation Server and Visual Studio Team System and

Comverse DiME. Each platform has identified a set of details including breadth of lifecycle support, integration, role-based views, traceability and reporting, platforms support, scalability, security, incremental implementation, integration and interoperability with existing solutions, success stories etc. McHugh et al. [7] confirm that as ALM structure has matured, two focal areas have been formed by leading vendors in this field as the integration of the ALM segment tools and improvement works of the quality processes. In addition, it is argued that adopting the right process for the efficient use of the ALM is an important step. This study also claims that the ALM process is the best approach for complex and large-scale projects with an experienced team.

Herden et al. [9], claim that ALM technology status includes paper-based, manually-operated, guided procedures to a larger extent. Schwaber [12] points out that IT organizations spend billions of dollars annually on the development life-cycle tools, which do not provide them with a proper development and state that for most of the companies, coordination of the application life-cycle management (ALM) activities that still remains to be manual to a large extent. While emphasizing that the recent ALM tools may not provide support beyond what is achievable with the integration of the ALM from tool to tool, it is argued that future ALM platforms will get better by providing joint service with the tools.

It is emphasized that Application Life-Cycle Management includes all processes related to the development, growth of the application, its customization in line with the feedback from the customer and its maintenance for a longer lifetime. For these matters, there is a need for a continuous process and the ALM is a concept that fulfills this requirement and the primary objective of the ALM structure is to increase the software quality and reduce the costs [13].

To sum up, an increasing number of software development organizations have started to adopt the ALM process, and therefore a number of ALM tools have been proposed [14]. Although recent developments in ALM have heightened the need for more attention from software engineering researchers, there is little published data on ALM [3,15]. However, this paper aims to enhance our understanding of the industrial adoption of ALM practices in a large-sized software company. Therefore, this study makes a major contribution to research on ALM by demonstrating a rigorous empirical assessment for exploring the implications of ALM tools and practices by (i) exploring the possible benefits of implementing ALM in large-sized software development environment, (ii) investigating the potential constrains, (iii) and ultimately documenting the experience gained during the transformation process.

3 Industrial Assessment

In the course of past twenty years, action research approach has been greatly used in social science and adopted in software engineering researches for information systems. Baskerville [16] argues that, these four common characteristics are shared in action research studies: (1) An action and change orientation (2) A problem focus (3) An organic process involving systematic and sometimes iterative stages (4) Collaboration among participants.

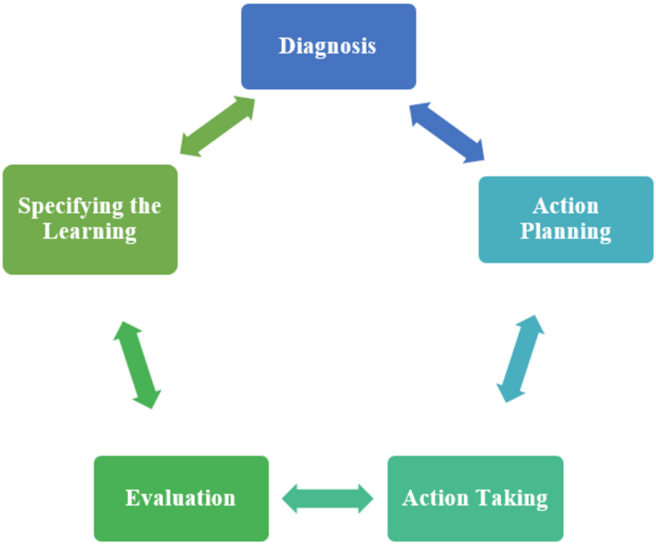


Fig. 1. Action research cycle

The Fig. 1 below shows the five basic steps of the action research:

1. **Diagnosis**: The primary and the most significant problems are identified in this step, which drive the organization to make a change. The formulation of the problem domain of the organization and the theoretical assumptions about it are made in this step.
2. **Action planning**: The action plans for the identified problems in diagnosis step are made in this step. We determine where we want to see the organisation at the end of the action research and we create a list of actions in order to achieve this.
3. **Action taking**: The planned actions in the previous step are executed in the action taking step.
4. **Evaluating**: The results of actions taken in the action taking step are evaluated in this step and the theoretical effects of the actions are evaluated by the studies to determine if they are accomplished.
5. **Specify the learning:** Whether or not the implementation is successful, the lessons learnt from the cycle has to be identified, which will be used for the determination of the prospective future cycles.

The action research on the impact of ALM adoption has been carried out in AKGÜN Software Inc., which has over 500 IT professionals, provides solutions and a wide range of products from IT, healthcare, education, telecommunication, government, defense, and the private sector. The company has a wide array of project experience in areas such as in R&D, Manufacturing Software, System Integration, Training and Support Services, Network, Communication and Information Security, Technology Consulting, Medical IT Consulting Services.

3.1 Action Research Implementation

Action research steps that have been taken consisted of diagnosis, action planning, action taking, evaluating, and specifying the lessons learned.

3.1.1 Diagnosis

To be able to carry out an effective diagnosis, the relevant groups and key personnel to carry out the works were identified within the company. The situation was assessed in terms of organizational, process and tool perspectives as a result of the reviews of the projects and interviews, meetings with 6 groups in total including Image Processing Group, Clinical Information Group, Web Technologies Group, Mobile Technologies Group, Integration Group, Decision Support Systems Group and the key personnel in these groups.

The results of the examinations were as follows from three perspective;

Organizational Perspective: It was observed that some of the employees in the organization, preferred different tools on the basis of each project, and some groups were still using their old working methods. With this study, it was ensured that they are turned into a part of the system and that they are working harmoniously.

Process Perspective: The company has been developing small, medium and large-scale projects for the last 20 years and mainly uses waterfall, agile and iterative software methodologies in significant number of its projects. In addition, a process tailoring roadmap was found essential in order to establish the requirements of these processes and the relevant documentation in all projects and a continuous monitoring and control mechanisms have been planned for the implementation.

Tool Perspective: The tools that support the processes implemented in the organization were found in more than one and in different in many projects. While some projects demanded manual documentation, other projects were likely use a set of tools with different features for the same processes (see Fig. 2).

3.1.2 Action Planning

Within the scope of the planning of the actions, firstly the focus was on the solution of the problems that were determined in the diagnosis phase and were identified as a result of the meetings held. In the action plan, we aimed to evaluate the process and tool perspective together. Correct tool selection and adaptation were taken into account. Several alternative tools have been investigated. One of the most important criteria in choosing a tool was to be compatible with our CMMI processes and to ensure *traceability* between projects and processes [17].

The requirements have been determined in the first place to select the tools and create the best tool set, the best possible alternatives have been identified, and then interviews with appropriate suppliers have been carried out. A committee has been established within the company to carry out the action planning. Most importantly expectations from the tool have been identified which was also

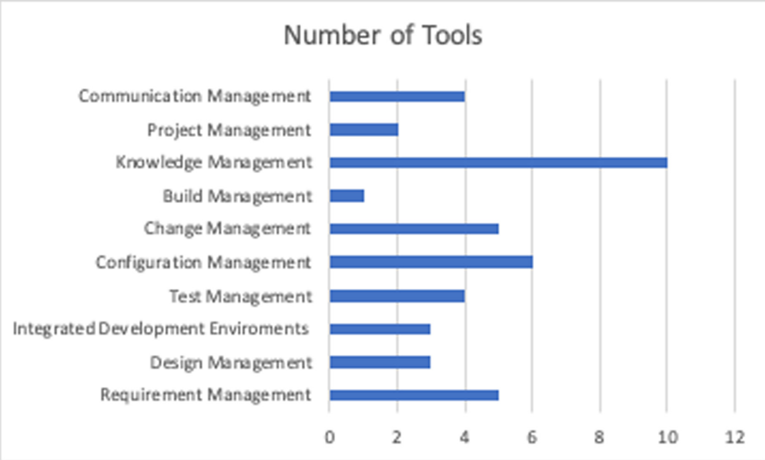


Fig. 2. Product life-cycle tools

considered as an input to the tool suppliers assessment and technical specifications of contracts with suppliers.

3.1.3 Action Taking

At this stage, primarily the analysis of the tools was carried out. In this context, a significant number of ALM tools have been examined.

Analysis: To ensure more effective management of action research studies within the scope of ALM and achievement of its goal, it has been decided to select a *requirement* and *test management* tool, which was able to operate the analysis and test processes that were already in use in a more efficient and effective manner, which were also expected to satisfy the CMMI requirements, in the first iteration.

The application development lifecycle management (ADLM) focuses on the tool market, software development lifecycle (SDLC) planning and governance activities. ADLM products focus on the "development" part of the life of an application. In recent tears, research criteria have been changed to measure the agile development and DevOps support in a robust manner.

The key elements of a solution include:

- Software Requirements Definition and Management
- Software Change and Configuration Management
- Item Management by Focusing on Software Planning, Agile Planning
- Quality Management including Error Management

The tools and equipment within the scope of ALM, which can meet these needs, have been researched and it was seen that there are many alternatives developed in this regard. In the case of software projects, specific application lifecycle management tools were used to ensure successful project management and successful results.

A review of the Gartner Survey has been conducted to assist ALM providers, application development managers, and other IT leaders to select existing technology partners in the industry where they work [18]. This research is based on the most preferred application lifecycle management tools used today and provides information about the tools. A world-wide enterprise that researches technology tools, Gartner uses a method called Magic Quadrant for these operations. The answer to which players in the market are superior to their competitors and in which areas they are superior is looked for by means of this method. By means of the method, all tools in any market can be compared. Research conducted by Gartner using this method is published on their own website[1].

The research report published on the Gartner official website has a detailed list of the tools used [8]. Companies in this list: Atlassian, IBM, Microsoft, Rally (now CA Technologies), Micro Focus, VersionOne (CollabNet), Jama Software, Perforce Software, CA Technologies, Rocket Software, CollabNet, PTC, Siemens, SUSE, Kovair Software, Beesion, Inflectra, Digite.

The products of the companies in the list are listed in the table and in this context; the information of ALM provider company names, products, tool score, license type and tool's appraisal score is detailed.

According to the results of the research conducted by Gartner and as can be seen in the Fig. 3, Microsoft, IBM and Atlassian are listed among the most popular ALM Tools.

Assessment and Acquisition: Based on the assumption that taking all of the above mentioned tools, which have been obtained during the analysis stage, among the alternatives may cause unnecessary studies and loss of labor force and at the same time, not all of them will meet the requirements and quality, these tools have firstly been researched on the internet and, if necessary, demonstrative versions have been downloaded and examined. The tools that are not qualified to meet our requirements have been identified and eliminated. As a result of the studies and examinations, it has been decided to consider the five ALM tools, which are likely to meet the requirements. As a result of this investigation, it has been decided to detail the works within the scope of each tool.

The following action plan has been developed during the detailed study of each tool;

- Determination of the user license requirements and usage frequencies,
- Obtaining contact information of ALM tool suppliers,
- Installation of demonstration versions of the tools,
- Examining of the tools by the technical evaluation committee of our company in the first place,
- Submitting a list of our user needs to the ALM suppliers and requesting their explanations for meeting each one of our needs,

[1] https://www.gartner.com/en/documents/2980720/magic-quadrant-for-application-development-life-cycle-ma.

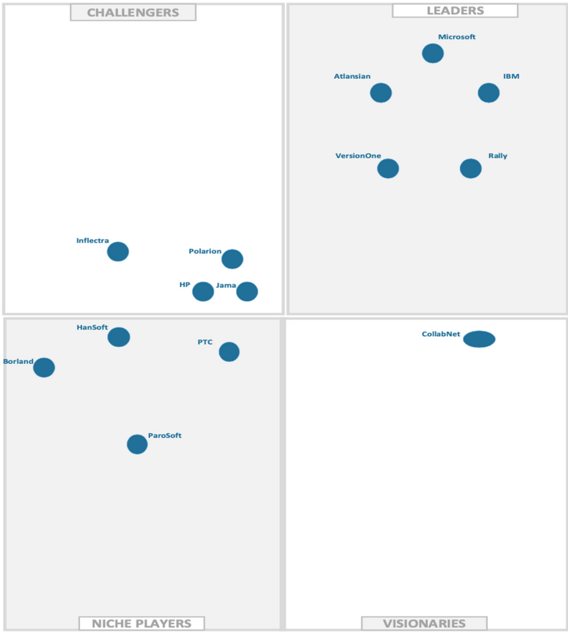


Fig. 3. Magic quadrant for application development life cycle management

- Receiving the license -based price offers from the ALM suppliers,
- Inviting the ALM supplier companies to the demonstration meetings and ensuring that demonstrations on the tool are made by the firms by holding meetings in this scope and evaluating the needs and solution methods in the form of questions-answers

During the evaluation and selection of the alternative ALM tools; selection criteria have been determined for these tools (See Table 1):

The weight points of each selection criterion and the score for tool's fulfillment of our selection criteria were given. The score of each relevant tool according to the each criteria has been calculated using the Tool Weight Rating * Score for the Alternative Tool's Fulfillment of the Selection Criteria formula. According to the selection criteria identified as a result of the evaluations made using this scoring technique, it has been decided to choose the IBM JAZZ ALM tool, which had the highest total score. After having submitted the evaluation results and approval of the management, procurement processes have been initiated. After this stage, the following actions have been completed by meeting with the supplier;

- Drafting of the Technical Specification,
- Preparation of the contract,
- Establishment of the installation and training plans for the commissioning of the tool,

Table 1. Alternative ALM tools selection criteria list

	Criteria description
1	Defining different requirement types
2	Defining custom fields
3	Associating different requirement types with custom fields
4	Establishing trace between the usage status steps and other requirement tests
5	Establishing trace between the test steps and the requirements
6	Warning for the associated requirement or tests when a requirement is updated
7	Recording of the history of the requirement and test updates
8	Creating baseline and changing baseline content
9	Display of relevant requirement and tests for impact analysis
10	Reporting skills
11	Flexibility and customization opportunities
12	Ratio of compliance to our processes
13	Potential of including design phase and relevant process(es)
14	Potential of including coding phase and relevant process(es)

- Determining and contracting the warranty process and licensing upgrade requirements,
- Mutual signing of the contract.

Adoption of Pilot Projects: Upon signing the contract with the IBM JAZZ Supplier, installation and adaptation works started within 3 days. The duration of the commissioning of the system within the scope of these activities including the installations is 4 months and we have worked with the supplier firm every week. As per our institutional processes, works of each working day have been recorded in the signed forms against *service delivery minutes*. Consequently, relevant teams have received training and their certificates at a ceremony upon completing the training successfully. In this way the use of the system is encouraged by means of motivation and advances have been made to achieve our goal of automation of the development benches.

Information on adaptations to the tool is summarized in the Table 2 below;

Table 2. Statistics of the IBM JAZZ configuration

# of Project templates	4
# of document templates	4
# of artifact types	32
# of artifact templates	26
# of document types	2
# of tracebility types	39

Going Live: Tool has been made ready for the use of the entire company and experiences in the pilot have been shared with the other projects. Within the scope of the activities carried out to go live, a presentation has been made to the management.

During a meeting held, it has been decided to plan the transfer of a number of projects to IBM JAZZ. Consequently, Table 3 shows the number of all affected entities.

Table 3. Usage statistics of the ALM platform

General usage	Numbers
# Projects and modules	111
# Users	104
Total # of work items/Artifacts	39.124
Revisions of work items	2.254

3.1.4 Evaluating

The selected product provides basic services that allow group applications to work together as a single logical server. The product ensures that two important parties of Quality Management (QM) and Requirements Management (RM) are managed and followed. After associating project areas in different applications, artifacts in those project areas can be linked e.g., the requirements in an RM project area can be linked to work items in the project area of Change and Configuration Management (CCM), requirements and work items can be linked to test plans and test scenarios in a QM project area.

In each application, the teams work within the space allocated for the project area. A project area defines the items to be delivered, the team structure, operations, and timing of the project. Project managers are responsible for creating and managing project areas. Reports can provide effective decision-making by allowing information to be seen at a glance. The report component contains a data warehouse to store read-only history and aggregated data.

Application Lifecycle Management and systematic management of software lifecycle processes according to international standards and methodologies are very important for project/business monitoring [19]. Process outputs guide the people who are interested in the issue, when there is missing information or an open issue and present a productive ecosystem [20].

3.1.5 Specifying the Learning

Throughout the action research, lessons learned and significant comprehension documentation have been created. Our action research study has enabled us to integrate the ALM requirements into the corporate system by taking an integrated review and ultimately addressing them. Software development practices in the company have been better understood and comprehended, and the team has acquired a significant amount of information about the best tool kits and ALM concepts for each process area.

Above all, it was seen that there were requirements and complex requirements for integration within the scope of determining the identified customer needs and the tools, which help the lifecycle activities were not at the level to meet all our needs. We found out that some of the problems we have faced can be very difficult to deal with. As a result of our research, we have identified some tools that can simultaneously support integration and satisfy the necessary lifecycle activity needs.

To mitigate risks associated with ALM adoption, we decided to implement an iterative and incremental approach. Based on our previous experience [21], we evaluated the implementation of pilot projects in various domains. This has not only helped to provide feedback to customize the platform but also to directly create experienced practitioners who are well-endowed with organizational and social capital to adopt new organizational forms and lead the agile transformation process [22]. In general, it was observed that socio-technical factors had a continuous effect on software development's social dynamics [23,24]. Therefore, the practitioners who utilized different tools resist the change from the very early stages of this transition. However, it has previously been observed that this resistance has resulted from the belief systems of software development professionals (i.e., the underlying beliefs of the key stakeholders [25]) who tend to reject some of the deployments. To adopt the ALM, it is very important to convince the key people in software teams. Besides the many benefits this adoption has to offer, there are still some drawbacks that need to be considered. From a technical perspective, the cross-platform compatibility is one of them. In order to cope with new concerns, it is necessary to integrate learning and development activities.

4 Conclusions

In this paper, we reported an action research study for ALM adoption in an industrial context. To document practitioners' experiences, action research acts as an important guideline for companies with similar needs. The findings of this research offered a preliminary analysis and provided insight into the ALM implementation roadmap. This study is an important milestone and a significant achievement for the company that has provided valuable long-term benefits such as traceability, integration, and process automation. In fact, development time and cost have been reduced, quality management processes have been improved, and all the other operations have been managed in harmony within the company. Although engineers, managers, and some other stakeholders have initially resisted the adoption of the ALM in the early stages, after implementation, they have realized the benefits of this effort.

In software engineering, adoption of action research for ALM is lately drawing more attention and becoming one of the most reliable research methodologies as opposed to the past. This study can be viewed as a milestone for a successful application of action research. One of the more significant findings to emerge from this study is that this approach guided the firm for some major changes in a limited amount of time. Further research could also be conducted to determine the effectiveness of action research to address different software engineering problems.

Acknowledgments. This research work is supported, in part, by Akgun Software Inc. During the research, we have received help and generous support from many people, to all of whom we would like to express our gratitude.

References

1. Clarke, P., O'Connor, R.V., Yilmaz, M.: In search of the origins and enduring impact of agile software development. In: Proceedings of the 2018 International Conference on Software and System Process, pp. 142–146 (2018)
2. Clarke, P., et al.: An investigation of software development process terminology. In: Clarke, P.M., O'Connor, R.V., Rout, T., Dorling, A. (eds.) SPICE 2016. CCIS, vol. 609, pp. 351–361. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-38980-6_25
3. Üsfekes, Ç., Yilmaz, M., Tuzun, E., Clarke, P.M., O'Connor, R.V.: Examining reward mechanisms for effective usage of application lifecycle management tools. In: Stolfa, J., Stolfa, S., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2017. CCIS, vol. 748, pp. 259–268. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64218-5_21
4. Lacheiner, H., Ramler, R.: Application lifecycle management as infrastructure for software process improvement and evolution: experience and insights from industry. In: 2011 37th EUROMICRO Conference on Software Engineering and Advanced Applications, pp. 286–293. IEEE (2011)
5. Sneed, H.M.: Economics of software re-engineering. J. Softw. Maintenance Res. Pract. **3**, 163–182 (1991)
6. Klespitz, J., Bíró, M., Kovács, L.: Aspects of improvement of software development lifecycle management. In: 2015 16th IEEE International Symposium on Computational Intelligence and Informatics (CINTI), pp. 323–327. IEEE (2015)
7. McHugh, M., McCaery, F., Casey, V., Pikkarainen, M.: Integrating agile practices with a medical device SDLC. In: International Conference on Software Process Improvement and Capability Determination (SPICE 2014) (2014)
8. Rossberg, J.: Beginning Application Lifecycle Management. Apress, Berkeley (2014)
9. Herden, S., Zwanziger, A., Robinson, P.: Declarative application deployment and change management. In: 2010 International Conference on Network and Service Management, pp. 126–133. IEEE (2010)
10. Chappell, D.: What is application lifecycle management? (2008). http://www.davidchappell.com/writing/white_papers/What_is_ALM_v2.0-Chappell.pdf. Accessed 3 Sept 2019
11. Kravchik, M.: Application lifecycle management environments: past, present and future. Ph.D thesis, Open University of Israel (2009)
12. Schwaber, C.: The Changing Face of Application Life-cycle Management, vol. 18. Forrester Research (2006)

13. Olausson, M., Rossberg, J., Ehn, J., Sköld, M.: Introduction to application lifecycle management. In: Pro Team Foundation Service, pp. 1–7. Apress, Berkeley (2013). https://doi.org/10.1007/978-1-4302-5996-1_1
14. Rossberg, J.: Introduction to application life cycle management. Agile Project Management with Azure DevOps, pp. 1–36. Apress, Berkeley (2019). https://doi.org/10.1007/978-1-4842-4483-8_1
15. Tüzün, E., Tekinerdogan, B., Macit, Y., İnce, K.: Adopting integrated application lifecycle management within a large-scale software company: an action research approach. J. Syst. Softw. **149**, 63–82 (2019)
16. Baskerville, R.L.: Investigating information systems with action research. Commun. Assoc. Inf. Syst. **2**, 19 (1999)
17. Manifesto: Software Process Improvement (SPI). https://2019.eurospi.net/images/eurospi/spi_manifesto.pdf. Accessed 3 Sept 2019
18. Wilson, W., Duggan, J., Murphy, T., Sobejana, M., Herschmann, J.: Magic Quadrant for Application Development Life Cycle Management (2015)
19. Korsaa, M., et al.: The people aspects in modern process improvement management approaches. J. Softw. Evol. Process **25**, 381–391 (2013)
20. Messnarz, R., et al.: Social responsibility aspects supporting the success of SPI. J. Softw. Evol. Process **26**, 284–294 (2014)
21. Yilmaz, M.: Observed effects of software processes change in three software firms: industrial exploratory case study. Pamukkale Univ Muh Bilim Derg **25**, 240–246 (2019)
22. Yilmaz, M., O'Connor, R.: Social capital as a determinant factor of software development productivity: an empirical study using structural equation modeling. Int. J. Hum. Capital Inf. Technol. Professionals (IJHCITP) **3**, 40–62 (2012)
23. Yilmaz, M.: A software process engineering approach to understanding software productivity and team personality characteristics: an empirical investigation. Ph.D thesis, Dublin City University (2013)
24. Yilmaz, M., O'Connor, R.V., Clarke, P.: Effective social productivity measurements during software development-an empirical study. Int. J. Softw. Eng. Knowl. Eng. **26**, 457–490 (2016)
25. O'Connor, R.V., Yilmaz, M.: Exploring the belief systems of software development professionals. Cybernet. Syst. **46**, 528–542 (2015)

Digitalisation of Industry, Infrastructure and E-Mobility

Automotive Mechatronics Training Programme - An Inclusive Education Series for Car Manufacturer and Supplier Industry

Mario Hirz(✉)

Institute of Automotive Engineering,
Graz University of Technology, Graz, Austria
mario.hirz@tugraz.at

Abstract. Increasing integration of electrics/electronics into modern products has led to a significant change in development processes and value creation of formerly mechanically oriented industries. With rising share of information technology in vehicles both product functionalities and customer benefit increase, which requires fundamental changes in the development and manufacturing processes. Due to the fact, that modern cars include a significant and still increasing share of mechatronics systems involving mechanical, electrics/electronics and software components, a smart and effective integration of these historically often separated disciplines plays an important role throughout the entire product lifecycle. To enable this integration, development processes as well as applied tools and methods are changing, which requires comprehensive skills of the involved human resources including management, engineers and a range of other concerned parties. The present publication discusses the progress of technology change in the automotive industry and the associated requirements on the applied development processes. Based on these, a comprehensive automotive mechatronics training programme is introduced, which addresses the needs of knowledge enhancement by know-how transfer between the different involved engineering disciplines.

Keywords: Automotive industry · Expert education · Mechatronics training programme · Automotive development · Cross-domain integration

1 Introduction

The steadily increasing share of E/E and software components in cars is motivated by aspects of comfort and communication, but also by continuously strengthened legislative boundary conditions for exhaust emissions and vehicle safety. In a modern midsize class car, the portion of averaged value creation of E/E reaches more than 30%, and this value will further increase during upcoming years. However, the share of E/E research and development has overtaken traditional mechanical-oriented development tasks and it will become even more important during the next years, [2]. The development processes did not follow the new requirements sufficiently so far, and applied development strategies, tools and methods as well as human resources skills lack in this context, too. Future comprehensive development and manufacturing will face these

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 341–351, 2020.
https://doi.org/10.1007/978-3-030-56441-4_25

new product characteristics by provision of enhanced development methods for mechatronics systems, which are able to support simultaneous development of mechanical, electrical and electronics systems including software. Especially software development follows different rules than traditional hardware development, because of differing behaviour in view of complexity management, occurring errors and failure, durability and lifecycle. In this way, new development approaches will provide functionalities not only to implement e.g. an electric motor or a camera system into a digital mock-up, but also features to consider functional aspects of implemented modules. For the applied development platforms, this requires interfaces between mechanical design, simulation and software development programs as well as evaluation tools for testing of complete mechatronics systems. The integration of the three domains results in the commonly known mechatronics systems, which conception, development, and production require comprehensive multi-domain knowledge. Figure 1 displays the interaction of the involved disciplines of mechatronics.

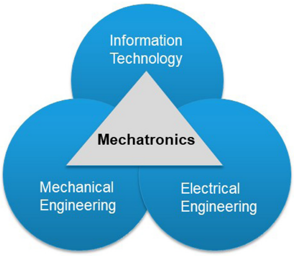


Fig. 1. Mechatronics integrating the domains mechanical- and electrical engineering as well as information technology.

According to VDI Guideline 2206, mechatronics is defined as synergetic interaction of the disciplines mechanical engineering, electrical engineering and information technology for the creation and production of industrial products as well as in process design [1]. In this way, mechatronics describes technical systems, which are composed of mechanic, electric and electronic components. In many cases, information processing technologies control primary (mechanical or electrical) functions of material and energy conversion. But in mechatronics systems, integrated functionalities and flexibilities are significantly larger than in conventional electromechanical systems.

2 Automotive Development Processes

Figure 2 shows an overview of the main development steps in a general automotive development process. The generic process is divided into five main phases: Definition Phase, Concept Phase, Pre-Development- and Series Development Phases as well as the Pre-Series and Start of Production (SOP). Throughout these development phases, a number of development disciplines come to use, e.g. requirements engineering (REQ ENG), which are performed according to the different work packages and disciplines involved. A successful implementation and conduction of these various tasks - performed in the different disciplines of mechatronics products development - requires both comprehensive process management as well as cross-domain knowledge of the involved staff.

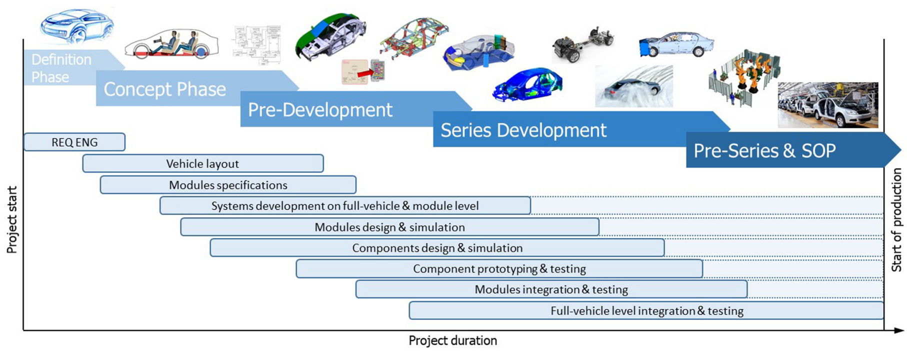


Fig. 2. Exemplary automotive development process, including different main steps of mechatronics product development.

Figure 3 shows a typical development process according to the so-called V-model, applied on automotive mechatronics systems, [1]. The V-model represents a special case of waterfall models that describe sequential steps of product development, [4]. The process begins at the top end of the left branch with product specifications that result from a list of requirements. The entire left branch focusses on product requirements definition, layout and design on system-, module- and component level including different areas of simulation and optimization. In this way, it is divided into a sequential chronology of increasing levels of detail. The system level includes systems engineering- related development, e.g. system architecture, definition of system functionalities, and system interfaces, [7]. After having defined main characteristics on system level, the module level provides a breakdown of complex systems into several modules according to the previously defined system architecture, e.g. mechanical structure, power supply, electrical sensor and actuators, software functionalities. Finally, modules are divided into their components, which are developed in the component level. Here, cross-domain implementation is performed at the bottom level of the V-model in the course of component integration.

The right branch of the V-model includes prototyping, testing and optimization at component, module and system level. After being tested, components are built together to modules, which are integrated and tested according to their specific functionalities. In the final system level at the top end of the right branch, all elements are assembled to a system prototype and tested for product confirmation according to initially defined specifications. Typically for development according to the V-model is a close interaction of design and testing. In this way, data and information exchange between product design (left branch) and integration & testing (right branch) support efficient improvements and optimization. In case of highly complex products, e.g. an automated

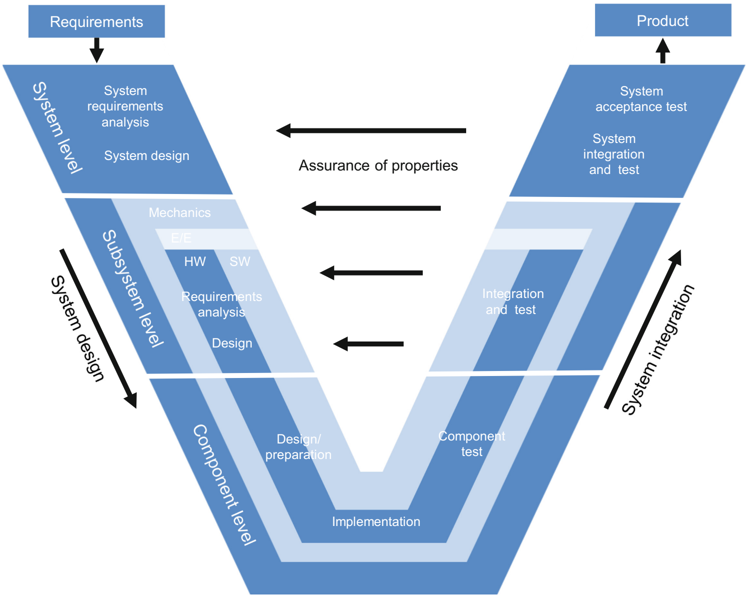


Fig. 3. Exemplary automotive development process according to the V-model, derived from [1].

transmission, the development process is run through several times, especially on module and component level. Both duration and complexity of these development cycles differ significantly in the three domains, which leads to varying levels of maturities in mechanics, electrics, hardware and software development.

In general, the V-model represents itself as a well-established approach for the development of mechatronics systems, but with rising complexity and increasing share of E/E, it requires deep understanding of the different disciplines to enable efficient cross-domain integration. To face these challenges, the involved parties have to provide cross-domain knowledge to be able to understand and support the corresponding development processes. The technological transition that takes place in the car industry today, e.g. the electrification of propulsion systems and the implementation of automated driving functions, requires new skills. Key of success lies in the introduction of flexible, interdisciplinary processes, which are able to consider different domain-specific methodological, functional and development cycle-time related characteristics. One important challenge in this context includes the provision of the required knowledge and skill for the involved staff.

Figure 4 visualises the challenge of cross-domain integration by a spiral model of mechatronics systems development processes following the V-model. Both duration and complexity of the domain-related development cycles differ significantly, which leads to varying levels of maturity throughout mechanic, hardware and software development. The different domain-specific development procedures in combination with changeable cycle frequency lead to a considerable increase of complexity in the development processes. In the exemplary spiral model, the subsequently performed process phases are shown in more detail for the system level. The processes of the three development domains are not shown in detail but follow the same principle, considering the specific procedures of mechanics-, electrics- and software development. In practical applications, the V-model is run through repeatedly in the different levels and domains. The phase of systems engineering plays an essential role in the development of complex mechatronics products, because it includes the general functional and structural layout. Out of knowledge gathered in the systems engineering phase, the subsequently performed domain-specific processes are supplied with required data and information for successful product development. In this way, the systems engineers involved have to understand cross-domain conceptual design of mechanical, electrics and software components as well as their integration into the product to be developed.

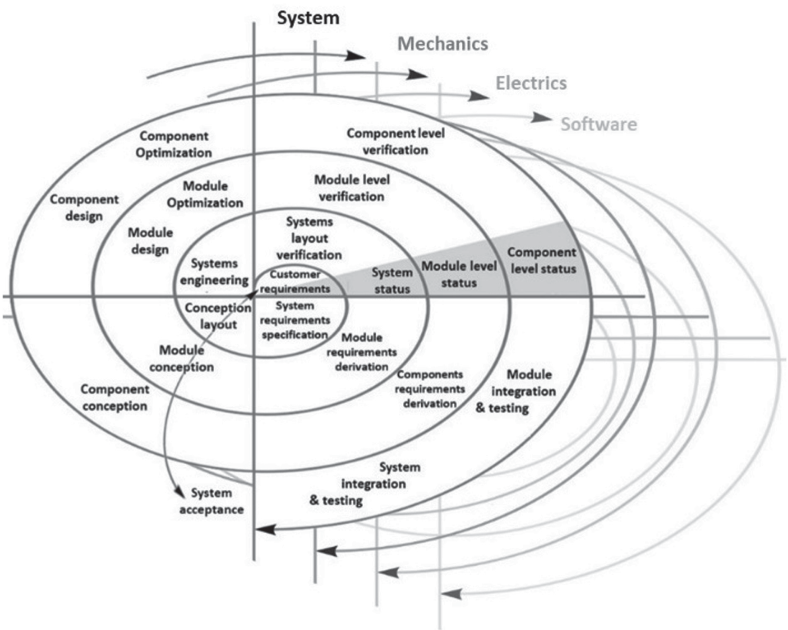


Fig. 4. Spiral model of mechatronics product development processes following the V-model [3].

Another important area of knowledge required for successful mechatronics product development includes the different standards and guidelines, which are defined within the three involved domains mechanics, electrics and informatics technologies. These standards and guidelines cover broad areas of development methods, technical aspects and boundary conditions within the domain they are made for. Exemplary, there are lots of standards covering the development of mechanical components and systems, e.g. crash- & safety standards for cars, [9]. Electric component development is also topic of several standards, e.g. regarding electromagnetic compatibility in automotive applications, [10]. Additionally, the informatics domain that includes computer hard- and software, is increasingly topic of guidelines and standardization, e.g. ISO 26262 Functional Safety [12], ISO/IEC 15504 ASPICE [14, 15], CMMI [15, 17].

These guidelines and standards fit into state of the art development processes, e.g. the V-model, and provide structures, strategies and methods to support the development processes in automotive applications. Besides the mentioned development process standards and guidelines that include comprehensive product development, monitoring and optimization measures in view of E/E development, there are various additional standards and guidelines addressing quality management in different phases of product lifecycle, e.g. ISO 9000 [21], ISO 9001 [22], ISO 9004 [23], IATF 16949 [24], ISO/IEC 90003 [25], ISO/IEC 25000ff [26], ISO/IEC 15504 [27], ISO 26262 [28], VDI 2221 [20], VDI 2206 [19]. Finally, specific quality management methods and tools are integrated into development processes, e.g. Quality Function Deployment (QFD) [5], Failure Mode and Effects Analysis (FMEA), Fault Tree Analysis (FTA) [18], Statistical Process Control (SPC) [16], Poka-Yoke [13] and Total Quality Management (TQM) [6].

One important aspect of successful mechatronics product development represents the provision of the required knowledge and skills. Know how is provided in the three domains in multifarious way, e.g. by specific university studies and job trainings. In this way, different job roles exist to supply the corresponding domain-specific development, e.g. mechanical engineers, electrical engineers, software development engineers. In these domains, clear roles and know-how of specified jobs have been established during the past decades. Not sufficiently available are job roles providing cross-domain knowledge and skills, as it is important in mechatronics products development. Besides being familiar with the development process landscape, e.g. the V-model, and the corresponding development standards and guidelines, these job roles demand understanding of mechatronics systems engineering from top to detailed view to serve as combining and networking institution between the different topic-related expertises.

3 Education Concept

The discussed rising degree of electrification and digitalization of automotive systems and the effective integration of the domains mechanics, electrics and information technology play a significant role in automotive development processes. Besides engineering expertise in each of the domains, vehicle manufacturer and supplier increasingly need human resources for the management, development and administration of mechatronics systems throughout the entire value-creation chain. This includes requirements- and systems engineering, conception, design, simulation, manufacturing engineering as well as production, logistics, maintenance and quality management of mechatronics products.

With the target to meet the increasing requirements on employee's knowledge in the automotive industry, a comprehensive training programme has been developed. The training programme is offered by the Life-Long-Learning Institute of Graz University of Technology [11] since 2018 and is in addition part of the ECQA DRIVES programme [8] since 2020. The programme has been successfully conducted at different international automotive supplier companies, e.g. a supplier of automotive drivetrain components and a supplier developing and producing automotive tank systems, with training locations throughout Europe, Asia and North America. In total, more than 600 people staff in industry have been educated and have enhanced their skills and knowledge during the past years (status 2020). Besides the mentioned industrial training programme, university and non-academic exchange has been established internationally in and beyond Europe. Providing comprehensive knowledge about mechatronics systems in automotive applications, including aspects of systems engineering, modules- and component development, industry benefits from the offered training programme in different aspects. On the one hand, employee's skills are enhanced to the domains of mechatronics systems to support a deep understanding of the different development disciplines and their interaction. On the other hand,companies are supported in their transition from single-domain expertise to enterprisesthat can handle multi-domain product development and manufacturing. It has shown in the past, that automotive supplier with long tradition in mechanically-oriented product development and production have to expand their product ranges to mechatronics products to remain competitive in the market. This requires an enhancement of skills that is offered by the mechatronics training programme.

The training programme is clustered according to four job roles, which correspond to the actual needs of the automotive industry. Within these job roles, maximum eight modules are offered, including one virtual introduction module and seven intensive modules taught by specified trainers comprising each one week (40 lesson-hours), see Fig. 5. The skill hierarchy for the job roles of the "Automotive Mechatronics Training Programme" has been designed based on a comprehensive demand analysis in the automotive industry. In this context, the corresponding job roles are embedded into the complex development processes in the automotive industry. These roles fit into both system- and vehicle development as well as production engineering. The training is designed as modular course; in this way the trainees can attend units and elements separately and independently, although it is recommended to attend the entire unit structure according to the targeted job role. In general, there are four levels offered:

1. Introduction to Automotive Mechatronics (online training course)
2. Automotive Mechatronics Manager Basic Level
3. Automotive Mechatronics Expert
4. Automotive Mechatronics Developer

The training of the first level "**Automotive Mechatronics Manager** – Introduction to Automotive Mechatronics" includes one unit comprising three main elements. The duration of training is three hours (online training course). After having accomplished the training, the trainees can take an exam and become a certified Automotive Mechatronics Manager on Introduction to Automotive Mechatronics Level. In this course it is intended to provide basic information to decision makers and interested persons and to introduce into the topics and challenges of mechatronics systems development. Participants who have accomplished the Automotive Mechatronics Manager on Introduction to Automotive Mechatronics Level have a knowledge about the challenges and demands in mechatronics systems development and understand the general requirements on the development and production of mechatronics systems and products.

Content Unit of the Training Skill Card	Introduction to Automotive Mechatronics	Automotive Mechatronics Manager Basic Level	Automotive Mechatronics Expert	Automotive Mechatronics Developer
Unit 1 - Introduction to Automotive Mechatronics (virtual)				
Introduction to Mechatronics Systems	comprehension			
Mechatronics Systems Development & Production	comprehension			
Required Skills and Boundary Conditions	comprehension			
Unit 2 - Automotive Mechatronics Systems Development				
Boundary Conditions in Automotive Industry		comprehension	comprehension	comprehension
Characteristics, Structures &Functions of Mechatr. Systems		comprehension	comprehension	comprehension
Mechatronics Systems Development Processes & -Standards		application	application	application
Mechatronics System Engineering Basics		application	application	application
Unit 3 - Introduction of Electrical Engineering & IT				
Basic Elements of Electric Circuits			comprehension	comprehension
DC-Circuits Layout and Calculation, AC-Current Basics			analysis	analysis
Basics of Magnetic and Electric Fields			analysis	analysis
IT Development Process and Embedded System Integration			comprehension	comprehension
Introduction of IT Hardware and Software Engineering			comprehension	comprehension
Automotive Communication Systems & Data Interfaces			application	application
Unit 4 - Sensor, Actuator and Electric Motor Technologies				
Sensing & Measurement Principles			analysis	analysis
Sensor System Data Processing			application	application
Sensor & Measurement Technologies			application	application
Types and Functions of Electric Motors and Actuators			analysis	analysis
Electric Motor / Actuator Design & Control Characteristics			application	application
Unit 5 - QM & Verification of Mechatronics Systems				
Mechatronics Development Processes, Standards & Guidelines			comprehension	comprehension
Basics of Functional Safety, System Verification & Optimization			comprehension	comprehension
Quality Management			application	application
ISO 26262, ASPICE, CMMI			application	application
Risk Assessment, System Test and Integration			application	application
Unit 6 - Introduction of Matlab-Simulink				
Introduction and Overview of Simulation Basics				comprehension
Program Introduction and User-Handling				comprehension
Constants, Variables, Vectors, Matrices				analysis
Modelling of Simple Systems, Linear Equation Systems				application
Data Processing and Programming Structures				application
Model-Based Simulation of Mechatronics Systems				application
Unit 7 - Introduction of Control Engineering				
Notion of Control				comprehension
Feedback Control vs. Feedforward Control				analysis
Basic Modelling and Simulation Concepts				analysis
Transfer Functions and Stability				application
PID controllers				application
Simulation of Feedback Loops / Applications				application
Unit 8 - Advanced Systems & Components Design				
Mechatronics System Layout and Design				application
Mechatronics System Simulation				application
Sensor, Actuators & ECU Hardware Simulation & Optimization				application
Investigations of Control Strategies				application
Application and Discussion based on a Project Work				application

Fig. 5. The skills set for ECQA certified automotive mechatronics roles (overview) [8].

Training of the second level "**Automotive Mechatronics Manager Basic Level**" includes one unit comprising four main elements. The duration of training is one week. After having accomplished the training, the trainees can take an exam and become a certified Automotive Mechatronics Manager Basic Level. In this job role, understanding of the legal background, complexity and behaviour of mechatronics functions and products is given. In addition, risks and impacts on design and costs, as well as required organisation roles and responsibilities in the development and production of mechatronics systems are understood. Automotive Mechatronics Manager Basic Level have a knowledge about the complex interactions of the development domains mechanics, electrics and IT and can plan and manage the proper methods to be applied in successful development projects. They know the required development standards and guidelines and can manage the involvement of the corresponding expertise and resources.

Training of the third level "**Automotive Mechatronics Expert**" includes four units comprising different main elements. The duration of training is four weeks. After having accomplished the training, the trainees can take an exam and become a certified Automotive Mechatronics Expert. Automotive Mechatronics Experts understand the legal background, complexity and behaviour of mechatronics functions and products. In addition, they understand risks and impacts on design and costs and required organisation roles and responsibilities in the development and production of mechatronics systems. Automotive Mechatronics Experts have a knowledge about the complex interactions of the development domains mechanics, electrics and IT and can plan and manage the proper methods to be applied in successful development projects. They know the required development standards and guidelines and can manage the involvement of required expertise and resources. In addition to the Job Role of Automotive Mechatronics Manager Basic Level, Automotive Mechatronics Experts have deep technical insights of mechatronics products. This includes knowledge in electrical- and computer engineering, detailed knowledge in sensor systems and electric actuators, as well as fundamental understanding of Quality Management and Verification of Mechatronics Systems.

The training of the forth level "**Automotive Mechatronics Developer**" includes seven units comprising different main elements. The duration of training is seven weeks. After having accomplished the training, the trainees can take an exam and become a certified Automotive Mechatronics Developer. Automotive Mechatronics Developer – as an enhanced level of the job role "Automotive Mechatronics Expert" - understand the legal background, complexity and behaviour of mechatronics functions and products. In addition, they understand risks and impacts on design and costs and required organisation roles and responsibilities in the development and production of mechatronics systems. Automotive Mechatronics Developer have a knowledge about the complex interactions of the development domains mechanics, electrics and IT and can plan and manage the proper methods to be applied in successful development projects. They know the required development standards and guidelines and can manage the involvement of required expertise and resources. In addition to the Job Role of Automotive Mechatronics Manager Basic Level and Automotive Mechatronics Expert, Automotive Mechatronics Developer have deep technical insights of mechatronics products. This includes knowledge in electrical- and computer engineering, detailed knowledge in sensor systems and electric actuators, as well as fundamental understanding of Quality Management and Verification of Mechatronics Systems. Automotive Mechatronics Developer have skills for the development of mechatronics systems, including layout, design, simulation and testing. They have the knowledge to develop control algorithms within comprehensive simulation tools and they understand the complex interaction between the components of embedded systems.

4 Conclusion

To face the challenges in multi-domain mechatronics product development, development processes and tools have to provide much more than the traditional functionalities and data structures. In automotive industry, the implementation of development standards for mechanical components and electrics/electronics systems took place during the past decades, but a successful integration of the three domains mechanics, electrics and electronics/software requires more than new development processes. Future-oriented process- and product models require highly educated staff in all involved hierarchies and development disciplines to enable effective cross-domain development of mechatronics systems.

In this context, a comprehensive educational programme has been developed and successfully offered to automotive manufacturer and supplier industry. Facing the upcoming challenges caused by large technology- and business model-related changes in this industry sector, knowledge transfer by effective training plays a crucial role. The introduced automotive mechatronics training programme represents an inclusive education series to support car manufacturer and supplier industry to strengthen their competitiveness and their know-how for development and manufacturing of new solutions and products.

References

1. Development Method for Mechatronic Systems, VDI Guideline 2206, Association of German Engineers (2003)
2. Hirz, M., Dietrich, W., Gfrerrer, A., Lang, J.: Integrated Computer-Aided Design in Automotive Development. Development Processes, Geometric Fundamentals, Methods of CAD, Knowledge-Based Engineering Data Management. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-11940-8
3. Hirz, M., Hofstetter, M., Lechleitner, D.: Electric propulsion systems design supported by multi-objective optimization strategies. In: Proceedings of the 16th European Automotive Congress (EAEC) 2019, Minsk, Belarus (2019). https://doi.org/10.21122/2227-1031-2019-18-6-461-470
4. Sell, R., Tamre, M.: Integration of V-model and SysML for advanced mechatronics system design. In: International Workshop on Research & Education in Mechatronics, At ANNECY, France (2005)
5. Ernst, M.: KPI-related analysis methods to optimize mechatronic product development processes; doctoral thesis, Institute of Automotive Engineering, Graz University of Technology (2016)
6. O'Regan, G.: Introduction to Software Quality. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-06106-1. ISBN 978-3-319-06105-4
7. Dietrich, W., Hirz, M., Rossbacher, P.: Integration von geometrischen und funktionalen Aspekten in die parametrisch assoziative Modellgestaltung in der konzeptionellen Automobilentwicklung. In: Grazer Symposium Virtuelles Fahrzeug, Austria (2010)
8. Homepage of ECQA – European Certification & Qualification Association. http://www.ecqa.org/. Accessed 07 Apr 2020

9. SAE Standard J224: Collision Deformation Classification. www.sae.org. Accessed 10 Dec 2017
10. SAE Standard J1113: Electromagnetic Compatibility Measurement Procedure for Vehicle Components. www.sae.org. Accessed 10 Dec 2017
11. Life-Long-Learning Institute at Graz University of Technology, homepage. https://www.tugraz.at/en/studying-and-teaching/degree-and-certificate-programmes/continuing-education/life-long-learning/. Accessed 05 Apr 2020
12. ISO Standard 26262 Road vehicles - Functional safety; International Organization for Standardization. www.iso.org. Accessed 10 Dec 2017
13. Shingo, S.: Zero Quality Control: Source Inspection and the Poka. B&T Productivity Press, Portland (1986). ISBN 9780915299072
14. Anderson, P.: Mehr Softwaresicherheit, Statische Analyse-tools und die ISO 26262. J. ATZ Elektronik, January 2017
15. Automotive Spice (ASPICE) process assessment and process reference model; SIG & VDA QMC. http://www.automotivespice.com. Accessed 10 Dec 2017
16. Shewhart, W.A.: Economic Control of Quality of Manufactured Product. ASQ Quality Press, Milwaukee (1980). ISBN 0-87389-076-0
17. CMMI Product Team: CMMI® for Development, Version 1.3; Carnegie Mellon University, Software Engineering Process Management Program, Pittsburgh (2010)
18. Ernst, M., Dallinger, P., Fabian, J., Hirz, M., Schnellbach, A.: Innovative Analysemethode in der Entwicklung mechatronischer Systeme am Beispiel elektrifizierter Antriebsstränge. Elektrotech. Inftech. **132**(3), 134–141 (2015). https://doi.org/10.1007/s00502-015-0292-7
19. The Association of German Engineers: VDI 2221, Methodik zum Entwickeln und Konstruieren technischer Systeme und Produkte; Germany; VDI (1993)
20. International Organization for Standardization: ISO 26262-1:2011 Road vehicles - Functional safety; Switzerland; ISO (2011)
21. Raubold, U.: Lebenszyklusmanagement in der Automobilindustrie - Ein Optimierungsansatz auf Basis der auf den Lebenszyklus wirkenden Einflussfaktoren; doctoral thesis, Technische Universität Cottbus, 2010, ISBN 978-3-8349-2862-7, Germany (2010)
22. International Organization for Standardization: ISO 9000 - Quality management systems - Fundamentals and vocabulary; Switzerland, ISO (2015)
23. International Organization for Standardization: ISO 9001: 2015 Quality management systems - Requirements; Switzerland: ISO (2015)
24. IATF 16949:2016 Anforderungen an Qualitätsmanagementsysteme für die Serien- und Ersatzteilproduktion in der Automobilindustrie, 1. Ausgabe, October 2016
25. International Organization for Standardization: ISO/TS 16949:2009 Quality management systems - Particular requirements for the application of ISO 9001:2008 for automotive production and relevant service part organizations; Switzerland, ISO (2009)
26. International Organization for Standardization: ISO/IEC 90003:2014 Software engineering - Guidelines for the application of ISO 9001:2008 to computer software; Switzerland, ISO (2014)
27. International Organization for Standardization: SO/IEC 25000:2014 Systems and software engineering - Systems and software Quality Requirements and Evaluation (SQuaRE) - Guide to SQuaRE; Switzerland, ISO (2014)
28. International Organization for Standardization: ISO/IEC 15504-1:2004 Information technology - Process assessment; Switzerland; ISO (2004)

Automotive Engineering Skills and Job Roles of the Future?

Jakub Stolfa[1(✉)], Svatopluk Stolfa[1], Richard Messnarz[2], Omar Veledar[3], Damjan Ekert[2], Georg Macher[4], and Utimia Madaleno[5]

[1] VSB TUO Ostrava, Ostrava, Czech Republic
jakub.stolfa@vsb.cz
[2] ISCN GesmbH, Graz, Austria
rmess@iscn.com
[3] AVL List GmbH Graz, Graz, Austria
[4] University of Technology Graz, Graz, Austria
[5] Eupportunity, Brussels, Belgium
um@eupportunity.eu

Abstract. Technological progress and new lifestyle demand is changing the way we live and affects the future development and usage of technologies in different areas including the automotive sector. According to this demand, the future car is aiming to be smart, clean and autonomous. To be able to achieve that in the relatively short time given by the social demand, we need to assess, evaluate and establish a combination of new, emerging and existing skills pools to define future job roles and skills required for the design, production and maintenance of cars. Mapping these skills onto the phases of engineering, production, and maintenance in the automotive life cycle outlines an interesting future picture of a vehicle that will integrate smart technologies, smart production, smart maintenance and continuous machine learning. In this paper the job roles supporting this future are proposed and discussed.

Keywords: Skills · Job roles · Automotive engineering

1 Introduction

Digital and other key technologies are reshaping the foundations of industry and the wider global economy. This offers tremendous opportunities for innovation, growth and jobs, but it also requires skilled and adaptable individuals who are able to drive and support change. In order to deliver sector-specific skills solutions, the New Skills Agenda for Europe has launched the action Blueprint for Sectorial Cooperation on Skills [47, 63]. DRIVES project [60, 68] address the automotive sector and was launched in the first wave of five sectors specific Blueprint projects. The aim is to develop concrete actions to satisfy short- and medium-term skills needs to support the overall sector - DRIVES project foresees:

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 352–369, 2020.
https://doi.org/10.1007/978-3-030-56441-4_26

1. Implementation of a common European automotive skills reference framework where industry, academia and VET-providers can access and get to the pool of skills and job roles needed to help them to grow.
2. Enabling mutual recognition of automotive skills across Europe by DRIVES digital badge in order to enhance and support the automotive workers and their applicability and comparability of skills across Europe (respectful to the current certification an formal and non-formal training networks).
3. Assessing and inclusion of existing, proven and newly developed trainings for the future job roles and/or needed skills. A first push of emerging job roles in Europe by ca. 1200 attendees in trail courses according to the future DRIVES reference framework scheme
4. Deployment of the Apprenticeship Marketplace by enhancing its effectiveness for automotive job seekers. Linking to the DRIVES framework of automotive skills and job roles to facilitate dissemination of common job requirements, which will be available for job seekers, training providers (namely universities), VET providers and other stakeholders.

When performing the skills analysis and creating the first pool of skills in Work Package 3 - Skills Framework of the DRIVES project, and in discussions with the working group SOQRATES, an interesting perspective of future automotive workforce has emerged and is discussed in the next sections.

2 Architecture for Future Automotive Engineering

Currently, the car is defined as components and functions [16, 30, 33, 41]. But in future, components and functions are inherently linked with competences of teams and smart technologies supporting these capabilities in cars. Human resources and skills will play a major role in launching new concepts and technologies into car design, production and maintenance. By empowering the skills, the new technologies get integrated into the cars and this would lead to a competence-based intelligence architecture of future cars.

Figures 1, 2, 3, 4, 5 and 6 show such an intelligence architecture of car design, production and maintenance based on the first pool of job roles on which DRIVES project is working on. It describes grouping of job roles into the three views – Engineering, Production and Series Maintenance.

This competence based architectural strategy would mean to first create a knowledge-based strategy of a future car and then co-develop the car based on shared competences [33, 38, 40, 41, 43, 50, 51, 53, 61–65].

Where can such a strategy be seen already on the market?

Papers published by leading companies developing autonomous driving concepts for European OEMs, outline that in future development and engineering partnerships are necessary and will replace traditional supply chains. E.g. Google and OEM partnership [69–71].

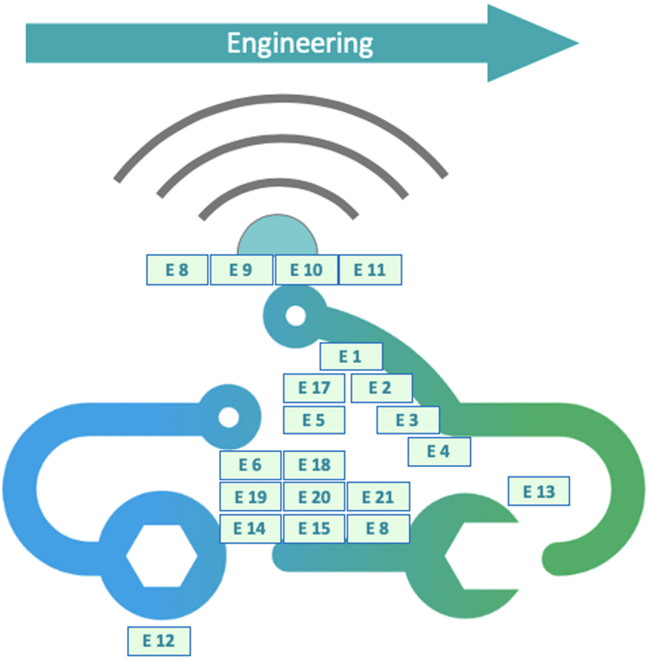


Fig. 1. Intelligence architecture in engineering

Engineering and R&D

E1: ADAS/ADF testing and validation engineer

E2: Artificial Intelligence (AI) Expert

E3: Computer Vision Expert (3 level in Dissemination material)

E4: Machine Learning Expert (3 level in Dissemination material)

E5: Sensor Fusion Expert (3 level in Dissemination material)

E6: Automotive Engineering CAD, CAE, CAM

E7: Automotive SPICE® Basics

E8: Connected Vehicles Expert

E9: Cybersecurity Engineer

E10: Cybersecurity Manager

E11: Cybersecurity Tester

E12: Rubber Technologist

E13: Advanced e-Powertrain Engineer

E14: Functional Safety Manager Strategy Level

E15: Functional Safety Manager Project Level (Advanced)

E16: Functional Safety Engineer (Advanced)

E17: Highly Automated Drive Engineer

E18: Automotive Mechatronics Manager - Introduction to Automotive Mechatronics

E19: Automotive Mechatronics Manager Basic Level

E20: Automotive Mechatronics Expert

E21: Automotive Mechatronics Developer

Fig. 2. Intelligence areas (DRIVES job roles) in engineering

Fig. 3. Intelligence architecture in production and innovation

Production

P1: Advances Manufacturing Press line Set-up

P2: Automotive Engineer in Quality and Metrology

P3: Lean Six Sigma Yellow Belt

P4: Lean Six Sigma Green Belt

P5: Lean Six Sigma Black Belt

P6: Robotic Engineer

P7: Robotic Technician

P8: Technical Cleanliness Expert

General

G1: Working in Automotive (Automotive Engineer)

G2: Automotive Quality Engineer (AQUA)

G3: Innovation Agent for Business Model Innovation

Fig. 4. Intelligence areas (DRIVES job roles) in production and innovation

The same approach can be observed in industry developments in Japan and the US. Featuring some, Nikola [67], is a new car maker developing hydrogen fuel cell powered trucks that rather organises a collaboration of suppliers and competence network strategy than a top down structured OEM relationship. The new Sony Vision S Concept [66] (2020) Prototype is a new electric car concept brought forward by SONY in

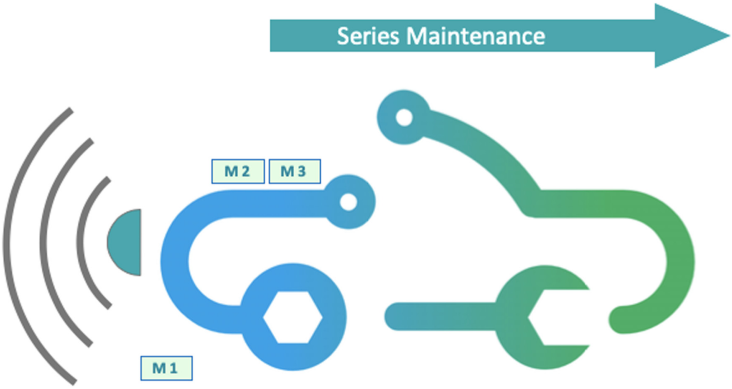


Fig. 5. Intelligence architecture in maintenance

Maintenance

M1: Automotive Engineer in Tool and Die Production and Maintenance

M2: Sustainability manager

M3: Predictive Maintenance Expert

Fig. 6. Intelligence areas (DRIVES job roles) in maintenance

cooperation with leading suppliers, combining the skills in sensor fusion, AI, and driving dynamics knowledge.

The authors of the paper are part of a group developing a pool of job roles and competencies that can be used in such an architecture of knowledge approach.

3 Short Outline of Each Competence Area

In this chapter there is a short outline how each of the knowledge areas provided in 2020–2021 contribute to the intelligence architecture outlined in Figs. 1, 2, 3, 4, 5 and 6.

3.1 Engineering and R&D Specific

E1: ADAS/ADF testing and validation engineer [24, 25, 72], [76]

As the Autonomous Driving (AD) is rapidly becoming more socially acceptable, the market pull is creating a demand for sound technical solutions. These are highly reliant on a range of intelligent and adaptable control strategies, which must be capable of taking over vehicle control from human drivers. In this context, simulation is widely required to specifically replace development and testing activities, which are often expensive, time-consuming, critical and/or seldom represent corner-cases.

E2: Artificial Intelligence (AI) Expert

Industry and society are experiencing the transformational impact of the autonomous systems revolution, empowered by automation capabilities offered by Artificial Intelligence (AI) in manifold manner. The performance of AI systems is reaching or possibly exceeding the human level on an increasing number of complex tasks. However, these highly successful AI models are usually experienced as black boxes. No information is provided about what exactly makes them arrive at their predictions. This lack of transparency and understanding, combined with domain expertise are required to make AI better applicable in practice and ensure trustworthiness on these systems.

E3: Computer Vision Expert (3 level in Dissemination material) [77]

In the automotive industry, computer vision systems have a vast potential in terms of autonomous/self-driving vehicles, driver assistance and safety precautions, and preventive measures. Computer vision acts as a sensor apparatus, that simultaneously takes in the environment around the vehicle and inside; analysing it for potential threats, obstacles, and other relevant situations that drivers need to react to while driving. Additionally, this technology can also be used to streamline traffic flow.

Using computer vision in transportation opens up exciting avenues for advanced image recognition technology, traffic flow management and making transportation easier and safer for all involved.

The field of expertise of a computer vision expert focuses on the technology, application domain knowledge and enhancement of established functionalities and services.

E4: Machine Learning Expert (3 level in Dissemination material)

With the availability of big data and deep learning methodology the performance of neuronal networks or AI systems is reaching or even exceeding the human. However, because of their nested and inscrutable learning and training approach, these systems are usually seen in black box manner and not fully trusted on how exactly decisions and interpretations are made. Due to this lack of transparency, experts on visualising, explaining and interpreting deep learning models have recently attracted increasing attention. Their job focuses on explaining predictions of deep learning models, decomposing the decision and help establishing an adequate level of trust on these systems

E5: Sensor Fusion Expert (3 level in Dissemination material) [73]

Autonomous Driving (AD) functionalities are extremely dependent on the input information about the vehicle environment. The consequence of such a setup is that today's vehicles are equipped with a growing number of environmental sensors. These include, but are not limited to, lidar, radar and camera. Through an intelligent fusion of sensor data, cutting-edge functions enable vehicle automation. From the perspective of automotive engineering, the evaluation of the appropriate sensor fusion extends far beyond the scope of traditional development and thus requires additional skills on environment, vehicle and computer vision.

E6: Automotive Engineering CAD, CAE, CAM

There are numerous methodologies for computer- and model-based development supports. In the context of novel automotive system, these engineering supports further increase in importance. The knowledge representation needs to be reusable, scalable and its construction needs to be distributed. These requirements are not easily covered by methodologies. Development of digital twins for large complex systems, components modelling and virtual simulation, multi-mode modelling and simulation are increasingly important and move from a USP to fundamentals of work.

Experts in this domain have experience in the coupling of simulation models for different platforms and engineering domains and help avoiding engineering silos and promote interdomain expertises.

E7: Automotive SPICE® Basics [1, 2, 15, 30, 31]

Modern cars have above 100 computers, software and networks, and build on mechatronic functions. Automotive SPICE has become a norm that is expected worldwide to be implemented by suppliers to prove their competence in tracing 100% of requirements to their design, implementation and testing.

E8: Connected Vehicles Expert [74]

Connected and automated cars communicate with each other, with the local environment, and with the world at large scale via radio networks and satellites. By implementing real-time connectivity to cloud computing services and new ways of providing information and entertainment, the primary function of these cars is evolving from simple transportation devices to integrated systems in a wireless connected world of things. Therefore, satellite positioning and wireless connectivity are at the heart of the bigger changes in the market. The driving experience is becoming ever more convenient. By connecting vehicles to the infrastructure also reliability, robustness and security features become of major importance. A V2X wireless connection must be operational at high speed, be equally effective in every direction, and have a range that goes well beyond line-of-sight and the boundaries of conventional vehicle development.

E9 – E11: Cybersecurity Related Competences

Cybersecurity dramatically impacts automotive systems and many security incidents are owed to design or implementation flaws. Therefore, system security evaluation, the definition of proper criteria for security evaluation and frameworks for collection and organisation of the criteria are crucial for design and development of these systems [20–23, 26, 48].

E9: Cybersecurity Engineer

This addresses elements such as

- System Threat Analysis and Cybersecurity Goals
- System Design and Vulnerability Analysis
- Software Design and Vulnerability Analysis
- Software Detailed Design and Cybersecurity
- Hardware Design and Vulnerability Analysis

E10: Cybersecurity Manager

This addresses elements such as

- Legal Aspects and Privacy
- Organisational Structure
- Cybersecurity Planning
- Life Cycle Assessment
- Cybersecurity processes and audits
- Incident Response Management
- Supply Chain Security

E11: Cybersecurity Tester

This addresses elements such as

- Cybersecurity Verification at SW level
- Safety in the Production, Operation and Maintenance
- Cybersecurity verification at system level

E12: Rubber Technologist

In modern cars, rubber is used for its unique elastic and damping properties. The performance of rubber components influences safety, lifetime and comfort, e.g. in tire, gasket and bushing applications. E.g., the design of tyres has a big influence on the energy consumption of the vehicle and thus at the battery consumption on electric cars etc. [41, 54–59].

Thorough understanding of rubber material, processing methods, behavioural phenomena and compounding methodology, is specifically required.

E13: Advanced e-Powertrain Engineer

Full electric, hybrid, and fuel cell-based powertrains require a different system architecture set up, specific electric components and control algorithms, which need to be transferred to the engineering teams.

It can impact the entire vehicle design to reach the optimum energy consumption profile for the car.

E14 – E16: Functional Safety Related Competences

The use of software and electronics in the car leads to potential hazards which need to be managed in automotive projects. Norms like ISO 26262:2018 - Road Vehicles Functional Safety Standards need to be considered in the homologation process of cars [13, 14, 16, 28, 29, 32, 42, 44, 49].

E14: Functional Safety Manager Strategy Level

This addresses elements such as

- Motivation and Introduction to Functional Safety
- Introduction to Hazard and Risk Analysis and Safety Goals
- Impact of Functional Safety on Product Design and Cost

E15: Functional Safety Manager Project Level (Advanced)

This addresses elements such as

- Safety management on organisational and project level
- Safety Requirements and Safety Case Definition
- Overview of Required Engineering and V&V Methods
- Establish and Maintain Safety Planning
- Reliability design on product and system level
- Safety in the Production, Operation and Maintenance
- Legal aspects of Safety
- Legal aspects and Liabilities
- Qualification and Certification
- Production, operation, service and decommissioning

E16: Functional Safety Engineer (Advanced)

This addresses elements such as

- System Hazard Analysis and Safety Concept
- Integrating Safety in System Design & Test
- Integrating Safety in Hardware Design & Test
- Integrating Safety in Software Design & Test

E17: Highly Automated Drive Engineer [77]

With highly automated driving the driver is not only supported but also single driving tasks are situationally covered by the system. This important interim step between assisted and autonomous driving implies multiple constraints and scenarios which makes these tasks highly sophisticate. Not only must the handover and changes between autonomous and human drive be adequately supported and timed, also certain safety and comfort measures must be taken into consideration.

Engineers working in this field of expertise do not only need to take care of HAD functionalities and smooth handover, but also need to consider user experiences and interfaces (safe and trustable change of driving mode), user comfort (perceived safety and comfort), and proper scenario identification for automation of control. Requirements for these systems become more demanding with the increasing number of road users and the increased complexity of situations.

E18 – E21: Automotive Mechatronics Related Competences

The trend of replacing traditional mechanical systems with modern embedded systems enables the deployment of more advanced control strategies, but at the same time brings new challenges in terms of understanding mechanical, electronical and embedded systems/software domains.

Automotive mechatronic systems require appropriate systematic approaches to support holistic system engineering, rather than traditional engineering approaches managing electronic, software, mechanics, and reliability features separated.

Automotive Mechatronics have the ability to evaluate and manage mutual impacts and interdisciplinary values of the domains in common and a considerable overlap among existing methods. Thus, bridges the gap between engineering domains and silos.

E18: Automotive Mechatronics Manager - Introduction to Automotive Mechatronics

Is able to address elements such as:

- Cross-domain implications
- Analysis of impact and effects of cross-domain decisions
- Impact analysis of Product Design decision and Cost implications
- Understanding of mutual implications of embedded and mechanical systems
- Reliability design on product and system level

E19: Automotive Mechatronics Manager Basic Level

Understands the basic implications of automotive mechatronic systems and the related engineering domains. On basic level the consequences of implications of decisions and cross-implications between electronic and mechanic systems can be traced and evaluated.

E20: Automotive Mechatronics Expert

An Automotive mechatronic expert has profound knowledge in the development and engineering of embedded automotive systems and mechanical engineering. The expert understands implications and is able to evaluate designs and concepts of automotive mechatronic systems for their implications on electronic components, mechanical components and control strategy requirements.

E21: Automotive Mechatronics Developer [78]

Automotive mechatronic developers are either experts on electronic hardware, software or mechanic system design and development; and know about implications for the other domains. Automotive mechatronic developers, although having a firmly field of expertise, know and understand the basic concepts and principles of the other domains and interaction with their engineering counter parts.

3.2 Production Specific

P1: Advanced Manufacturing Press line Set-up

Engineers working in an Advanced Manufacturing Press line Set-up require knowledge about an integrated Press Line, give support in a production line and repair any issue in a production line. Using data analysis tools, they understand and control a production line and prevent future issues.

P2: Automotive Engineer in Quality and Metrology

The most advanced automotive OEMs are working to automate inspection and integrate metrology data with product lifecycle management systems, statistical process control and supply chain management software.

P3 – P5: Lean Six Sigma Yellow, Green, Black Belt

Six Sigma focuses on eliminating defects in a process. Lean focuses on cutting out waste in a process by keeping only those steps that add value to the end user. With the big data that IoT will provide organisations, the tools and techniques provided by Six Sigma are more important than ever.

P6 – P7 Robotics

Within the Automotive Sector use of Robotics is now commonplace in relation a wide range of shop floor processes from welding to assembly and painting. Robotic Process Automation is also now increasingly used in relation to other areas of the automotive sector including car financing and distribution. There is a need for qualified robotics personnel, given the increasing use of robotics technology.

P6: Robotic Engineer

A robotic engineer understands automated manufacturing systems, maintenance of robotic systems, implementing new machine software and technologies to further improve the robotic system, can do software simulations and design, and integrate new technologies.

P7: Robotic Technician

Robotics Technician have a particular focus on practice in a work place with a robotics environment of production. They can program logics of robots, can understand the machinery process controlled by the robot and can identify problems and solve them. They also understand the related health and safety norms.

P8: Technical Cleanliness Expert

Nowadays production of highly complex electronic components requires an extreme cleanliness in the production and work place.

3.3 Innovation Specific

G1: Working in Automotive (Automotive Engineer)

Automotive industry applies specific norms, has specific understanding of mechatronic functions, root cause analysis, error reporting, design principles and quality principles to be followed. Therefore, a very important step is to motivate young engineers to enter automotive and to prepare already at school and university to the automotive specific understandings.

G2: Automotive Quality Engineer (AQUA) [17–19, 46]

Automotive Quality has to deal with the integrated complexity of software, electronics, mechanics and services. Automotive Quality Integrated skills combine Automotive SPICE [12], Functional Safety, Six Sigma based production quality, and cybersecurity.

G3: Innovation Agent for Business Model Innovation

New use of smart technologies and car architectures lead to new business models. It is decisive to create new business innovation models for creating new value chains and new services that can be sold [3–11, 27, 34–37, 39, 52].

3.4 Maintenance Specific

M1: Automotive Engineer in Tool and Die Production and Maintenance

Any smart production nowadays is using CAD and digital elements (DIEs) designed and production automatically adjusted to the digital parameters of the elements. This will impact the design of elements and tools by having a more and more digital view on the production.

M2: Sustainability manager

Sustainability manager are able to reflect the expectations of stakeholders to the requirements of the business development of an organisation for environmentally friendly behaviour of the employees, their social security and the production environment.

Sustainability manager provide important targets, measures and business strategies on the way to a sustainable development. The focus is on increasing sustainable production and a reduction of energy and material consumption. The massive automotive supply chain implies additional means for sustainability management and improvement. In the past, the main focus of sustainability was the reduction of environmental impacts in the use phase and later also decommission. Now the topics of second life and alternative reuse emerged. When the targets for achieving sustainable development are set, the methodology and tools to be used and to assess the sustainability performance are the focus of sustainability managers.

M3: Predictive Maintenance Expert [78]

Logging and monitoring are the basic capabilities of an application to track and trace the state of objects, discover information regarding its past states and potentially estimate future states. Predictive maintenance experts are investigating techniques of non-intrusive logging and tracing of events from systems and establish adequate monitoring for getting run-time data from traceable objects, storing data and make data analysis, as well as ensuring the required level of safety and prediction of maintenance. In terms of providing a high quality of answers to predictive maintenance, aggregate monitoring queries, finding surprising levels of a data stream, and detecting bursts is not sufficient. Predictive maintenance experts also include domain knowledge, design expertise, and similarity queries, such as detecting correlations and finding interesting patterns to predict remaining lifetimes and future failure and fault detection.

4 Outlook

New technologies and needed competences will be emerging more often than in the past and therefore the continuous assessment of knowledge base and its update is needed to stay competitive on the market. Both from the producer perspective, as an automotive/mobility provider, and from the human resources point of view, as getting knowledgeable workforce.

Like some of the known examples e.g. the already mentioned Nikola and SONY Vision car S illustrated, there will be new and forthcoming knowledge and competence partnerships driving vehicle design. It is important for European car makers to transform their system development strategy into a knowledge intelligence architecture for future car design and add this view on top of the traditional component and functional view of a car. More than any time before, it is key challenge for all stakeholders to cooperate on upskilling and reskilling of automotive workforce to tackle those changes.

Acknowledgements. We are grateful to the European Commission which has funded the Blueprint project Development and Research on Innovative Vocational Skills (DRIVES, 2018-2021, www.project-drives.eu). The project DRIVES – Project number 591988-EPP-1-2017-1-CZ-EPPKA2-SSA-B is co-funded by the Erasmus + Programme of the European Union.

The European Commission support for the production of this publication under the Grant Agreement Nº 2017-3295/001-001 does not constitute an endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

We are grateful to the partners in DRIVES project, especially of the Work Package 3 - Skills Framework of the DRIVES project, for their inputs during the project work.

Work is partially supported by Grant of SGS No. SP2020/62, VŠB - Technical University of Ostrava, Czech Republic.

We are grateful to a working party of Automotive suppliers SOQRATES [45] (www.soqrates.de) who exchange knowledge about such assessment strategies. This includes: Böhner Martin (Elektrobit), Brasse Michael(HELLA), Bressau Ernst (BBraun), Dallinger Martin (ZF), Dorociak Rafal (HELLA), Dreves Rainer (Continental Automotive), Ekert Damjan (ISCN), Forster Martin (ZKW), Geipel Thomas (BOSCH), Grave Rudolf (Elektrobit), Griessnig Gerhard (AVL), Gruber Andreas (ZKW), Habel Stephan (Continental Automotive), Hällmayer Frank (Software Factory), Haunert Lutz (Giesecke & Devrient), Karner Christoph (KTM), Kinalzyk Dietmar (AVL), König Frank (ZF), Lichtenberger Christoph (MAGNA ECS), Lindermuth Peter (Magna Powertrain), Macher Georg (TU Graz & ISCN), Mandic Irenka (Magna Powertrain), Maric Dijas (Lorit Consultancy), Mayer Ralf (BOSCH Engineering), Mergen Silvana (TDK/EPCOS), Messnarz Richard (ISCN), Much Alexander (Elektrobit), Nikolov Borislav (msg Plaut), Oehler Couso Daniel (Magna Powertrain), Riel Andreas (Grenoble INP & ISCN), Rieß Armin (BBraun), Santer Christian (AVL), Schlager Christian (Magna ECS), Schmittner Christoph (Austrian Institute of Technology AIT), Schubert Marion (ZKW), Sechser Bernhard (Process Fellows), Sokic Ivan (Continental Automotive), Sporer Harald (Infineon), Stahl Florian (AVL), Wachter Stefan (msg Plaut), Walker Alastair (Lorit Consultancy), Wegner Thomas (ZF).

References

1. Automotive SPICE © 3.1, Process Assessment Model, VDA QMC Working Group 13/Automotive SIG, November 2017
2. Automotive SPICE © Guidelines, 2nd edn. Nov 2017, VDA QMC Working Group 13, November 2017
3. Miklos, B., Richard, M., Davidson Alfred, G.: The impact of national cultural factors on the effectiveness of process improvement methods: the 3rd dimension. In: Proceedings of the 11th ICSQ Conference, ASQ, USA (2001)
4. Messnarz, R., Tully, C.J., Biro, M.: Better Software Practice for Business Benefit: Principles and Experiences. IEEE Computer Society Press, Los Alamitos (1999)
5. Korsaa, M., et al.: The SPI manifesto and the ECQA SPI manager certification scheme. J. Softw. Evol. Process **24**(5), 525–540 (2012)
6. Korsaa, M., et al.: The people aspects in modern process improvement management approaches. J. Softw. Evol. Process **25**(4), 381–391 (2013). Special Issue: Selected Industrial Experience Papers of EuroSPI 2010
7. Messnarz, R., Ekert, D., Reiner, M., O'Suilleabhain, G.: Human resources based improvement strategies—the learning factor. J. Softw. Evol. Process **13**(4), 355–362 (2008)
8. Messnarz, R., Nadasi, G., O'Leary, E., Foley, B.: Experience with teamwork in distributed work environments. In: Smith, B.S., Chiozza, E. (eds.) Proceedings of the E2001 Conference, E-Work and E-commerce, Novel Solutions for a Global Networked Economy. IOS Press, Amsterdam (2001)
9. Messnarz, R., et al.: Social responsibility aspects supporting the success of SPI. J. Softw. Evol. Process **26**(3), 284–294 (2014). Special Issue: Software Process Assessment and Improvement (EuroSPI 2011)
10. Messnarz, R., Spork, G., Riel, A., Tichkiewitch, S.: Dynamic learning organisations supporting knowledge creation for competitive and integrated product design. In: Proceedings of the 19th CIRP Design Conference – Competitive Design, Cranfield University, 30–31 March 2009, p. 104 (2009)
11. Biró, M., Messnarz, R.: Key success factors for business based improvement. In: Proceedings of the EuroSPI 1999 Conference, Pori 1999, (Pori School of Technology and Economics. Ser. A., 25) (1999)
12. Höhn, H., Sechser, B., Dussa-Zieger, K., Messnarz, R., Hindel, B.: Software Engineering nach Automotive SPICE. Entwicklungsprozesse in der Praxis-Ein Continental-Projekt auf dem Weg zu Level 3, Kapitel: Systemdesign, dpunkt. Verlag (2015)
13. ISO - International Organization for Standardization. ISO 26262 Road vehicles Functional Safety Part 1–10 (2011)
14. ISO – International Organization for Standardization. ISO CD 26262-2018 2nd Edition Road vehicles Functional Safety (2018)
15. ISO/SAE 21434, Road vehicles – Cybersecurity engineering, ISO and SAE, Committee Draft (CD) (2018)
16. KGAS, Konzerngrundanforderungen Software, Version 3.2, Volkswagen LAH 893.909: KGAS_3602, KGAS_3665, KGAS_3153, KGAS_3157, November 2018
17. Christian, K., Messnarz, R., Riel, A., et al.: The AQUA automotive sector skills alliance: best practice in an integrated engineering approach. Softw. Qual. Professional **17**(3), 35–45 (2015)
18. Messnarz, R., et al.: Integrating functional safety, automotive spice and six sigma – the AQUA knowledge base and integration examples. In: Barafort, B., O'Connor, Rory V., Poth, A., Messnarz, R. (eds.) EuroSPI 2014. CCIS, vol. 425, pp. 285–295. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-662-43896-1_26

19. Kreiner, C., et al.: Automotive knowledge alliance AQUA – integrating automotive SPICE, six sigma, and functional safety. In: McCaffery, F., O'Connor, Rory V., Messnarz, R. (eds.) EuroSPI 2013. CCIS, vol. 364, pp. 333–344. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-39179-8_30
20. Macher, G., Sporer, H., Brenner, E., Kreiner, C.: Supporting cyber-security based on hardware-software interface definition. In: Kreiner, C., O'Connor, Rory V., Poth, A., Messnarz, R. (eds.) EuroSPI 2016. CCIS, vol. 633, pp. 148–159. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-44817-6_12
21. Macher, G., Messnarz, R., Kreiner, C., et al.: Integrated safety and security development in the automotive domain, working group 17AE-0252/2017-01-1661. In: SAE International, June 2017
22. Macher, G., Much, A., Riel, A., Messnarz, R., Kreiner, C.: Automotive SPICE, safety and cybersecurity integration. In: Tonetta, S., Schoitsch, E., Bitsch, F. (eds.) SAFECOMP 2017. LNCS, vol. 10489, pp. 273–285. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66284-8_23
23. Macher, G., Sporer, H., Berlach, R., Armengaud, E., Kreiner, C.: SAHARA: a security-aware hazard and risk analysis method. In: Design, Automation Test in Europe Conference Exhibition (DATE), 2015, pp. 621–624, March 2015
24. Macher, G., Brenner, E., Messnarz, R., Ekert, D., Feloy, M.: Transferable competence frameworks for automotive industry. In: Walker, A., O'Connor, Rory V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 151–162. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_12
25. Messnarz, R., Kreiner, C., Macher, G., Walker, A.: Extending automotive SPICE 3.0 for the use in ADAS and future self-driving service architectures. J. Softw. Evol. Process **30**(5), March 2018
26. Messnarz, R., Kreiner, C., Riel, A.: Integrating automotive SPICE, functional safety, and cybersecurity concepts: a cybersecurity layer model. Softw. Qual. Professional **18**, 13–23 (2016)
27. Messnarz, R., Spork, G., Riel, A., Tichkiewitch, S.: Dynamic learning organisations supporting knowledge creation for competitive and integrated product design. In: Proceedings of the 19th CIRP Design Conference – Competitive Design, Cranfield University, p. 104, 30–31 March 2009
28. Messnarz, R., Kreiner, C., Riel, A., et al.: Implementing functional safety standards has an impact on system and SW design - required knowledge and competencies (SafEUr). In: Software Quality Professional (2015)
29. Messnarz, R., et al.: Implementing functional safety standards – experiences from the trials about required knowledge and competencies (SafEUr). In: McCaffery, F., O'Connor, Rory V., Messnarz, R. (eds.) EuroSPI 2013. CCIS, vol. 364, pp. 323–332. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-39179-8_29
30. Messnarz, R., Sehr, M., Wüstemann, I., Humpohl, J., Ekert, D.: Experiences with SQIL – SW quality improvement leadership approach from Volkswagen. In: Stolfa, J., Stolfa, S., O'Connor, Rory V., Messnarz, R. (eds.) EuroSPI 2017. CCIS, vol. 748, pp. 421–435. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64218-5_35
31. Messnarz, R., König, F., Bachmann, V.O.: Experiences with trial assessments combining automotive SPICE and functional safety standards. In: Winkler, D., O'Connor, Rory V., Messnarz, R. (eds.) EuroSPI 2012. CCIS, vol. 301, pp. 266–275. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-31199-4_23
32. Messnarz, R., Kreiner, C., Riel, A.: Integrating automotive SPICE, functional safety, and cybersecurity concepts: a cybersecurity layer model. Softw. Qual. Professional **18**(4), 13–23 (2016)

33. Messnarz, R., Much, A., Kreiner, C., Biro, M., Gorner, J.: Need for the continuous evolution of systems engineering practices for modern vehicle engineering. In: Stolfa, J., Stolfa, S., O'Connor, Rory V., Messnarz, R. (eds.) EuroSPI 2017. CCIS, vol. 748, pp. 439–452. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64218-5_36
34. Messnarz, R., O'Suilleabhain, G., Coughlan, R.: From process improvement to learning organisations. J. Softw. Evol. Process **11**(3), 287–294 (2006). https://doi.org/10.1002/spip.272. Special Issue: Special Issue on SPI Industrial Experience
35. Messnarz, R., Horvat, R.V., Harej, K., Feuer, E.: ORGANIC-Continuous Organisational Learning in Innovation and Companies. In: Proceedings of the E2005 Conference, E-Work and E-commerce, Novel solutions for a global networked economy, eds. Brian Stanford Smith, Enrica Chiozza, IOS Press, Amsterdam, Berlin, Oxford, Tokyo, Washington (2005)
36. Gavenda, M., et al.: Fostering innovation and entrepreneurship in European VET: EU project "from idea to enterprise". In: McCaffery, F., O'Connor, Rory V., Messnarz, R. (eds.) EuroSPI 2013. CCIS, vol. 364, pp. 282–293. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-39179-8_25
37. Sauberer, G., Riel, A., Messnarz, R.: Diversity and PERMA-nent positive leadership to benefit from industry 4.0 and Kondratieff 6.0. In: Stolfa, J., Stolfa, S., O'Connor, Rory V., Messnarz, R. (eds.) EuroSPI 2017. CCIS, vol. 748, pp. 642–652. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64218-5_53
38. Biró, M., Messnarz, R., Colomo-Palacios, R.: Software process improvement leveraged in various application domains. J. Softw. Evol. Process **26**(5), 465–467 (2014). Special Issue: Software process improvement leveraged in various application domains
39. Messnarz, R., et al.: InnoTEACH – applying principles of innovation in school. In: Stolfa, J., Stolfa, S., O'Connor, Rory V., Messnarz, R. (eds.) EuroSPI 2017. CCIS, vol. 748, pp. 294–301. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64218-5_24
40. Messnarz, R., et al.: Shifting paradigms in innovation management – organic growth strategies in the cloud. In: Walker, A., O'Connor, Rory V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 28–42. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_3
41. Messnarz, R., Ekert, D., Grunert, F., Blume, A.: Cross-cutting approach to integrate functional and material design in a system architectural design – example of an electric powertrain. In: Walker, A., O'Connor, Rory V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 322–338. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_25
42. Messnarz, R., et al.: Integrated automotive SPICE and safety assessments. WILEY Softw. Process Improv. Pract. **14**(5), 279–288 (2009). https://doi.org/10.1002/spip.429. Special Issue: Part 1: Special Issue on SPI Experiences and Innovation for Global Software Development
43. Much, A.: Automotive Security: Challenges, Standards and Solutions, Software Quality Professional, September 2016
44. SAE J3061. Cybersecurity Guidebook for Cyber-Physical Vehicle Systems, SAE - Society of Automotive Engineers, USA, January 2016
45. SOQRATES. Task Forces Developing Integration of Automotive SPICE, ISO 26262 and SAE J3061. http://soqrates.eurospi.net/
46. Stolfa, J., et al.: Automotive quality universities - aqua alliance extension to higher education. In: Kreiner, C., O'Connor, Rory V., Poth, A., Messnarz, R. (eds.) EuroSPI 2016. CCIS, vol. 633, pp. 176–187. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-44817-6_14

47. GEAR 2030, European Commission, Commission launches GEAR 2030 to boost competitiveness and growth in the automotive sector, 2016. http://ec.europa.eu/growth/tools-databases/newsroom/cf/itemdetail.cfm?item_id=8640
48. Riel, A., Kreiner, C., Messnarz, R., Much, A.: An architectural approach to the integration of safety and security requirements in smart products and systems design. CIRP Ann. **67**(1), 173–176 (2018)
49. Riel, A., et al.: EU project SafEUr – competence requirements for functional safety managers. In: Winkler, D., O'Connor, Rory V., Messnarz, R. (eds.) EuroSPI 2012. CCIS, vol. 301, pp. 253–265. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-31199-4_22
50. Riel, A., Tichkiewitch, S., Messnarz, R.: The profession of integrated engineering: formation and certification on a European level. Acad. J. Manuf. (2008)
51. Riel, A., Draghici, A., Draghici, G., Grajewski, D., Messnarz, R.: Process and product innovation needs integrated engineering collaboration skills. J. Softw. Evol. Process **24**(5), 551–560 (2012)
52. Feuer, É., Messnarz, R.: ORGANIC - continuous organisational learning in innovation and companies. In: ICETA 2005. 4th International Conference on Emerging E-Learning Technologies and Applications, Proceedings. Kosice (2005)
53. Clarke, P., et al.: An investigation of software development process terminology. In: Clarke, Paul M., O'Connor, Rory V., Rout, T., Dorling, A. (eds.) SPICE 2016. CCIS, vol. 609, pp. 351–361. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-38980-6_25
54. Bratzel, S.: Kautschuk Gummi Kunststoffe, vol. 03, pp. 10–11 (2019)
55. Kunststoffe im Auto – was geht noch? GAK 4/2013 Jg. 66, pp. 248–258 (2013)
56. Reina, G.: Tyre pressure monitoring using a dynamical model-based estimator. Veh. Syst. Dyn. **29** (2015). https://doi.org/10.1080/00423114.2015.1008017
57. McIntosh, J.: Run-Flats vs Self-Sealing Tires. Autotrader Canada, 24 October 2017. Accessed 15 Feb 2019
58. https://2020.eurospi.net/index.php/manifesto. Accessed 17 Apr 2020
59. https://www.energie-lexikon.info/rollwiderstand.html. Accessed 6 Apr 2019
60. EU Blueprint Project DRIVES. https://www.project-drives.eu/. Accessed 6 Apr 2020
61. Much, A.: Automotive security: challenges, standards, and solutions. Softw. Qual. Professional **18**(4), 4–12 (2016)
62. GEAR 2030, High Level Group on the Competitiveness and Sustainable Growth of the Automotive Industry in the European Union (2017)
63. European Sector Skill Council: Report, Eu Skill Council Automotive Industry (2013)
64. Messnarz, R., Ekert, D.: Assessment-based learning systems - learning from best projects. Softw. Process Improv. Pract. **12**(6), 569–577 (2007). https://doi.org/10.1002/spip.347. Special Issue: Special Issue on Industrial Experiences in SPI
65. Messnarz, R., Ekert, D., Zehetner, T., Aschbacher, L.: Experiences with ASPICE 3.1 and the VDA automotive SPICE guidelines – using advanced assessment systems. In: Walker, A., O'Connor, Rory V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 549–562. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_42
66. SONY Vision S car. https://www.derbrutkasten.com/sony-ces-2020-auto-vision-s/. Accessed 17 April 2020
67. Nikola. https://de.wikipedia.org/wiki/Nikola_Motor_Company. Accessed 17 April 2020
68. Stolfa, J., et al.: DRIVES—EU blueprint project for the automotive sector—a literature review of drivers of change in automotive industry. J. Softw. Evol. Process **32**(3), March 2020. Special Issue: Addressing Evolving Requirements Faced by the Software Industry

69. Google Announces Major Automotive Partnership to Integrate Its Android System Into Future Cars. https://interestingengineering.com/google-plans-to-integrate-its-android-system-into-future-renault-nissan-mitsubishi-cars. Accessed 24 Apr 2020
70. Google partners with the world's biggest auto group to bring Android to cars. https://www.theverge.com/2018/9/18/17873502/google-android-os-cars-renault-nissan-mitsubishi-alliance. Accessed 24 Apr 2020
71. Volvo Cars collaborates with Google on a brand new infotainment system. https://group.volvocars.com/news/connectivity/2018/volvo-cars-collaborates-with-google-on-a-brand-new-infotainment-system. Accessed 24 Apr 2020
72. Macher, G., Diwold, K., Veledar, O., Armengaud, E., Römer, K.: The quest for infrastructures and engineering methods enabling highly dynamic autonomous systems. In: Walker, A., O'Connor, Rory V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 15–27. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_2
73. Macher, G., Druml, N., Veledar, O., Reckenzaun, J.: Safety and security aspects of fail-operational urban surround perceptION (FUSION). In: Papadopoulos, Y., Aslansefat, K., Katsaros, P., Bozzano, M. (eds.) IMBSA 2019. LNCS, vol. 11842, pp. 286–300. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32872-6_19
74. Schmittner, C., et al.: Smart industrial indoor farming-Technical and societal challenges. In: IDIMT 2019: Innovation and Transformation in a Digital World-27th Interdisciplinary Information Management Talks, Trauner Verlag Universitat, pp. 401–409 (2019)
75. Schmittner, C., Macher, G.: Automotive cybersecurity standards - relation and overview. In: Romanovsky, A., Troubitsyna, E., Gashi, I., Schoitsch, E., Bitsch, F. (eds.) SAFECOMP 2019. LNCS, vol. 11699, pp. 153–165. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-26250-1_12
76. Armengaud, E., et al.: Development framework for longitudinal automated driving functions with off-board information integration arXiv preprint arXiv:1906.10009 (2019)
77. Innerwinkler, P., et al.: TrustVehicle – improved trustworthiness and weather-independence of conditionally automated vehicles in mixed traffic scenarios. In: Dubbert, J., Müller, B., Meyer, G. (eds.) AMAA 2018. LNM, pp. 75–89. Springer, Cham (2019). https://doi.org/10.1007/978-3-319-99762-9_7
78. Veledar, O., Damjanovic-Behrendt, V., Macher, G.: Digital twins for dependability improvement of autonomous driving. In: Walker, A., O'Connor, Rory V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 415–426. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_32

Analysis of Asymmetry in the Resonant Stage of Multiphase Converters for E-Mobility Charging Station

Dimitar Arnaudov[1(✉)], Marin B. Marinov[2], and Maria Stefanova-Pavlova[3]

[1] Department of Power Electronics, Technical University of Sofia, 8, Kliment Ohridski Blvd., 1756 Sofia, Bulgaria
dda@tu-sofia.bg
[2] Department of Electronics, Technical University of Sofia, 8, Kliment Ohridski Blvd., 1756 Sofia, Bulgaria
mbm@tu-sofia.bg
[3] Center for Innovation and Technology Transfer-Global, Jahroslav Veshin Str., Bl.32/B, 1408 Sofia, Bulgaria
stefanova-pavlova@citt-global.net

Abstract. The paper presents a study of a multiphase converter for charging stations for electric vehicles. The converter consists of parallel operating units based on high-frequency resonant inverters. Simulation models of multiphase converters have been developed. Simulation studies have been conducted to demonstrate the advantages and disadvantages of these systems. The impact of the asymmetry of the individual units on the pulsations of the charging current is investigated, as well as the possibilities for regulating the value of the charging current by changing the parameters of one or several units. The advantages of a five-phase converter over six-phase one at the same power output are shown.

Keywords: Charging station · Multiphase converter · Energy storage · Asymmetry · Resonant converters

1 Multiphase Converters in Charging Stations

1.1 Introduction

Charging stations for electric vehicles are implemented in a modular structure. In this way, the output power is increased by connecting in parallel operating modules, while at the same time higher reliability is achieved. Galvanic separation is achieved through high-frequency transformers. The connection of the individual units so that they work on a common load is realized by connecting the secondary windings of the transformers. The connection may be in series [1], in parallel [2], or a common point of the secondary windings [3], thus realizing a multiphase source. A single-phase rectifier is connected to the multiphase source. Most often, this rectifier is a three-phase one. The specifics of the operation of the rectifier are because it is powered by current sources. The current sources are realized through resonance inverters. Changing the current in

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 370–380, 2020.
https://doi.org/10.1007/978-3-030-56441-4_27

one source causes a change in the other, which is a disadvantage and that is the reason why they are less commonly used.

Another way of realizing a multiphase source is by connecting the secondary coils of the high-frequency transformers to separate rectifiers, and the outputs of the rectifiers are combined and connected to the load - the energy storage element of the electric vehicle (battery or supercapacitor) [4, 5].

1.2 Multiphase Topologies

The DC fast-charging stations for electric vehicles include rectifiers with PFC systems and multiphase DC-DC converters. Figure 1 shows the block diagram of a multiphase DC-DC converter as part of an electric vehicle charging station system [6]. The concept of using a common rectifier powered by a multiphase source is shown. Figure 2 shows a circuit diagram of a multiphase rectifier used in this concept, the rectifier being powered by power sources. The secondary windings of the individual high-frequency transformers are connected in a star and supply the multiphase rectifier. The features of the circuit operation mode, advantages and disadvantages are discussed in [6, 7]. A major disadvantage of that topology is the strong interplay between units in the multiphase system. Very often, in addition to transmitting energy through high-frequency transformers, system management information is also transmitted, such as in the case of contactless charging of electric vehicles [8]. This is another reason to investigate the effect of asymmetry on the harmonic composition since high-frequency noise is mixed with the information signal of the control systems [9].

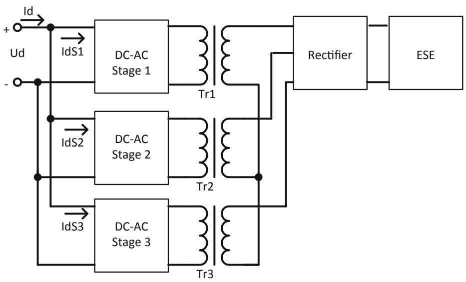


Fig. 1. Block diagram of the single multiphase rectifier topology (Topology I).

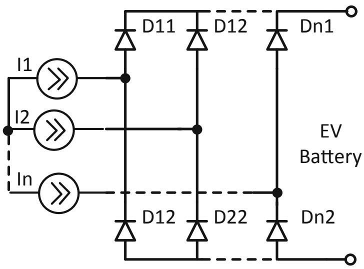


Fig. 2. Circuit of the single multiphase rectifier topology (Topology I).

The block diagram of a multiphase DC-DC converter with parallel-connected output rectifiers is shown in Fig. 3 [6]. Again, three-phase systems are most commonly used here. Figure 4 shows the wiring diagram of the output rectifiers.

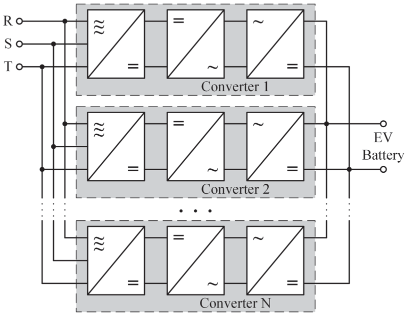


Fig. 3. Block diagram of the converter with individual rectifiers topology (Topology II).

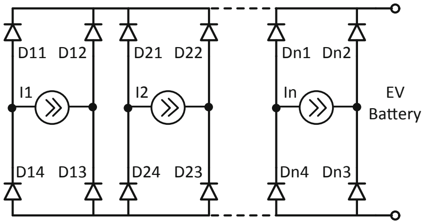


Fig. 4. Circuit of the multiphase rectifier topology connected in parallel (Topology II).

This work will explore the five-phase and six-phase variants of the schemes of Fig. 3 and Fig. 4. In practice, synchronous rectifiers are used to reduce losses, where the diodes are replaced by a MOS transistor with a built-in reverse diode. To study this scheme, simulation models in LTSpice software have been developed.

2 Modeling of a Multiphase Converter for Charging of Energy Storage Elements

Figure 5 shows a simulation model of a six-phase power source – (Topology IV). It can also be realized as two three-phase sources, whose controls are defaced at an angle of $2\pi/6$.

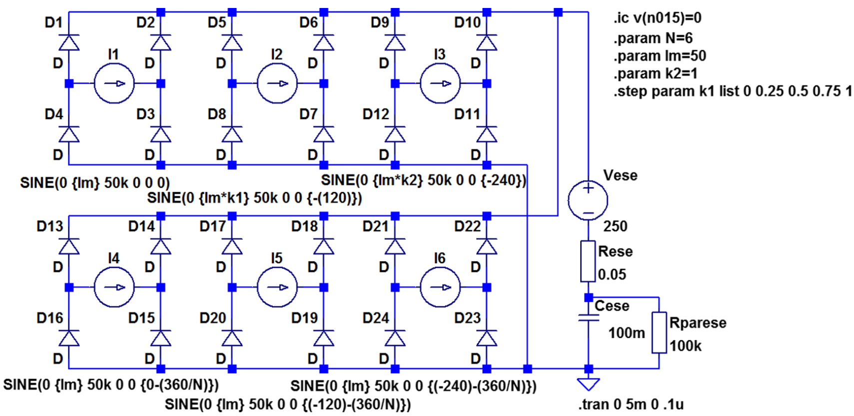


Fig. 5. Six stage rectifier circuit (six-phase rectifier – Topology IV).

In the simulation model, the sinusoidal current sources $I1 \div I6$ replace the secondary windings of the high-frequency transformers. The primary windings are connected in series in the resonant circuit of resonant inverters. The $Vese, Rese, Cese$, $Rparese$ elements are equivalent to the replacement of an energy storage unit (battery). The $Cece$ value is significantly reduced to illustrate the charging processes for smaller simulation times.

Figure 6 shows a simulation model of a five-phase power source – (Topology III). Here, the dephasing between the sinusoidal sources is $2\pi/5$. Simulation models have been developed with the ability to perform parametric analysis.

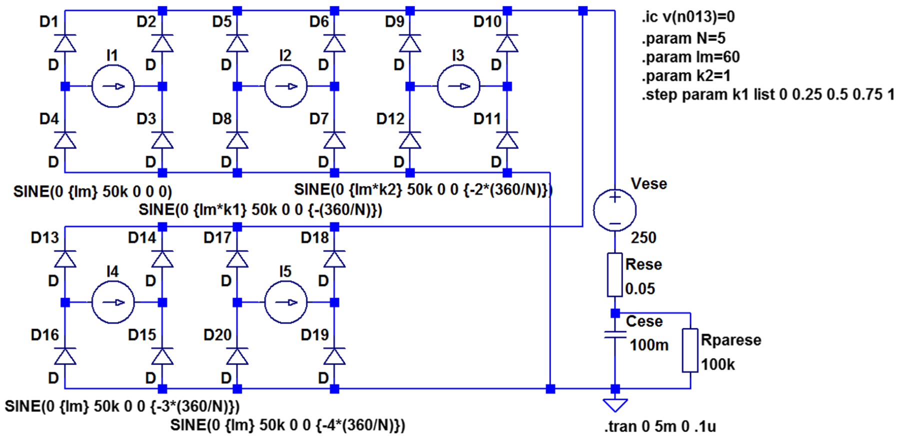


Fig. 6. Five stage rectifier circuit (six-phase rectifier – Topology III).

Figure 7 shows the waveforms of the currents of the five current sources.

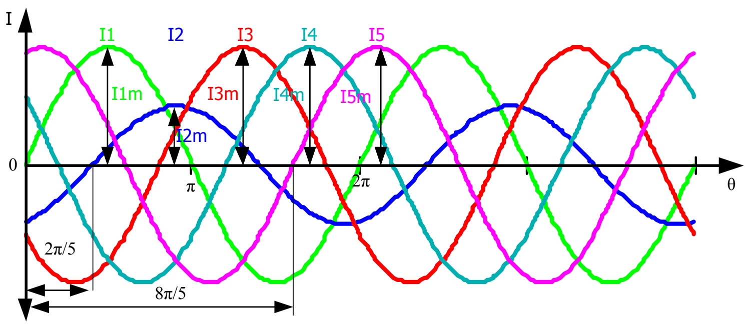


Fig. 7. Five stage rectifier circuit (six-phase rectifier – Topology III).

3 Experimental Results

Simulation studies have been conducted to change the amplitude of one and two of the sources supplying the individual rectifiers. The ripples of the current through the battery and the change in its average value were measured.

Figure 8 and Fig. 9 show the results of the simulation studies in the symmetrical operation of the units (the amplitudes of the currents of the individual sources are the same). Tests were made for both converters (five-phase and six-phase) with an average load current of $190A$. In order to obtain the same output currents of the converter, the maximum current value of each of the sources in the five-unit variant must be by 20% higher than that of the six-unit converter. Figure 8 shows the currents through the diodes $D1, D5, D9, D13, D17 - I(D1), I(D5), I(D9), I(D13), I(D17)$ and the current through the capacitor $Cese - I(Cese)$ for a five stage converter. Figure 9 shows the results for the currents through diodes $D1, D5, D9, D13, D17, D21$ -$I(D1), I(D5), I(D9), I(D13), I(D17), I(D21)$ and the current through the $Cese$ - $I(Cese)$ capacitor for six stage circuits.

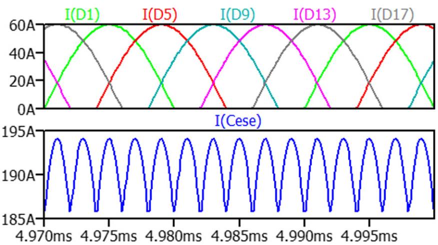


Fig. 8. Results for the currents through diodes and the current through the charged battery for Topology III.

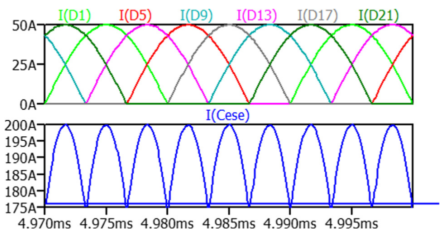


Fig. 9. Results for the currents through diodes and the current through the charged battery for Topology IV.

Figure 10 and Fig. 11 show the results of a parametric analysis when changing the amplitude value of one source.

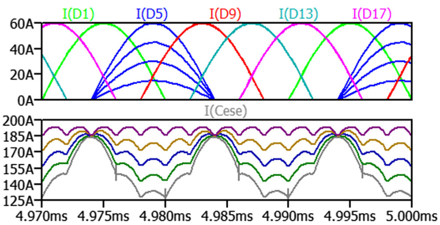


Fig. 10. Simulation results Topology III.

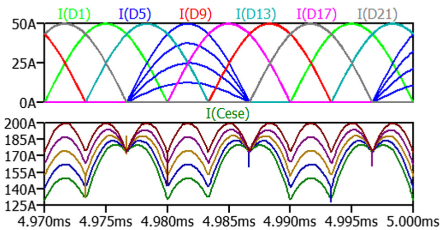


Fig. 11. Simulation results Topology IV.

The results are given at amplitude $Im0, 0.25, 0.5, 0.75$, and 1 relative to the nominal value of the remaining currents. Source $I2$ current value is changed. This is seen in the results as a change in the value of the current through diode $D5$ - $I(D5)$. The waveforms of battery current $I(Cese)$ have the smallest ripples at symmetric sources and the highest ripples at an amplitude equal to 25% of the nominal. A peculiarity of the waveforms for the battery current is that the Topology III (rectifier with 5 stages) ripple rate is 10 times higher than the frequency of the power sources, whereas for Topology IV (rectifier with 6 stages) - the frequency is 6 times higher. When the amplitude of one of the sources is reduced below 50% of the nominal for Topology III, the pulsations are twice the frequency of the power sources. The same is observed with Topology IV.

Simulation studies were performed at a given amplitude of one of the power sources - 50% of the nominal (current source $I3$), and at another source (current source $I2$) the amplitude Im was changed respectively to $0, 0.25, 0.5, 0.75$ and 1 relative to the nominal value of other currents. The results for Topology III are shown in Fig. 12, and for Topology IV in Fig. 13. The change in the maximum value of the current through the source $I2$ can be seen from the time diagrams of the current through the diode $D5$ - $I(D5)$. The fact that the current through source $I3$ has an amplitude of 0.5 relative to the nominal current can be read from the current diagram through diode $D9$ - $I(D9)$.

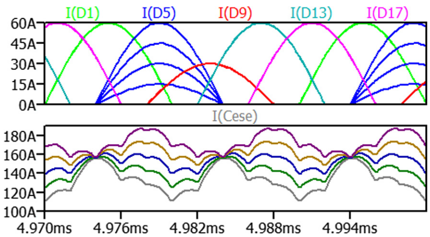


Fig. 12. Simulation results Topology – III.

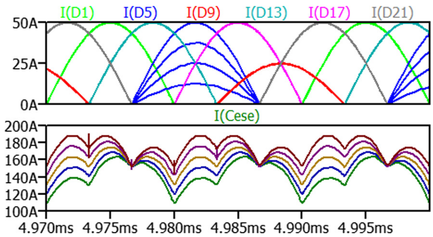


Fig. 13. Simulation results Topology – IV.

From the analysis of the results of the waveforms in Fig. 12, it is again visible for Topology III that we have a pulse rate 10 times higher than that of the current source, as well as harmonics with twice the frequency. For Topology IV, the harmonics with the highest amplitude are with 6 times higher frequency and twice higher frequency, respectively.

A simulation study was performed to obtain the control characteristic for changing the amplitude of one of the units. The load current dependence was measured as a function of the maximum current value through one of the sources $Idr = F(Imr)$. This dependence in relative units is plotted in Fig. 14. The y-axis is the load current in relative units. In producing the graph the load current value is related to the value when all sources have the same maximum values. The abscissa is the maximum current value of the source at which we change the value. It is also in relative units, relative to the value at equal maximum values of all sources. The study clearly shows that the dependence is linear.

Figure 14 presents the *Id*5 characteristic for a five-unit converter Topology III, and *Id*6 is for the six-unit Topology IV converter. The characteristics also show that in the five-unit scheme, the average load current through the load decreased by 20% $(1/5)$ and in the six-unit scheme by 16.6% $(1/6)$. The linearity of the characteristics is also preserved when the values of the two units are changed, which confirms the possibility of controlling the battery charging current by adjusting the values of the individual units.

$$Idr = \frac{Id_{asymetry}}{Id_{symetry}} \tag{1}$$

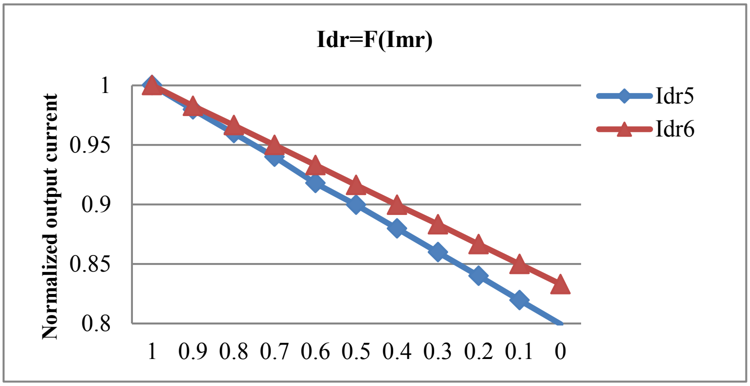


Fig. 14. Control characteristic.

In the simulation study of the two topologies, the pulses of the peak-to-peak currents were also measured, as shown in Fig. 15. Figure 16 depicts the dependences on the ripple coefficients as a function of the change in the maximum current value of one of the sources $Kr = F(Imr)$. The change in the maximum value is presented in relative units.

$$K_{RIPPLE} = \frac{\Delta I_{P-P}}{2I_d}. \tag{2}$$

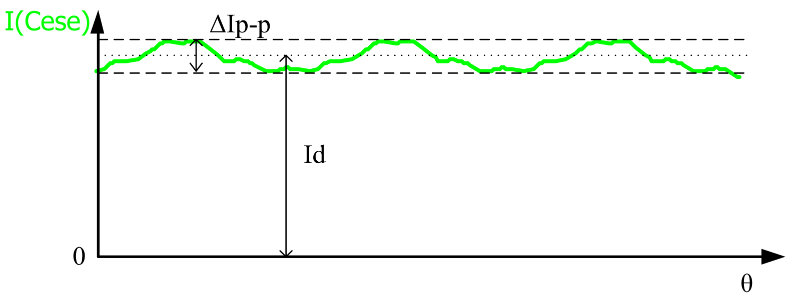


Fig. 15. Ripple.

From the results shown in Fig. 16, it can be seen that, with a symmetrically working circuit, the current ripples for Topology III are three times smaller than that of Topology IV.

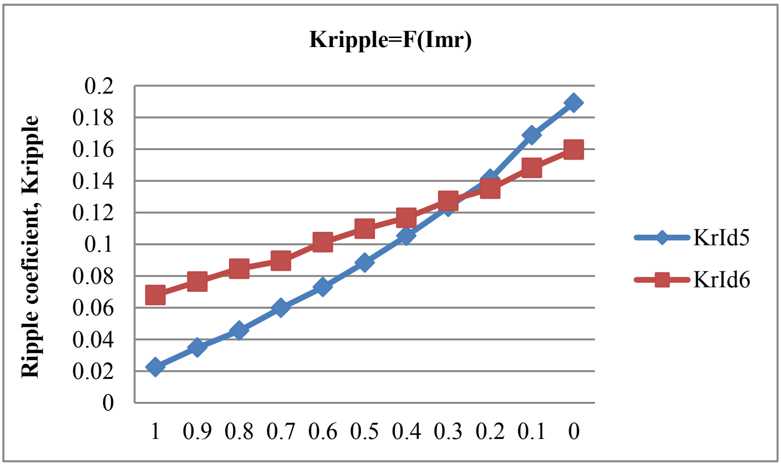


Fig. 16. Ripple coefficient.

Figure 17 and Fig. 18 show the currents through the charged batteries $I(Cese)$ and the voltage across them during charging. In the simulation model, the value of the battery capacity is greatly reduced to illustrate the change in battery voltage for short times of simulating the operation of the converters.

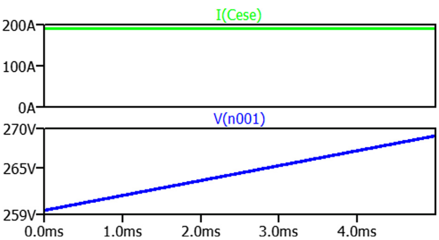


Fig. 17. Simulation results for Topology III.

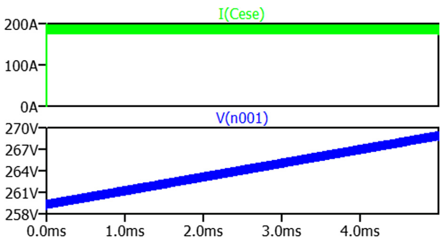


Fig. 18. Simulation results for Topology IV.

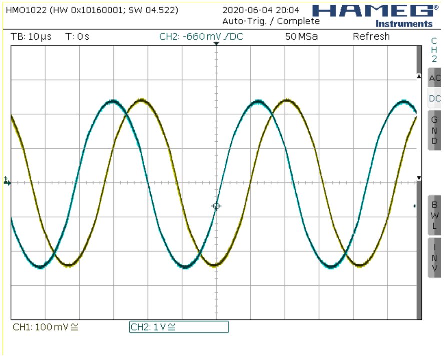


Fig. 19. Currents from tow stages.

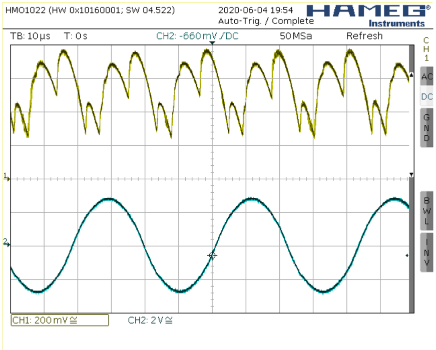


Fig. 20. Current through the energy storage component and a current power source.

Figure 19 and Fig. 20 show the results of an experimental study of a five-unit laboratory stand with low power. Figure 19 shows the currents through two of the units. The measurements were made by current probes with settings 1 mA/mV and 10 mA/mV. The currents of two units with equal amplitudes are shown. Figure 20 shows the currents through the energy storage element (channel 1) measured on a 0.1 ohm resistor. The current through one of the units is also measured (channel 2), with a current probe of 1 mA/mV.

4 Conclusion

This paper presents a study of multiphase power sources for an electric vehicle charging station. The control characteristics and the ripple coefficients for adjusting the current value of one of the units are compared. The research has shown the benefits of a five-unit topology. It has a lower ripple coefficient and a higher frequency compared to the six-unit topology. This makes it possible to smooth the current with smaller inductors. The ripple coefficients for the two topologies are nearly the same at a regulated source amplitude of 30% relative to the nominal. For this reason, it is recommended that the output power be adjusted by adjusting the amplitudes of several sources in a smaller range. As a disadvantage of the five-unit topology, the need to be implemented as one five-phase block or five single-phase blocks can be indicated. In the six-phase variant, it can also be realized by modules of two three-phase units.

When using a six-phase topology (Topology IV), even at a current value of one unit 0 (dropped or disconnected), the battery current is never equal to 0, as would be the case for three-phase topology [5].

Further recommendations can be made after spectral analysis of the output currents in the modes studied above. The models of multiphase power sources developed and shown above allow this analysis to be simulated.

The results presented herein provide information on the effect of asymmetry in five- and six-link topologies and can be used to define the requirements for converter control systems for electric vehicle charging stations or energy storage systems.

The most recent interpretation of the SPI term is System, Software, Services, Safety, and Security Process and Product Improvement, Innovation, and Infrastructure [10]. EuroAsiaSPI has developed as experience and knowledge exchange platform for Europe and worldwide, where Software Process Improvement (SPI) practices and knowledge can be gathered and shared. The connected SPI Manifesto [11] defines the required values and principles for most efficient SPI research and industry projects and is a Strategy Paper for further development on the base of those principles. One main goal is to support changes by innovation and include all stakeholders. There are two important values, which support the whole development process namely "SPI is inherently linked with change" means that change is part of the normal development process which leads to new business opportunities [12]. A current major change in the automotive industry affects the mobility concepts and the charging infrastructure. The presented paper gives new insights into the implementation of multiphase converters for the e-mobility charging station. "Create a learning organization" is an important principle for the exchange of best practices on the researched topics [13].

Acknowledgment. The research was carried out within the framework of the project "Optimal design and management of electrical energy storage systems", КП-06-Н37/25/18.12.2019, Bulgarian National Scientific Fund.

References

1. Soeiro, T., Bauer, P.: Research in E-mobility Infrastructures. www.tudelft.nl. Accessed 10 Mar 2020
2. Stecca, L.M., Ramirez, E., Batista Soeiro, T., Bauer, P., Palensky, P.: A comprehensive review of the integration of battery energy storage systems into distribution networks. IEEE Open J. Ind. Electron. Soc. **1**, 46–65 (2020)
3. Colak, K., Asa, E., Czarkowski, D.: A comparison analysis of CLL and LLC resonant converter for multi-phase applications. In: 2015 IEEE Transportation Electrification Conference and Expo (ITEC), Dearborn, MI (2015)
4. Ahn, S., Jang, S., Ryoo, H.: High-efficiency bidirectional three-phase LCC resonant converter with a wide load range. IEEE Trans. Power Electron. **34**, 97–105 (2019)
5. Arnaudov, D., Vuchev, S.: Modelling and investigation of multi-phase rectifiers supplied by resonant converters. In: 41st International Spring Seminar on Electronics Technology – ISSE 2018, Zlatibor, Serbia, 16–20 May 2018
6. Arnaudov, D.: Influence of asymmetry in multiphase resonant converters for energy storage systems. In: XXVI International Scientific Conference Electronics – ET 2017, Sozopol, Bulgaria, 13–15 September 2017
7. Vuchev, A.S., Arnaudov, D.D.: Design considerations for stages of modular topology for fast charging of electric vehicles. In: 2019 X National Conference with International Participation (ELECTRONICA), Sofia, Bulgaria (2019)
8. Madzharov, N., Petkov, V.: Innovative solution of static and dynamic contactless charging station for electrical vehicles and Ener. In: PCIM Europe 2016; International Exhibition and Conference for Power Electronics, Intelligent Motion, Renewable Energy and Energy Management, Nurnberg, Germany (2016)

9. Christen, D., Tschannen, S., Biela, J.: Highly efficient and compact DC-DC converter for ultra-fast charging of electric vehicles. In: 2012 15th International Power Electronics and Motion Control Conference (EPE/PEMC), Novi Sad (2012)
10. Manifesto. http://2018.eurospi.net/index.php/manifesto. Accessed 2 Apr 2020
11. The SPI Manifesto, Alcala, EuroSPI 2009. http://2019.eurospi.net/images/eurospi/spi_manifesto.pdf
12. Messnarz, R., Ekert, D., Grunert, F., Blume, A.: Cross-cutting approach to integrate functional and material design in a system architectural design – example of an electric powertrain. In: Walker, A., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 322–338. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_25
13. Korsaa, M., et al.: The people aspects in modern process improvement management approaches. J. Softw. Evol. Process **25**(4), 381–391 (2013)

A Seamless Self-configuring EtherCAT Master Redundancy Protocol

Jürgen Dobaj(✉), Matthias Seidl, Thomas Krug, Michael Krisper, and Georg Macher

Graz University of Technology, 8010 Graz, Austria
{juergen.dobaj,t.krug,matthias.seidl,michael.krisper, georg.macher}@tugraz.at

Abstract. The ever-increasing demand for system autonomy in all kinds of applications is accompanied by the need for increased fault-tolerance to build highly-reliable systems. EtherCAT technology is one of the most widely used industrial Ethernet solutions for process control and automation. While EtherCAT provides some fault-tolerance mechanisms, it does not protect an industrial control system against the malfunction of its master node. In this paper, we present a software-based master redundancy protocol for EtherCAT, which enables an EtherCAT network to seamlessly recover from master node failures within the same communication cycle. Moreover, the protocol integrates seamlessly into new and existing applications without the need to re-implement functionality and without running into compatibility problems. Experimental results showed that communication cycles in the sub-millisecond range are supported. Hence, the presented master redundancy protocol provides an easy to use, scalable, and cost-effective solution to increase the reliability of EtherCAT networks.

Keywords: Industrial ethernet · Reliability · Fault-tolerance · Real-time · EtherCAT · Automation · Process control

1 Introduction

The ever-increasing demand for system autonomy in all kinds of applications is accompanied by the need for increased fault-tolerance (FT) to build highly-reliable real-time systems. To fulfill these demanding requirements, automation system vendors no longer favor Fieldbus technologies; instead, they increasingly rely on industrial Ethernet solutions [14,16]. A market share analysis conducted by HMS[1] confirms this trend, showing that the annual share of Fieldbus installations is declining by −5%, while industrial Ethernet solutions show a steady annual growth rate of 20%, reaching a global market share of 59% in 2019.

[1] https://www.hms-networks.com/news-and-insights/news-from-hms/2019/05/07/industrial-network-market-shares-2019-according-to-hms.

Supported by the H2020 project *TEACHING* (n. 871385) - www.teaching-h2020.eu.

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 381–395, 2020.
https://doi.org/10.1007/978-3-030-56441-4_28

While most industrial Ethernet technologies support safety, they are not designed to provide high-reliability [1]. However, the increasing system autonomy and the trend towards smart and flexible manufacturing demands cost-efficient solutions that foster the reliable automation of the shop floor to achieve low down-times. Ethernet Control Automation Technologies (EtherCAT)[2] is the third most deployed industrial Ethernet technology. It implements a master/slave architecture and offers basic FT mechanisms including seamless path redundancy, as well as fail-over procedures to provide time and clock redundancy [1].

However, EtherCAT does not provide mechanisms that protect an industrial control network against a malfunction of its master node. Since the master node depicts a (SPoF) within EtherCAT networks, EtherCAT is not suited to provide high-reliability.

In this paper, we address this issue and present a software-based (MRP) for EtherCAT, which enables an EtherCAT network to seamlessly recover from master node failures.

2 EtherCAT Technology

Ethernet Control Automation Technologies (EtherCAT) is a data link-layer protocol that provides highly accurate clock-synchronization and high-speed cyclic communication in a sub-millisecond range. These features make EtherCAT optimal suited for high-end industrial applications such as motion control or power electronics [5,8,9,13]. EtherCAT is based on a single-master/multi-slave network architecture, where the master node (i.e., the control node) manages all slave nodes in the network. The master can be implemented as software on conventional embedded devices or industrial PCs, whereas the slaves use a dedicated integrated circuit (IC) to take part in the communication, the so-called EtherCAT processing unit (EPU). While EtherCAT supports various physical topologies, including line, tree, star, and ring, the EPU is used to always enforce a **logical circular topology**, which allows EtherCAT to seamlessly tolerate slave network faults, if the network is configured as a physical ring, like shown in Fig. 1. To enforce the logical circular redundancy, the EPU forwards received frames from one of its ports to the other port and uses its "loop-back" feature to tolerate faults.

In the following, we describe the EtherCAT protocol execution, which is depicted in the diagram shown in Fig. 2. The diagram illustrates the protocol steps by visualizing the network traversal of frames in space and time for specific communication cycles. For the sake of simplicity, we decided to start the explanation of a communication cycle in the *SEND* state. However, it should be mentioned that the *SLEEP* state is used to synchronize the cyclic protocol execution, which follows the pattern *SLEEP*, *RECEIVE*, *PROCESS*, *SEND*, *SLEEP*.

[2] www.ethercat.org.

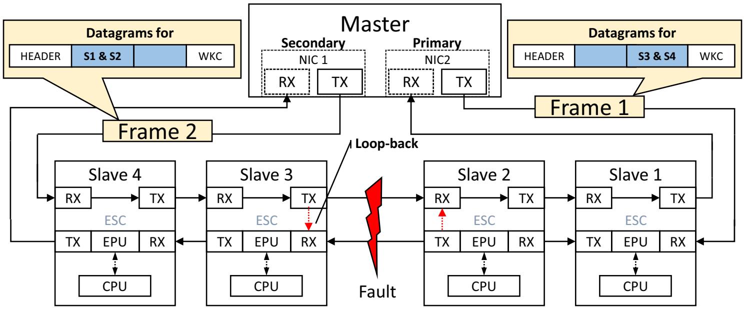


Fig. 1. EtherCAT communication redundancy.

The protocol execution in a **fault-free network** is illustrated in Fig. 2 starting at time point t_0. In a physical ring topology, as shown in Fig. 1, the first and the last slave are connected to the master node, and the master simultaneously sends out $frame_1$ and $frame_2$ over the primary and the secondary port in both directions. $Frame_1$ passes the network in *processing direction* and $frame_2$ passes the network in *forwarding direction*. In the fault-free case, $frame_2$ is forward by all slaves remaining unmodified.

$Frame_1$, on the other hand, passes all EPUs in processing direction, meaning that an EPU does not only forward the frame, but also modifies it by performing specific actions (i.e., read and/or write) on-the-fly. These actions are specified within the same frame in the header-part of an EtherCAT datagram. An EtherCAT frame can consist of multiple datagrams, where each datagram is addressed to a specific slave.

The protocol execution in a **faulty network** is illustrated in Fig. 2 starting at time point t_8. If a **network fault** occurs, as exemplified in Fig. 1, the physical ring topology degrades to two separate lines. The EPUs of the slaves react to this fault by activating an internal switch, which sends the received frames back to the master node via the same port. The master, on the other hand, recognizes such a fault by checking the (WKC) of received frames. Since the WKC is modified by slaves whenever they are addressed, an unexpected working counter value indicates a fault in the slave network.

Since $frame_1$ did not pass all slaves, the master re-transmits $frame_1$ over the secondary port to ensure that the frame passes all slaves in the correct order. This feature is called “loop-back”. It enforces the **logical circular topology** to ensure that all frames traverse the network in deterministic order. This also guarantees that the slaves perform all read/write actions in a deterministic order. Since the “loop-back” feature and the sensing of cable failures are implemented in hardware within the EPU, it is possible to seamlessly tolerate transient or permanent faults in the redundant network configuration.

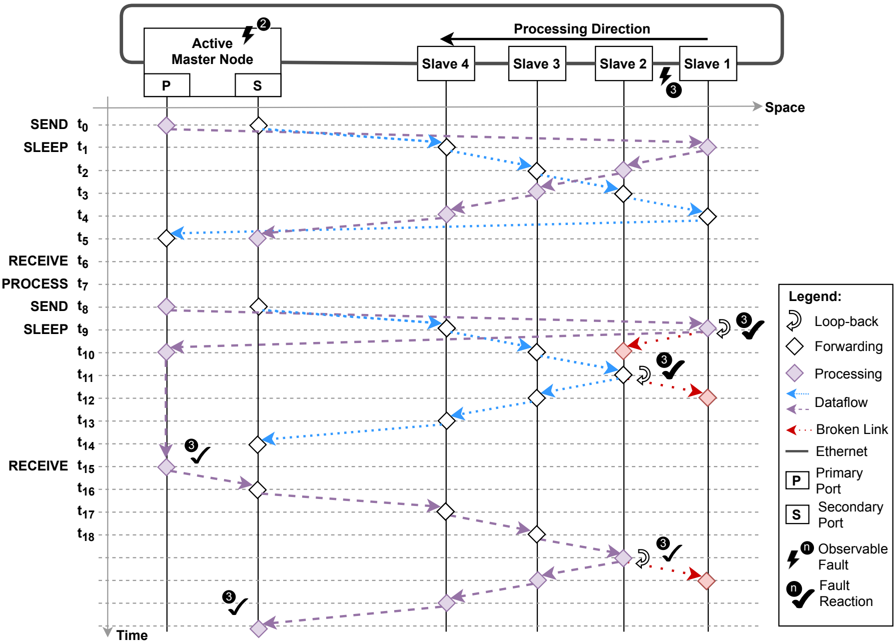


Fig. 2. Space-time diagram: EtherCAT protocol

3 Enhancing EtherCAT's Fault-Tolerance

To ensure (RT) capabilities EtherCAT relies on clock synchronization. The clock synchronization is also periodically triggered by the master. Hence, the master is key to the correct system operation, since the master manages both the communication and the clock-synchronization. Nevertheless, **EtherCAT provides no mechanisms to tolerate faults of its master node**. Furthermore, the **frame used for clock synchronization constitute a SPoF**, which is especially critical, because there is neither a time redundancy mechanisms present to tolerate frame losses, nor a mechanism to synchronize the line networks in case the physical ring topology breaks.

In [7], Prytz and Skaalvik present a solution for the described time synchronization problem by integrating two additional slaves into a master, whereas one slave is integrated in reverse order. This enables the master to provide synchronized clock sources for the distinct line networks.

The patent [12] describes a hardware-based method to implement redundant master nodes in an EtherCAT network. To that aim, specialized hardware is used to allow a secondary master to participate in bus communication as a slave node passively. In the case the secondary master detects a fault, the secondary master becomes active and can seamlessly take over the bus control.

The fault-treatment in EtherCAT networks is managed in software by the master node, which might be a limiting factor for high-performance applications.

The approach presented in [4], moves the fault-treatment code into a field programmable gate array (FPGA) located near to the network ports, accelerating the fault-reaction speed of master nodes.

To the best of the authors' knowledge, the presented paper is the first publication that offers a MRP for EtherCAT that can detect and seamlessly compensate master node faults in the value and time domain. Furthermore, the MRP is compatible with the solution presented in [7], enabling to build a system not only offering redundant master nodes, but also redundant clock sources.

Since the MRP is software-based, the authors claim that the protocol can be implemented on a FPGA (similar to [4]). This would decrease the jitter and latency introduced by the operating system scheduler and the memory operations, making the proposed MRP suitable for high-performance control tasks supporting cycle times in the sub-millisecond range. The evaluation of the MRP on an FPGA and its integration into the redundant clock architecture is subject to future work.

4 EtherCAT Master Redundancy Design Rationale

The concept that underlies the software-based MRP presented in this paper builds upon the fundamental communication principles of EtherCAT, which we briefly recall in the following. EtherCAT is a master/slave network implementing a hop-by-hop communication protocol. In the context of industrial RT control systems, this means the master node sends out EtherCAT frames in strict periodic time intervals. Each frame circulates through the network and can be observed by all connected slaves, which allows addressing multiple slaves by sending a single frame. Whenever a slave is explicitly addressed, the slave executes an action specified by the frame content and updates the WKC (and the data) of the frame "on-the-fly".

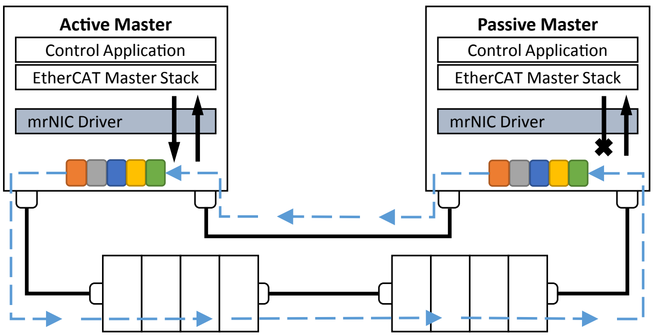


Fig. 3. EtherCAT master redundancy protocol concept.

While this "*summation frame*" technique is intended to increase the bandwidth efficiency of the EtherCAT protocol, we explicitly exploit this technique to implement the FT mechanisms of our software-based MRP. In essence, every device connected to the bus can transparently sniff the data that is transferred between the master and the slave devices, making it possible for every device to become aware of both the master and slave activities in the network.

Figure 3 sketches how we utilize this property to introduce one (or more) passive master nodes, which act as deterministic replicas of the active master. During regular system operation, a passive master node only observes the network activity without issuing its own frames to the network, relinquishing the network control to the active master node. If the passive master encounters a fault of the active master, the passive master starts its fault-management procedure to seamlessly take over the network control, to maintain the operational state of the system. To do so, the passive master first deactivates or disconnects the active master from the network to become the active master by sending its control commands to the slaves within the same communication cycle.

4.1 Failure Mode Coverage and Fault-Compensation

Before we continue with the description of the protocol details, we briefly introduce some FT concepts. A system can fail in various ways, which is categorized as *failure modes*. A failure mode is defined through the effects, as perceived by the system user. A *fault-detection mechanisms* is used to observe the system behavior to detect such effects. Hence, the fault-detection mechanism determines which failure modes can be covered. Once a fault is detected, an appropriate fault-compensation procedure can be triggered to obtain a fault-tolerant system [2,3,6].

A system composed of a set of distinct components is said to be *t fault-tolerant* if the system provides a correct service given that no more than t components become faulty during some *interval of interest*. The redundant EtherCAT system, shown in Fig. 1, for example, provides *0 fault-tolerance*; although it can tolerate single faults in the slave network, it does not provide means to tolerate master node faults.

By introducing additional master nodes into the EtherCAT network, as shown in Fig. 3, the FT against master node faults can be improved. Although it is possible to add multiple master nodes to the network, at maximum, a *1 fault-tolerant* system configuration can be obtained since, at maximum two master nodes can be connected to the slave network simultaneously. Hence, only one of those masters is allowed to fail, before the connection to the slave network is lost.

The fault-detection and fault-compensation mechanisms of the MRP are based on the concept of *deterministic state machine replication* [10], more precisely, on passive replication as described in Sect. 4. Due to the inherent non-determinism of (CPSs), it is impossible to obtain fully deterministic replicas in such systems. However, using techniques such as time-synchronization, as well as, deterministic algorithms, operating systems, hardware components, and communication protocols, it is possible to bound the level of non-determinism, allowing to build CPSs that offer relaxed replica determinism.

EtherCAT relies on such techniques to build so-called real-time synchronous systems. Real-time synchrony offers relaxed replica determinism, where an apriori known time interval bounds the inconsistency between replicas. Real-time synchrony is based on a discrete time grid, which is realized in distributed systems through clock synchronization means, such as provided by the EtherCAT protocol, the (NTP), or the (PTP).

4.2 Fault-Detection Based on Replica Determinism

The fault-detection algorithm of the MRP is based on two different approaches: a **content-based** and a **time-based approach**. Both approaches rely on the real-time synchrony of the EtherCAT protocol, which guarantees the deterministic ordered delivery of messages (frames) within a certain time-bound. Exactly this deterministic behavior enables the efficient implementation of the software-based redundancy protocol, because only little communication overhead is introduced to implement *replica coordination* between the active and the passive master nodes.

Replica coordination requires that all replicas start in the same initial state, execute the same deterministic algorithm, and all replicas receive and process the same sequence of requests. This can be summarized by the statement that *agreement* and *order* must be achieved between replicas, which means that every non-faulty replica receives every request (i.e., agreement), and every non-faulty replica processes the request it receives in the same relative order (i.e., order) [10].

While these properties typically require complex and resource-intensive consensus or broadcast protocols, the EtherCAT communication protocol and network architecture provide these properties by default through restricting its communication topology to a **logical circular topology**. This logical circular topology guarantees that every network participant receives each frame in a deterministic order, as described in Sect. 2.

More specifically, the **content-based fault-detection** approach of the MRP is based on the deterministic behavior of the control algorithm executed by the active and the passive master nodes. Since both nodes execute the same deterministic algorithm, the passive master can compare its self-generated EtherCAT frames, to the frames received from the active master. A difference in the frames indicates a failure in the communication link or some master node. Without the exchange of additional messages, which are not part of the EtherCAT protocol specification, it is impossible to determine which component actually failed (goal of the authors was to not violate the EtherCAT protocol specification). Under such fault conditions, the correct system operation can not be guaranteed, which typically results in triggering a fail-safe procedure that shuts down the system in a safe and controlled manner.

The **time-based fault-detection** approach, on the other hand, allows the passive master to react to faults of the active master by seamlessly taking over the network control. This is possible by learning the timing-behavior of the cyclic communication pattern controlled by the active master. Once the pattern is learned, the passive master continues to observe the timing behavior of the communication. Whenever the passive master does not receive an expected frame in time, the passive master triggers its fault-compensation protocol for recovery.

To that aim, the passive master first executes a call-back function. Usually, application developers use such call-back functions to trigger alerts and to disconnect faulty devices from the network to ensure fault-containment. After executing the call-back function, the passive master becomes active and sends the EtherCAT frames already prepared for the present communication cycle. This allows the passive master to take over the system control within the same exection cycle, which ensures to preserve the system's operational state.

4.3 Self-configuring Fault-Detection

Since the protocol is software-based, the jitter introduced by the operating system scheduler and the used software and hardware stack varies between platforms. Furthermore, communication patterns also vary for different applications. In order to use the time-based fault mechanisms, it would, therefore, be necessary to determine the specific platform and communication timings of each use case individually. Since this would introduce significant engineering overhead, it should be prevented.

Thus, the MRP learns the timing behavior of the specific application and platform by observing the strictly periodic receive intervals in the cyclic communication. The passive master measures the receive timings on all ports and calculates the average expected receive timing of frames within a cycle, using a moving window algorithm. The watchdog timer for triggering the fault-compensation mechanism is set to a fixed percental offset based on the observed timing behavior. This offset value was determined by executing several test runs. While this approach is feasible for the operation of platforms, which do not frequently change, further investigation is required to prevent the generation of false positives. For example, if the system load or the application behavior changes over time.

5 EtherCAT Master Redundancy Protocol

In the following, we describe the protocol execution for a fault-free network and three fault scenarios, as well as the protocol's self-configuration mechanisms in more detail. The fault-free scenario and the three fault scenarios are illustrated in the space-time diagram in Fig. 4 and are representative of all faults that can be detected and compensated by the protocol. For the presented scenarios, the following assumption (i.e., pre-conditions) are made:

- The active master does not know the passive master or the protocol executed by the passive master node.
- The master nodes are synchronized using a distributed clock synchronization mechanisms.
- The master nodes execute the same deterministic control software.

– The master nodes start their protocol execution in the same initial state at the same point in time, meaning that the nodes build a real-time synchronous communication system.
– Since the passive master must learn the cyclic communication pattern before faults can be detected, it is required that the network is non-faulty at the beginning of the protocol execution.

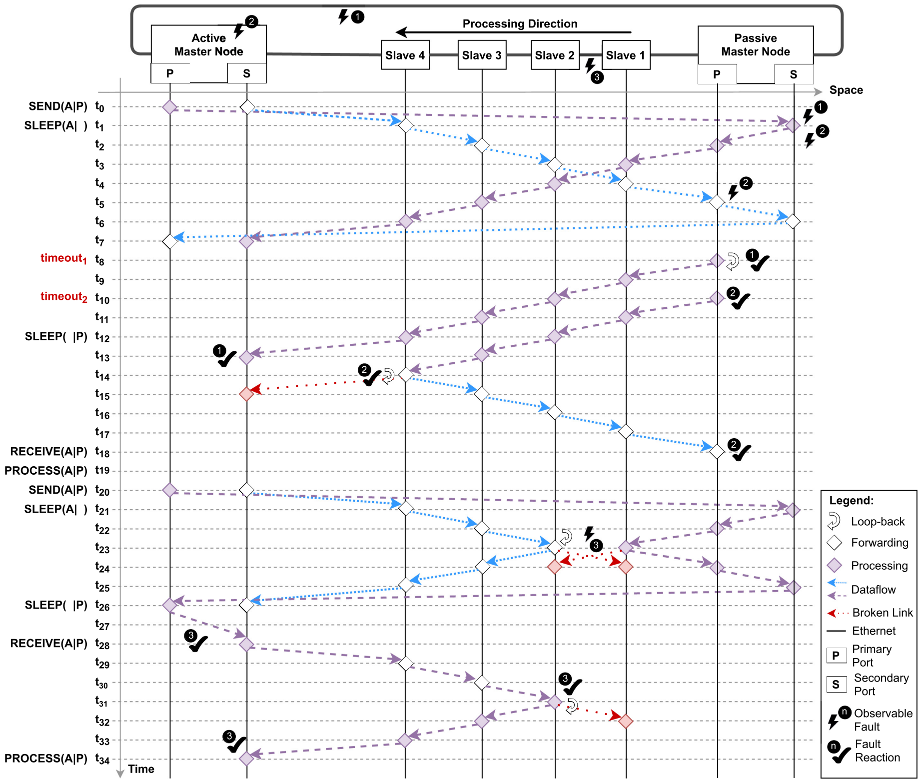


Fig. 4. Space-time diagram: EtherCAT master redundancy protocol (MRP)

Scenario 1: Fault-free Operation. The fault-free protocol execution is shown in Fig. 4 starting at time point t_0. Both master nodes enter their *SEND* state and generate the frames to be transmitted to the slave nodes by writing the output data into the corresponding *transmit-buffers*. However, only the active master sends its frames into the network. If the communication paths between the master nodes are non-faulty, **all frames sent by the active master are received by the passive master without modifications**. The frames transmitted via the primary port remain unchanged, since the active master's primary port is directly connected to the passive master's secondary port. The frames send via the active master's secondary port remain unchanged, because they traverse the slaves in *forwarding direction*.

When receiving the frames at time point t_1 and t_5, the passive master stores a *timestamp for each receive-event*, *compares* the content of the secondary port's *receive-buffer* to the content of the primary port's *transmit-buffer*, which was written by the passive master in its *SEND* state, but not transmitted to the network. Since the received frames are not modified when traversing the network, the content of the receive- and transmit-buffers must be equal, if the communication path is fault-free (the fault scenarios are discussed in the following).

If the buffers are equal, the passive master can immediately send the content of the primary and secondary port *transmit-buffers* to the network. The costly memory-copy operation from the receive- to the transmit-buffers can be omitted, increasing the efficiency of the software implemented MRP.

Next, the frames sent via the passive master's primary port traverse the slaves in *processing direction*, meaning that the frame content is updated on-the-fly according to the actions specified within the datagrams contained in the frame. The frames sent via the secondary port of the passive master traverse the network and remain unchanged.

Comparing the fault-free sequence of Fig. 2 and the described sequence reveals that the active master receives the same data (i.e., frames) later in time, at time point t_7 instead of t_5. This means, that the execution of the MRP introduces latency into the communication, but still remains transparent to the active master.

Scenario 2: Disconnected Master Nodes. Figure 4 illustrates a scenario where the connection between the master nodes is faulty. This scenario starts at time point t_0. The sequence of operations is similar to the sequence of *scenario 1*, yet the passive master does not receive the frames on its secondary port at time point t_1. However, the frames on the primary port are received correctly at time point t_5 and can thus be treated equally to *scenario 1*.

To compensate the described fault, the passive master already used the *timestamps created for each receive-event* to learn the cyclic communication pattern of the active master and to calculate the $timeout_1$, as described in Sect. 4.3. If the passive master does not receive the expected frame before the $timeout_1$ occurs, the passive master initiates its fault-compensation mechanisms to seamlessly recover from the fault (see also Sect. 4.1).

To do so, the passive master immediately sends the already prepared frames in its primary port's *transmit-buffer* to the slaves. These frames traverse the slaves in *processing direction* and are received by the active master on its secondary port.

Comparing the described sequence to the fault-free sequence of *scenario 1* reveals that the active master receives the same data later in time, at time point t_{13} instead of t_7. Although the active master receives the correct data, it is essential to adequately set the $timeout_1$, to on the one hand prevent false-positives and on the other hand to ensure that the frames are received by the active master in time, meaning before the active master leaves its *RECEIVE* state.

To sum up, the waiting phase to $timeout_1$ introduces additional latency to the communication. However, if the timeout is adequately set, the MRP can seamlessly compensate the described fault. Again, the protocol execution is still transparent to the active master.

Scenario 3: Master Node Fault. In the following, we describe two fault scenarios. In scenario 3.a, the passive master can compensate the fault, while in scenario 3.b the fault can only be detected.

In **scenario 3.a**, we assume that the active master has crashed and does not send messages to the network. This scenario is shown in Fig. 4 and starts at time point t_0. The sequence of operations is similar to those of *scenario 1*, yet the passive master does not receive the frames on its primary and secondary port at time point t_1 and t_5.

To detect this failure mode, the same technique can be applied as described in *scenario 2*. By introducing an additional $timeout_2$, the passive master can trigger its fault-compensation mechanisms for recovery. However, in this case, the passive master must assume that the primary master has a performance failure or has crashed. To ensure that the active master does not corrupt the communication cycles in case the master resumes operational, the passive master first deactivates the active master (see Sect. 4.1 for more details), degrading the network to a line topology.

Then the passive master takes over the network control and sends the already prepared frames to the slaves at time point t_{10}. These frames pass the slaves in *processing direction*. Since the active master is deactivated, the EtherCAT network operates no longer in redundant mode. This is detected by *slave 4*, which seamlessly activates its "loop-back" functionality and sends the received frames back to the passive master in *forwarding direction*.

Summing up, in the described sequence, the passive master seamlessly recovers by first identifying the correct fault-scenario (i.e., running into $timeout_1$ and $timeout_2$). Then the passive master deactivates the active master to ensure error containment. As a third and final step, the passive master takes over the network control by sending out the frames prepared for the specific cycle.

Again, it is essential to set the $timeout_2$ adequately, which prevents false positives, and ensures seamless recovery by re-transmitting the frames before the passive master leaves its *RECEIVE* state.

In **scenario 3.b**, we assume that the active master is still operational, but it sends out incorrect frames over one of its ports, and correct frames over the other port, or some frames too late or too early. This behavior describes a *byzantine or arbitrary failure*. The passive master can detect the faulty behavior of the primary master using the previously described time-based and value-based fault-detection mechanisms. However, the passive master can not recover from this failure mode because the passive master can not determine which component in the network has introduced the fault. In such a situation, it is common to trigger a fail-safe procedure.

Scenario 4: Slave Network Fault. In this scenario we describe a failure that occurs in the slave network, and degrades the network form a physical ring topology to two separate lines. Although this fault scenario is not related to master node faults, we think it is important to show that the presented MRP does not violate the FT mechanisms of EtherCAT.

The scenario of interest is shown in Fig. 4 and starts at time point t_{20}. The sequence of operations in this scenario starts similar to the one explained by *scenario 1*. However, in the present scenario $slave_1$ and $slave_2$ are disconnected due to a cable fault. Thus those slaves activate their "loop-back" feature and send the received frames back to the master nodes.

The passive master node receives the frames at time point t_{24} and recognizes that a slave network failure has occurred by checking the WKC of the received frame already modified by $slave_1$. Since the active master is responsible for reacting to slave network faults, the passive master forwards the received frame over its secondary port at time point t_{25} and enters its *SLEEP* cycle afterward.

It is essential to notice that the passive master is not allowed to immediately enter its *SLEEP* cycle after sending the frames at time point t_{22}, such as the active master is doing at time point t_{22}. Otherwise, the passive master would not be able to immediately forward the "loop-back" frame, which the active master consumes in its next *RECEIVE* cycle.

In the presented scenario, the active master receives both frames at time point t_{26} during its *SLEEP* cycle and continues to sleep until time point t_{28}. At time point t_{28} a new communication cycle starts and both master nodes enter their *RECEIVE* cycle. Like the passive master node before, the active master checks the WKC and realizes that a fault in the slave network occurred.

Similar to the faulty slave network scenario shown in Fig. 2, the active master takes the frame received on its primary port and forwards it through its secondary port at time point t_{28}. Here it is essential to notice that this frame has already been modified by $slave_1$. Therefore, to preserve the request order $slave_2$ must be next to modify the frame, followed by $slave_3$ and $slave_4$. Since the frame transmitted at time point t_{28} first traverses the network in *forwarding direction*, it is ensured that $slave_2$ is the next, to modify the frame. $Slave_2$ is also the one that performs the "loop-back" of the frame, which ensures that the requested order is preserved. At time point t_{34} the active master finally receives the correct frame and can enter its *PROCESS* cycle.

If we compare the presented scenario to the fault-scenario shown in Fig. 2, it can be seen that the MRP does not introduce any additional latency, since the passive master forwards its frame to the active master, before entering its *SLEEP* cycle. Additionally, it is shown that the passive master protocol does not violate the EtherCAT protocol specification.

6 Integration

As a proof of concept, we integrate the MRP into the Simple Open EtherCAT Master (SOEM)[3] software stack. We make sure that the integration is seamless, i.e., that no changes to the SOEM core functionality or application programming interface (API) must be made. This allows developers to extend and enhance existing EtherCAT-based applications to support redundant masters without the need to re-implement functionality and without running into compatibility problems.

To ensure seamless integration, we employ the architecture sketched in Fig. 3. We encapsulated the entire MRP functionality within the (NIC) driver of the SOEM, named mrNIC driver in the architecture. The mrNIC driver extends the API of the SOEM NIC driver offering additional interfaces to pass parameters to the MRP and to register call-back functions. Application developers can use this call-back functions to enhance the MRP fault-compensation algorithm with application specific actions, such as alerting, error containment routines, and fail-safe procedures.

The mrNIC driver operates in the user space and uses a raw socket interface to communicate with the (OS) network library, providing the following benefits:

– The mrNIC driver is easily portable across OSs and platforms.
– Developers can integrate the MRP by (re-)compiling their application code against the mrNIC library.
– The mrNIC driver can be ported to run on an FPGA, which eliminates the latency introduced by the OS network stack and the OS scheduler.

7 Evaluation

We evaluate the MRP experimentally and showcase the MRP functionality by setting up an EtherCAT control network comprised of two master nodes and several slave nodes. Then we analyse the robustness and timing behavior of the MRP based on the four scenarios presented in Sect. 5. In the evaluation we determine the communication *jitter*, which describes the fluctuations of a specific cycle time due to uncertainty in the communication path [15]. This is especially interesting, since the MRP protocol increases the communication uncertainty due to its software-based implementation and by adding an additional hop (i.e., the passive master) to the network.

In the experimental setup the master nodes are equipped with an industrial-grade x86 multi-core processor (i.e., Intel®Atom™x5-E3940, 2M Cache, up to 2.00 GHz) and 4GB system memory. We only use a limited number of slave devices, as it does not influence the communication jitter introduced by the MRP. Hence, the analysis can be extended to network configurations with an arbitrary number of slaves. Our experimental setup is similar to the one used in [11] to empirically evaluate the protocol performance of EtherCAT.

[3] https://github.com/OpenEtherCATsociety/SOEM.

The evaluation results are provided in Table 1, showing the points in time when a master node enters a new state relative to the cycle start in the case of a fault-free network. The results reveal, as expected, that a single execution cycle on the passive master is longer due to the introduction of the additional waiting time required for the time-based fault-detection mechanisms. An execution cycle on the active master requires at maximum 459 µs, while the execution of the MRP on the passive master requires at maximum 714 µs. The table shows that although the MRP introduces additional latency the worst case execution time of a cycle was 714 µs. Hence, the software-based MRP supports communication cycles in the sub-millisecond range for the experimental setup.

Table 1. Time points of the master nodes entering a new state within one execution cycle.

	Active master				Passive master			
	RECEIVE	PROCESS	SEND	SLEEP	RECEIVE	PROCESS	SEND	SLEEP
MIN	15 µs	76 µs	77 µs	97 µs	19 µs	118 µs	119 µs	169 µs
AVG	180 µs	254 µs	254 µs	279 µs	185 µs	348 µs	349 µs	419 µs
MAX	339 µs	412 µs	413 µs	459 µs	339 µs	500 µs	501 µs	714 µs

8 Conclusion

In this paper, we have presented a master redundancy protocol (MRP) for EtherCAT, which is capable of detecting and seamlessly compensating master node faults in the value and time domain. We illustrated the protocol execution and explained that it can be integrated into existing EtherCAT-based applications to add support for redundant master nodes without the need to re-implement functionality and without running into compatibility problems. We showcased that the MRP supports communication cycles in the sub-millisecond range and that the protocol increases the reliability of EtherCAT networks.

Future work includes improving the self-configuration algorithm and evaluating the protocol robustness under different load and fault scenarios.

References

1. Alvarez Vadillo, I., Ballesteros, A., Barranco, M., Gessner, D., Djerasevic, S., Proenza, J.: Fault tolerance in highly reliable ethernet-based industrial systems. Proc. IEEE **107**(6), 977–1010 (2019). https://doi.org/10.1109/JPROC.2019.2914589
2. Avizienis, A., Laprie, J.C., Randell, B., Landwehr, C.: Basic concepts and taxonomy of dependable and secure computing. IEEE Trans. Dependable Secure Comput. **1**(1), 11–33 (2004). https://doi.org/10.1109/TDSC.2004.2
3. Goodloe, A., Pike, L.: Monitoring distributed real-time system - a survey and future directions (2010)

4. Maruyama, T., Yamada, T.: Spatial-temporal communication redundancy for high performance ethercat master. In: Proceedings of the IEEE, pp. 1–6 (2018). https://doi.org/10.1109/ETFA.2017.8247720
5. Orfanus, D., Indergaard, R., Prytz, G., Wien, T.: Ethercat-based platform for distributed control in high-performance industrial applications. In: 2013 IEEE 18th Conference on Emerging Technologies & Factory Automation (ETFA), pp. 1–8. IEEE, 10 September 2013–13 September 2013. https://doi.org/10.1109/ETFA.2013.6647972
6. Poledna, S.: Fault-Tolerant Real-time Systems: The Problem of Replica Determinism. The Kluwer International Series in Engineering and Computer Science, vol. 345. Real-Time Systems, Kluwer Academic, Boston (1996)
7. Prytz, G., Skaalvik, J.: Redundant and synchronized ethercat network. In: International Symposium on Industrial Embedded System (SIES), pp. 201–204. IEEE, 7 July 2010–9 July 2010. https://doi.org/10.1109/SIES.2010.5551386
8. Prytz, G. (ed.): A Performance Analysis of EtherCAT and PROFINET RT. IEEE (2008). https://doi.org/10.1109/ETFA.2017.8247720
9. Rostan, M., Stubbs, J.E., Dzilno, D.: Ethercat enabled advanced control architecture. In: 2010 IEEE/SEMI Advanced Semiconductor Manufacturing Conference (ASMC), pp. 39–44. IEEE, 11 July 2010– 13 July 2010. https://doi.org/10.1109/ASMC.2010.5551414
10. Schneider, F.B.: Implementing fault-tolerant services using the state machine approach: a tutorial. ACM Comput. Surv. **22**(4), 299–319 (1990). https://doi.org/10.1145/98163.98167
11. Seno, L., Vitturi, S., Zunino, C.: Real time ethernet networks evaluation using performance indicators. In: 2009 IEEE Conference on Emerging Technologies & Factory Automation, pp. 1–8. IEEE, 22 September 2009–25 September 2009. https://doi.org/10.1109/ETFA.2009.5347135
12. Spiegel, G., Vysotski, V.: Bus participant device and method for operating a bus subscriber device (2018)
13. Toh, C.L., Norum, L.E.: A performance analysis of three potential control network for monitoring and control in power electronics converter. In: 2012 IEEE International Conference on Industrial Technology, pp. 224–229. IEEE, 19 March 2012–21 March 2012. https://doi.org/10.1109/ICIT.2012.6209942
14. Vitturi, S., Zunino, C., Sauter, T.: Industrial communication systems and their future challenges: next-generation ethernet, IIoT, and 5G. Proc. IEEE **107**(6), 944–961 (2019). https://doi.org/10.1109/JPROC.2019.2913443
15. Willig, A., Wolisz, A.: Ring stability of the profibus token-passing protocol over error-prone links. IEEE Trans. Industr. Electron. **48**(5), 1025–1033 (2001). https://doi.org/10.1109/41.954567
16. Wollschlaeger, M., Sauter, T., Jasperneite, J.: The future of industrial communication: automation networks in the era of the internet of things and industry 4.0. IEEE Ind. Electron. Mag. **11**(1), 17–27 (2017). https://doi.org/10.1109/MIE.2017.2649104

A Study of Electric Powertrain Engineering - Its Requirements and Stakeholders Perspectives

Georg Macher[1(✉)], Eugen Brenner[1], Richard Messnarz[2], Damjan Ekert[2], Jakub Stolfa[3], Svatopluk Stolfa[3], Joerg Niemann[4], Claudia Fussenecker[4], Marius Schoening[4], Alexander Paul[4], Nikolay Pavlov[5], Boyko Gigov[5], and Maria Stefanova Pavlova[5]

[1] Graz University of Technology, Graz, Austria
{georg.macher,brenner}@tugraz.at
[2] I.S.C.N GmbH, Graz, Austria
{rmess,dekert}@iscn.org
[3] FEECS, VSB - Technical University of Ostrava, Ostrava, Czech Republic
{jakub.stolfa,svatopluk.stolfa}@vsb.cz
[4] Hochschule Düsseldorf, Düsseldorf, Germany
{joerg.niemann,claudia.fussenecker,marius.schoening, alexander.paul}@hs-duesseldorf.de
[5] Technical University of Sofia, Sofia, Bulgaria
{npavlov,bgigov}@tu-sofia.bg, stefanova-pavlova@citt-global.net

Abstract. The automotive domain is a key industrial sector for Europe with 13.8 million jobs, representing 6.1% of total EU employment, and generating a yearly trade balance of over 84.4 billion €. Simultaneously the automotive industry is and has for many years been the key driver of knowledge and innovation in Europe and is the second biggest R&D sector.

In addition to this the automotive market is of a highly competitive nature and is now being reorganized on a very broad front in a process which has been continuing for some years. One of the major Drivers-of-Change in the recent time period was related to powertrain electrification and novel energy-efficient propulsion systems. Due to massive societal challenges of CO2 reduction regulations the entire industry has given electrified mobility first priority.

This paper will concentrate on the evaluation of electric powertrain related skills required by stakeholder. The aim of this analysis is to establish a basis of skills required by those companies which are developing electric powertrain solutions of the future and develop an adequate training for future electric powertrain engineers. To that aim, the analysis presented in this paper integrates the different viewpoints of stakeholders (including EU perspective, industry representatives, researchers, and niche domain perspectives). Within this work, we present the different stakeholder view points and implications for the development of training concepts related to a smooth upskilling of engineers for the electronic powertrain engineering topics.

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 396–407, 2020.
https://doi.org/10.1007/978-3-030-56441-4_29

Keywords: Automotive · Skill training · Electric powertrain · Qualification profiles

1 Introduction

The European automotive industry generates 13.8 million (direct and indirect) jobs in the EU with a total trade balance of 84.4 billion € [8] and is the EU's top investor in R&D, with 57.4 billion € invested per year. The industry produces close to 20 million vehicles per year, accounting for 20% of worldwide production through its 15 manufacturers and associated supply chains. At the same time, Europe's automotive industry remains committed to address tomorrow's challenges.

Recent trends indicate an automotive revolution driven by the concept of connected and automated cars, but the European automotive industry is also massively challenged by CO2 reduction regulations (phased in from 2020 the EU fleet-wide average emission target for new cars will be 95 g CO2/km; corresponding a fuel consumption of around 4.1 l/100 km of petrol or 3.6 l/100 km of diesel) [9]. Thus, the entire industrial sector needs to evolve and adapt at a very fast pace to also stay ahead of global competition, while including all stakeholders and addressing societal needs. Within the past few years electrified mobility has been given first priority in the US, Japan, China, Korea and the EU.

It is not only the environmentally conscious societies of Europe that are eagerly looking forward to the integration of clean mobility into their urban lives. The global trend for sustainability is very clearly apparent and electro mobility is moving forward driven by a significant progress with attractive market oriented cars now providing enhanced drive dynamics, cruising capabilities and mileage overcoming the meagre car concepts of the past. The key to reducing the emissions of the mobility sector is the increase the electrification to profit from the vastly improved energy efficiency of electrified powertrains. This will facilitate more green cars and less emissions. Electric vehicles will definitely have more "positive" impact on our lives. It is perceived that they are going to start changing our way of life and our mobility.

Global automotive Drivers-of-Change such as climate change due to CO2 emission and related future innovative developments (electric powertrain, new infrastructures to support electric drive, e-mobility concepts) have led to an own agenda of the automotive manufacturers and actions under the new skills agenda for Europe until 2030. One of the analysed key job roles to support the future developments in the automotive industry is to train knowledge and skill set for an electric powertrain engineer.

This paper will concentrate on this evaluation of skill set for electric powertrain engineers. To that aim the analysis presented in this paper integrated the different viewpoints of stakeholders from (a) an EU policy maker and EU commission perspective, (b) representatives of the German and Austrian automotive industry, (c) the industry researcher and developer perspective (including two diverse views of development engineers and production & management engineers) and (d) the non-classical passenger vehicle perspective (including public transport and alternative highly fuel-efficient vehicles). Within the scope of this work, we present the different stakeholder view points and implications for the development of training concepts related to a smooth upskilling of engineers for the electronic powertrain engineering topics. The aim of this analysis is to establish a basis of skills required by those companies which are developing electric powertrain solutions of the future and develop an adequate training for future electric powertrain engineers.

This paper is organized as follows: Section 2 describes the general survey and questionnaire structure. Followed by Sect. 3 to Sect. 7, which present the overviews of the responses from the different stakeholder groups. In Sect. 8 the implications for the ECEPE training program deducted from the stakeholder needs is given. Finally, the relation to the SPI manifesto is provided and the work is concluded in Sect. 10.

2 Survey and Questionnaires

The aim of the industry survey is to obtain a deeper knowledge on the topic from the view of different industry stakeholders. For the success of the ECEPE project, it is necessary to collect and evaluate different viewpoints on the topic of electric powertrains in order to present tailored solutions.

The acquisition of the engineer's viewpoint was achieved either through personal meetings and discussions on the ePowertrain engineering topic and a brief presentation of the proposed structures of the course, followed by a discussion on the individual points of interest (for the development engineers view and niche view, Sect. 5 and Sect. 7), or the answers to the questions were collected online. This was done in an anonymous form to secure data privacy. The online questionnaires were tailored for the specific view points and also used different tools. As such, the questionnaire for the EU representatives[1] (Sect. 3), the automotive industry representatives[2] (Sect. 4), and the management and production view[3] (Sect. 6) used different approaches. The complete results created by the different tools are provided in the public report of ECEPE [6]. In general, the survey was divided into four parts, but then tailored for the specific stakeholder group.

The first part focuses on the participants and their expertise in their job area. Including the industry in which they are working and how many years of expertise they have.

The second question is considered the most relevant in the survey, providing firsthand information about which topics are relevant within the respective domain. This is aiming to achieve a better knowledge on which topics should play a more prominent role in designing the new course and how the course should be structured. The second question includes in total 27 elements of the electric powertrain development course material. The participants were asked to provide

[1] https://drives-survey.vsb.cz/.
[2] https://www.1ka.si/a/243083.
[3] http://umfrageonline.com.

their opinion based on a Likert scale from not important to very important for electric powertrain-development.

The third part of the survey focuses on different job roles and their linkage to the topics which are necessary to fulfil the development of electric powertrains. The third question aims to assign different job roles to the related topics.

The last section of the questionnaire aims to collect any missing topics which should be part of the final curriculum. In this context two open questions regarding cross-domain- knowledge and sustainability, Corporate Social Responsibility (CSR) and financial issues in the field of electric powertrain engineering were also added.

3 European Roadmaps for Electric Powertrain Engineering

To reach the EU climate neutrality goal by 2050, one of the main actions that must be taken is "rolling out cleaner, cheaper and healthier forms of private and public transport". This also includes and also greatly influences the move from the current widespread form of power transition (e.g. combustion engine etc.) to CO neutral methods for powering the private and public transport vehicles [10].

As stated in [11], 3 main challenges have been identified for the industry: (a) a globally competitive and world-leading industry, (b) an industry that paves the way to climate-neutrality, and (c) an industry shaping Europe's digital future.

Therefore, Europe must react and prepare a sustainable approach to overcome these challenges. E-powertrain is also part of it for some time [7].

All of that is also visible in EU strategic Sector Skill Alliance project DRIVES, which is the Blueprint for strategic cooperation on skills in the automotive sector. DRIVES is also mentioned in the 10 points plan how to tackle the Green Deal by ACEA [3]. Another strategic project in this area is ALBATTS project, which is the Blueprint project for strategic cooperation on skills in batteries for e-mobility sector. Industry Drivers-of-Change were and are discussed and pilot strategies defined in these two still running projects (2020/4).

"Drivers-of-Change" are factors which are key to transforming an industry. Specifically, a literature review of available automotive reports was undertaken in DRIVES project in order to create an overview of current Drivers-of-Change and their relevance (performed in the years 2018–2019). The analysis compares the outcomes of the European Automotive Skill Council report and GEAR 2030 report, with other available intelligence/reports related to the EU automotive sector in order to identify the main Drivers-of-Change within the European automotive sector. Through this approach five main 'macro' Drivers-of-Change have been identified, these being:

- new technologies and business models
- climate goals, environmental and health challenges
- societal changes and changes in the way that consumers access, purchase and use cars

– structural changes
– globalisation and the rise of new players

Each ‘macro’ Driver comprises several more specific Drivers-of-Change that were identified as relevant. Analysis also focused on the identification of emerging Drivers-of-Change. The literature review enabled the mapping of each initial macro Drivers-of-Change against wider research evidence, based on the number of times each Driver was cited in those reports included in the review. The literature review is described in [5]. The analysis indicates that most respondents recognize that each Driver of Change has either an urgent (by 2020) or a moderately urgent (by 2025) impact. Based on the results of the DRIVES demand survey, the importance of training and education provision mechanisms clearly shows the need for work-based learning as a delivery approach of training and education. Furthermore, the most important aspect mentioned is the possibility of tailoring the training for combination that reflects specific needs of each stakeholder. As a last point, specifically electric powertrain engineer was announced as one of the “most wanted” and urgent job roles of the DRIVES demand survey.

4 Automotive System Architecture and Engineering Viewpoint on Electric Powertrain Development

SoQrates working party [1] members, constituted from leading automotive suppliers (Robert BOSCH, ZF Friedrichshafen AG, Continental Automotive, Magna Powertrain, etc.), were surveyed to ascertain the viewpoint of German and Austrian automotive industry representatives. The full list of respondents is available via [16]. From this group of automotive suppliers a subset answered the questions at the general meeting at BOSCH Engineering GmbH in Abstatt, Germany, on 5.3.2020. At this meeting the study was presented and the partners were asked to answer specific questions related to the knowledge required for system control of an e-powertrain.

The survey revealed that electric powertrain engineer is seen as a job role in which the suppliers expect a retraining for (a) software engineers, (b) hardware engineers, and (c) electrical engineers. Furthermore, the respondees identified the following specific requirements as the focus for the training of electric powertrain engineers:

– standards/directives/regulations/High Voltage and fire considerations
– ePowertrain architectures (BEV, HEV, plugin HEV, fuel cell EV, in-wheel concepts)
– building blocks of ePowertrain (battery, eMotor, inverter, etc)
– safety view point (systems design & item thinking)
– signal flow/interfaces definitions (real-time thinking)
– hybrid control systems (eMotor/inverter/ECU) - SW Arch.
– eMotor control

5 Electric Powertrain Engineering from the Development Engineers Perspective

According to the work of [4], universities have the potential to play a vital role in the way Europe is equipped with the skills and competences necessary to succeed in a globalised, knowledge-based economy. University programs should be structured to enhance directly the employability of graduates and to offer broad support to the workforce more generally in order to overcome mismatches between graduate qualifications and the needs of the labour market. Only in very rare cases is a university degree a direct entry entitlement to a specific profession, so there has been no real pressure coming from professional bodies to define the requirements. Developments in the recent years have led to a different view of university education, reflecting the changing world with its growing mobility, and the necessity of transparency and comparability [15].

The contribution of this section is to present the viewpoint of automotive engineers and their employers on the usefulness and employment-related skill training provided by universities of technology. With this aim in view, the automotive partner network of Graz has been surveyed to reveal its needs and the skill gaps of graduates to be employed directly in the automotive cluster companies of the Graz region. The focus was set on the analysis of the viewpoint and perspective of the entire supply-chain; (a) large suppliers, (b) large engineering companies, (c) device suppliers, (d) SMEs, and (e) research institutions. The acquiring of the engineer's viewpoint was done via personal meetings and discussions on the ePowertrain engineering topic and a brief presentation of the proposed structures of the course, followed by a discussion on the individual points of interest.

Major common expectations of the different organisations were:

- Contributions to the breaking of silos in a manner which avoids unproductive and error prone processes
- Ability to communicate with engineering teams involved
- Engineers with the skill to communicate and document their core objective (safety, design, etc.)
- Fast intro to core developments and the core knowledge of the company. Doing so could save significant time, money and risks.

Except for these commonly expressed issues, specific differences related to the core work topic of the different companies could be observed. The group of large suppliers had two different views related to either production-oriented team focus or development-oriented engineering teams. The view of the latter was relatively aligned with the viewpoint of large engineering companies. As expected the viewpoint and focus of device suppliers and SME's also differed in a relatively drastic manner from that of the others in terms of the importance rating given to the engineering aspects. A more overarching view of the other viewpoints was provided by the industry research institutions. Industry research institutions act as bridge between university research and industrial development and focus on transferring technologies from low technology readiness level (TRL) to TRLs close to prototype and series development. Due to that focus the view of industry research centers is mainly aligned with the needs of the local automotive community and also tightly coupled to the research roadmaps and strategic research plans.

The five company groups provided feedback for the proposed curriculum structure. To that aim, the stakeholders have been asked to provide feedback for the following main questions:

- What skills should be part of the curriculum (general/specific)?
- What do you expect as basic knowledge in ePowertrain?
- How would you structure the course?

All the companies were asked to give feedback generally on the level of details that should be given to the teaching on the specific topic (basic or detailed level) and if the knowledge of this specific topic should be compulsory and mandatorily included or optional. Additionally, if they find additional topics to be provided and/or if a rearrangement of the topics/sub-topics would be required. In general the feedback on the proposed curriculum structure of the ECEPE training material was well appreciated and accepted; no major topic lacking has been identified and the overall structure fits the general context of automotive ePowertrain engineering. The questionnaire did not reveal any major gaps or shortcomings.

6 Management and Production Engineers View of Electric Powertrain Engineering

The different viewpoints collected for establishing the perspective of management and production engineers are from the automotive industry, automotive industry suppliers, the agricultural machinery industry and from the research environment. In order to provide a full view of the topic, all the participants are experts in the field of electric powertrains and work either in management or the production department. The impact on the agricultural machinery industry was also evaluated in this part of the survey. The participants from this branch had expertise based on at least 5 years in this working environment.

Furthermore the points of view represented by theses participants indicated that each of the selected subject areas is evaluated as being at least moderately important by them, while over 50% of these subject areas are ranked as being very important. Five out of the 27 subject areas were ranked as moderately important. These very important areas were the building blocks of electric powertrains, the e-Motor/generator, e-Motor control, electric power converter and the fuel cell. Those ranked simply as important were more broad subject areas, reaching from societal and environmental issues, through electromagnetic capabilities to the transmission.

As also indicated in this survey no average score lies under moderate importance. Every individual subject area has its own reason to be listed and to be considered for the curriculum of the ECEPE project. Only 5 out of the 27 subject could achieve a score above 4,0, and these cover the topics of batteries, safety regulations and the e-Powertrain architecture. The distribution of the participants implies that in every industry more or less knowledge in the same subject areas is needed, in the view of management and production engineers. From a management and production perspective the preliminary work in the project was successful due to a broad coverage of all the necessary subject areas.

7 Niche Market Views on Electric Powertrain Engineering

Electric powertrains are not only very important for passenger cars, they also have great prospects for development in the field of commercial vehicles. Electric vehicles have long been used in **public transport**, both rail and road. The new trends are mainly related to the pursuit of using autonomous power sources on board the vehicle to replace the overhead contact lines used so far. Operating vehicles without a catenary, for example, with a power source on board the vehicle by charging the power storage device periodically, can provide various benefits compared with the conventional systems [13]:

- Reductions in the level of visual intrusion
- Reduction in the cost of overhead infrastructure
- Reduction in power usage and CO2

As an example, according to the National Framework Program for Alternative Fuels and their Infrastructure, Bulgaria is expected to introduce 400 hydrogen busses by 2025. With the successful construction of a prototype, a Bulgarian production through retrofitting and modernization of vehicles would assist in the fulfillment of this commitment[4]. Figure 1 shows the most used innovative zero-emission bus technologies.

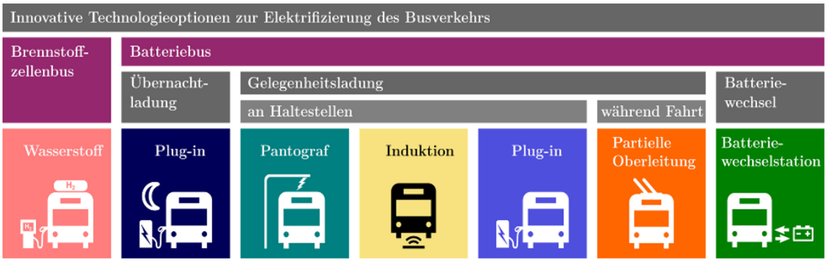


Fig. 1. Overview of the most used innovative zero-emission bus technologies [2].

Trucks are another large group of commercial vehicles. The very first eHighway was launched in Sweden in 2016. The concept here is the same - the trucks

[4] https://eplus.bas.bg/bg/results-110.

use pantographs (the pickups on their roofs) to latch on to the overhead contact lines and draw electricity. Trucks can feed electricity into the grid when they brake, making the system particularly useful if there's ever a jam.

To meet the challenges of an electric vehicle's short range, eRoadArlanda[5] combines battery power with direct power feeds while in motion. The test track is located on a ten-kilometer section of Road 893 between Arlanda Cargo Terminal and the Rosersberg logistics area, between Uppsala and Stockholm, of which two kilometers will be electrified for the demonstration project. The vehicle that are primarily planned to use the electrified road is an 18-ton truck that will be carrying goods for PostNord. As noted in [12] the high moving masses of commercial vehicles offer great potential for recuperation and therefore for an increase of the total efficiency. A system developed by SAF-Holland and the Institute of Mobile Machines and Commercial Vehicles at Technische Universität Braunschweig shows that electric systems can be used sensibly not only for the truck but also in the **trailer**. An electrified trailer axle generates electrical energy during braking, which can be temporarily stored and used for auxiliary consumers or for traction support later. Battery-electric trucks are very well suited for urban use in waste management due to the comparatively short and plan-able daily routes of up to 100 km with a high proportion of stop-and-go in inner-city traffic. With an anticipatory driving style, electrical energy can be recovered during braking to charge the battery, which further improves range and efficiency.

Motor sport has always been one of the main generators of new ideas and technologies in automotive technology. The increasing need to address the problem of depleting fossil fuels, as well as global warming and pollution, are putting strict limits on CO2 emissions from cars.

An example of the competitive developments in this sector is Shell Eco-marathon, which is a world-wide energy efficiency competition sponsored by Shell. Participants build automotive vehicles to achieve the highest possible fuel efficiency. There are two vehicle classes within Shell Eco-marathon: (a) Prototype - ultra-efficient, lightweight vehicles (generally 3 wheels) and (b) UrbanConcept - vehicles that have familiar road car features (always 4 wheels). They can choose the engine, determining the energy category: battery-electric, hydrogen fuel cell, and internal combustion engine (gasoline, ethanol, or diesel).

In other race events too electric cars were used for racing as early as 1899. At that time, it was not yet clear whether the internal combustion engine would be required as the main propulsion unit for mobile applications. Today, a revival of interest in electric racing cars are developing. They are mainly divided into two categories - speed racing cars in order to achieve the least amount of tour time (such as Formula E) and economical racing cars to minimize the energy consumed per distance unit (such as Shell Eco Marathon). To name a few prominent examples:

[5] http://eroadarlanda.com.

- Formula 1 - hybrid concepts
- Formula E - battery electric vehicle
- Formula Student Electric - battery electric vehicle
- Shell Eco-Marathon - hybrid/battery/fuel - cell vehicles

From the literature review and the conducted meetings and interviews, it can be concluded that the standard requirements for the commercial vehicle and race car electric powertrain engineers are the same as for passenger cars.

8 Implications Deducted for ECEPE Training Program

As presented in the previous sections of this paper, the new ePowertrain concept comes not only from the need for clean mobility, but also from the structural changes and new technologies Drivers-of-Change. In more general terms, it is one of the very specific domain concepts combining the needs of the new modern, clean, digital, and competitive solutions for the future of mobility. ePowertrain competence is a very much needed requirement of the present time and the ECEPE project approach is a possible reaction to this need. In training mechanism terms, the concept of work based learning combined with online training is taken into account and skills definitions and training materials of the ECEPE project should reflect that.

Thus, the survey shows that electric powertrain engineering is a job role where the industry expects a retraining of software, hardware, and electrical engineers with a specific focus on special topics needed in the different business fields. The survey of stakeholders clearly indicate that a tailoring of the topics is of utmost importance, as well as, a general adaptation on the audience.

The following topics are generally of greatest importance for the employing industries concerned:

- Overview on standards/directives/regulations/HV & fire considerations
- ePowertrain architectures (BEV, HEV, plugin HEV, fuel cell EV, in-wheel concepts)
- Building blocks of ePowertrain (battery, eMotor, inverter, etc.)
- Signal flow/interfaces definitions (real-time thinking)
- Hybrid control systems (eMotor/inverter/ECU) - SW Architecture
- Safety view point (systems design & item thinking)
- eMotor/generator control
- Battery systems

The feedback on the proposed curriculum structure of the ECEPE training material was much appreciated and accepted; no major topic lacking has been identified and the overall structure fits the general context of automotive ePowertrain engineering. The questionnaire did not reveal any major missing points or shortcomings. The distribution of the participants implies that more or less knowledge in the same subject areas is needed in every industry concerned. This shows that the demand exists for a course of the kind in all of the industries questioned, with the logical consequence that these industries should be addressed in the development work for a course of this kind.

In addition to these results, further knowledge in quality management, raw material sourcing, sourcing countries and salvage of used material should be included in the planned course curriculum. It can also be concluded from the literature review and from the meetings and interviews that the standard requirements for the commercial vehicle and race car electric powertrain engineers are the same as for passenger cars.

It should be stressed that the objectives are important in terms of environmental protection and the introduction of increasingly stringent emission requirements not only for the cars themselves, but also from the generation of electricity and the production of hydrogen, alternative fuels, etc. All the potential factors of influence are examined in the plan with the objective of minimizing the losses from a single car (electric vehicle) under different driving modes. Along with issues related to the electrical requirements, factors related to construction, the mechanical aspects together with the exploitation and the use made of the vehicle also need to be taken into account. This includes also basic questions of vehicle theory and design such as:

- impact of weight of construction
- influence of mass inertia of the rotating parts
- types of moving resistance and performance factors
- friction losses and hydraulic losses in transmission
- gearbox selection - regarding the control, type and range

9 Relation to SPI Manifesto

With this work we are contributing to the principles and values described in the SPI manifesto of the community [14]. Specifically we aim to enhance the involvement of people through training formats and thus improve the competitiveness of organisations (A.2). Our further objectives is to enhance learning organisations and learning environments (4.1) and thereby also to support the vision of different organisations and empower additional business objectives (5.1).

10 Conclusion

The automotive domain is a key industrial sector for Europe. Not only in Europe, but in all other parts of the world society is eager to achieve a significant improvement in environmental conditions and people everywhere are looking forward to the integration of clean mobility in the pattern of their urban lives. The global trend for electro mobility is moving forward driven by a significant progress with attractive market oriented cars. One of the key requirements analysed for supporting future developments in the automotive industry is training knowledge and skill sets for electric powertrain engineers.

In terms of the projected training mechanism, the concept of work based learning combined with online training have central roles and this should be reflected in the skill definitions and training materials of the ECEPE project.

To that aim, this paper has presented a stakeholder analysis of the different view points and their implications for the development of such a training concept. The aim of this analysis was to establish the skills basis required by companies developing the electric powertrain solutions of the future and to develop an adequate training concept for coming generations of electric powertrain engineers. The target audience for the training material is related to higher education of students participating in any automotive related courses program.

Acknowledgments. This work is partially supported by Grant of SGS No. SP2020/62, VŠB - Technical University of Ostrava, Czech Republic. The ECQA Certified Electric Powertrain Engineer project (*ECEPE*) is co-funded by the Erasmus+ Call 2019 Round 1 KA203 Programme of the European Union under the agreement 2019-1-CZ01-KA203-061430.

References

1. SOQRATES Task Forces Developing Integration of Automotive SPICE, ISO 26262 and SAE J3061
2. Kunith, A.: Elektrifizierung des urbanen oeffentlichen Busverkehrs. Springer, Wiesbaden (2017). https://doi.org/10.1007/978-3-658-19347-8
3. ACEA. Paving the way to carbon-neutral transport. Homepage (2020)
4. Cendon, E., Winkler, E., Prager, K., Schacherbauer, E.: Implementing Competence Orientation and Learning Outcomes in Higher Education - Processes and Practices in Five Countries. Project-Team HE_LeO (2008)
5. DRIVES Consortium. Erasmus+ BluePrint Project DRIVES. Homepage (2020)
6. ECEPE Consortium. IO1 - Study about the requirements for an electric powertrain engineer. Technical report, ECEPE Consortium (2020)
7. ERTRAC. European Roadmap Electrification of Road Transport. Homepage (2017)
8. European Automobile Manufacturers Association. The Automobile Industry Pocket Guide 2019–2020. Technical report, European Automobile Manufacturers Association (2019)
9. European Commission. Reducing CO2 emissions from passenger cars - before 2020. Homepage (2015)
10. European Commission. A European Green Deal. Homepage (2020)
11. European Commission. A New Industrial Strategy for Europe. Homepage (2020)
12. Ritters, K., Frerichs, L., Heuser, T., Drewes, O.: Investigation of electrified trailer axles by use of a functional model. ATZheavy duty worldwide **12**(3), 60–65 (2019). https://doi.org/10.1007/s41321-019-0039-1
13. Kolhe, M., Doyle, A., Muneer, T.: Electric Vehicles: Prospects and Challenges. Elsevier (2017)
14. Korsaa, M., et al.: The SPI manifesto and the ECQA SPI manager certification scheme. J. Softw. Evol. Process **24**(5), 525–540 (2012)
15. Macher, G., Brenner, E., Messnarz, R., Ekert, D., Feloy, M.: Transferable competence frameworks for automotive industry. In: Walker, A., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 151–162. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_12
16. SoQrates Working Group. ECEPE Questionnaire SoQrates. Homepage (2020)

Optimizing System Design Through Thermal Analysis with Cubesatellite as Case Study

André Wendland[1](✉), Marc Bumann[2], Shane Martin[3], Jörg Niemann[2], and Stephen Bosman[4]

[1] Department of Electrical Engineering and Information Technology, University of Applied Sciences Düsseldorf, Münsterstraße 156, 40476 Düsseldorf, Germany
andre.wendland@study.hs-duesseldorf.de

[2] Department of Mechanical and Process Engineering, University of Applied Sciences Düsseldorf, Münsterstraße 156, 40476 Düsseldorf, Germany
marc.bumann@study.hs-duesseldorf.de, joerg.niemann@hs-duesseldorf.de

[3] Department of Mechanical Engineering, Cape Peninsula University of Technology, Symphony Way, Bellville, Cape Town 7535, South Africa
martinsh@cput.ac.za, shanemartin92@gmail.com

[4] Department of Industrial and Systems Engineering, Cape Peninsula University of Technology, Symphony Way, Bellville, Cape Town 7535, South Africa
bosmans@cput.ac.za

Abstract. ZACube-2 is 3-Unit nanosatellite operating in low earth orbit. As for many electronic devices, the temperature within the satellite is an important factor for the function of some components. Following the assumption that thermal behavior can be modelled by software, ZACube-2 needs to be modelled, simplified, meshed and analysed in NX 12, a computer aided design and simulation software. To do so required an accurate but simple finite element model of the printed circuit boards (PCBs) for the various satellite subsystems. Specific material properties and finite element methods for modelling PCBs were found. These properties and methods were found by simulating the PCB behavior under thermal loading and comparing to that of tested specimens. The results showed an accuracy between the simulated and tested results of within 3%. The PCBs are used for almost all electronic hardware. They are complex and the approach used introduces a way of modelling them in a simplified but accurate method and using the findings for thermal modelling. These methods and procedures can be used in various industries that have electronics that are at the risk of overheating to improve the designing process and the efficiency of the product.

Keywords: Satellite · ZACube-2 · CubeSat · Thermal analysis · PCB

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 408–420, 2020.
https://doi.org/10.1007/978-3-030-56441-4_30

1 Introduction

1.1 Background

A CubeSat is a small satellite, specifically a so-called Nanosatellite [1]. They were initially developed by California Polytechnic University (Cal Poly) and Stanford University for educational purposes and space research [1]. CubeSats became increasingly popular amongst educational institutions. When talking about a CubeSat one should know that specific requirements need to be met. The cooperation of the two universities introduced the CubeSat Design Specification (CDS) which since publishing has been adopted by several organizations, educational and non-educational, from all over the world. As did the French South African Institute of Technology (F'SATI), based at Cape Peninsula University of Technology (CPUT) campus in Cape Town where the research of this paper has been done. A CubeSat can consist of a combination of various Units. One Unit (1U) is dimensioned 100 mm $*$ 100 mm $*$ 100 mm, like a cube [2]. A 1U CubeSat structure by Innovated Solutions in Space (ISS), a Dutch company, is shown in Fig. 1.

Fig. 1. One unit CubeSat structure by ISS

ZACube-2 is the second satellite that the African Space Innovation Centre (ASIC) at F'SATI launched as a follow-on mission of ZACube-1. It was launched on the 27th December 2018 and initially developed for tracking South Africa's coastal ocean traffic as well as providing real time information on fires to detect veld fires in advance [3, 4]. Meeting the CDS, the nanosatellite consists of three units with different payloads including a medium resolution matrix imager and various communication and computing subsystems which are shown in Fig. 2.

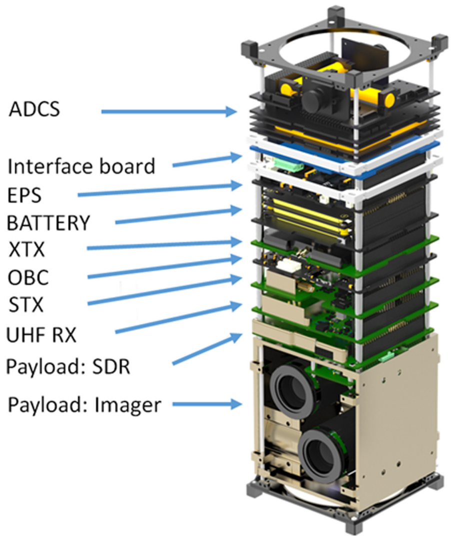


Fig. 2. Components of ZACube-2

Space is one of the most extreme environments existing, especially regarding thermal conditions and heat transfer. On Earth, heat can be transferred in three different ways:

1. Conduction: Heat is transferred by the physical contact between two bodies
2. Convection: Heat is transferred through a fluid, caused by molecular motion
3. Radiation: Heat is transferred in form of electromagnetic waves

The fact that there is vacuum in space makes convection through space almost entirely non-existent [5], so the primary way heat can be transferred through space is radiation and conduction within the satellite itself. Because of radiation, the satellite experiences huge temperature differences while orbiting earth. Additionally, components heat up while operating and transfer heat through the satellite, including the imager payload. At this point it is important to know, that the focus of the camera lens is set on earth based on a specific temperature and cannot be adjusted once placed into orbit.

1.2 Problem Statement and Objectives

As stated in 1995 by [6], "[..] thermal control has become a critical factor in the design of electronic equipment [..]". This is due to increased miniaturization of components and heat dissipation. As [7] describes, overheating of electronic components leads primarily to damage and therefore disfunction. This is also confirmed in a research of the US Airforce, which evaluated that more than 50% of failure of electronic components is due to temperature [8]. Since this research was performed, computational power increased. Nowadays it is possible to simulate heat development using software.

As pointed out within the with the SPI Manifesto, to be motivated for process improvement, one has to figure out the benefits [9]. The process of designing electronic components can be improved by simulating the heat transfer using software. This can have positive impact, especially regarding better heat control. The benefit here is, that failure can be prevented before manufacturing. By performing software based thermal analysis of Cubesatellites, the thermal effect of the space environment can be predicted to optimize the design of the satellite. As stated in [3], the operating temperature for some components within the satellite is of concern. If subsystems within the satellite are to overheat, this could lead to the subsystem failing which in the case of satellites, would highly likely result in a mission failure, resulting in a loss of resources invested into the mission as well as loosing the research output that would have been acquired.

Cubesatellites have multiple different subsystems that are made up of various different materials. One of the primary paths of heat conduction within the satellites is through the PCB board of the sub-systems. A PCB board however, is made up of a combination of materials and not in standard material databases.

Considering this fact, the project scope aims on providing material properties to be used in developing thermal finite element models (FEM) to perform a thermal analysis of ZACube-2 and future satellites. It will specifically focus on simplifying the CAD models of sub-systems and setting up a static thermal analysis to be compared and validated against physical tests. This can be proven by comparing the results of the analysis to the measured samples. The project requirement is to stay within 5% deviation in the analysis compared to the samples measured experimentally. The results can be used to recommend set up methods and procedures for analysis to compare against ZACube-2 temperature data and for future satellites.

2 Thermal Analysis

The simulation in this study is done by using Siemens NX as computer aided design and engineering (CAD/CAE) software. The analysis is performed by making use of the finite element method. ZACube-2 is objective of this study and is used to validate the assumption.

> *"Finite element methods provide a widely used tool for the solution of problems with an underlying variational structure."* [10]

Within this analysis the underlying structure is the PCB board itself and each of its components and the problem to be solved is the heat transfer through the board due to heat loads and environmental constraints. The following limitations needed to be faced and will be explained further:

1. Computing time when running a simulation
2. Accessibility of the satellite's sub-systems and their temperature data
3. Heat flow through layers of a printed circuit board (PCB) within the simulation

To reduce the computing time, every PCB board component was simplified and remodeled. The simplification is done based on measurements obtained with a FLUKE Infrared camera. The process of the simplification is shown in Fig. 4 and will be explained in the following paragraphs. For this analysis, only the heat sources on the components are of interest, as well as the parts which dissipate the heat. The simplification process is demonstrated based on the sample shown in Fig. 3. The measurement of the components was performed under operating conditions at different room temperatures. The total transfer in heat, being the constraint input for the simulation, can be simplified as heat load = $\mathrm{power}_{\mathrm{in}} - \mathrm{power}_{\mathrm{out}}$.

The setup of a static analysis consists of four different constraints.

1. The environmental conditions were measured and the temperature values fed into NX.
2. Convection to environment on earth is an important form of heat transfer [11]. Heat will be transferred from the hotter area, in this case the chips on the PCB, to the cooler area, the air. As explained briefly in the introduction, convection occurs by molecular movement. The convection coefficients can be looked up in data sheets depending on the material. This value in $[\frac{W}{m^2 K}]$ was implemented to the analysis.
3. Radiation occurs everywhere and does not require a medium in which to transfer heat but rather its view factor to the environment around it. To consider radiation within the analysis it needs to be set up on the board. This is done by adding it as a constraint to the surfaces of the board and the chips which radiate to the environment. To calculate heat transfer through radiation the solver needs the absorptivity and emissivity coefficients of the different materials. Those coefficients are shown in Table 1. The values of the coefficients can be implemented to the material library and will be considered when adding a radiation constraint.

Table 1. Thermo-optical properties [12–14]

Material	Emissivity	Absorptivity
Silicon	0.3	0.7
Printed circuit boards (material mix)	0.85	0.85
Aluminium	0.05	0.15
Titanium	0.35	0.55

4. In Fig. 3, the crucial heat sources are shown in red as well as the highest temperatures. They conduct to the printed circuit board and need to be considered as well. The component was measured several times at operating conditions to acquire this data. As mentioned before, the total heat load is simplified as $heat\,load = power_{in} - power_{out}$. The total heat load was then allocated to the single crucial heat sources on the simplified model by calculating the ratio of temperature increase from initial status to operating status. Within this sample one can see three crucial heat sources. The values mentioned above need to be determined individually considering the material, initial temperatures, room temperature and heat loads and then implemented to the analysis. When applying the FEM to each of the simplified models, the structure gets divided into a finite number of elements with a certain amount of nodes. These then will be used by the solver of the simulation software to

calculate the heat transfer at each node, the more nodes and elements, the longer the calculation. Logical consequence to keep the computing time low but still achieving an appropriate accuracy is to suppress all components which are not crucial for the heat flow through the satellite.

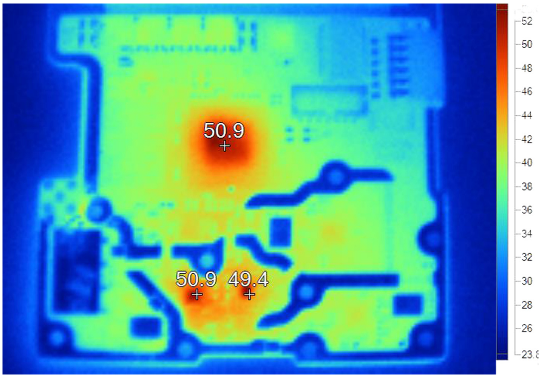


Fig. 3. Measured sample

The process of simplifying the CAD models is shown in Fig. 4. Model (A) in Fig. 4 is the original CAD model, (B) and (C) are meshed simplifications. (D) shows a section few of (A) to demonstrate how the components are attached to the PCB.

A static analysis of the board in Fig. 4 was setup with three applied heat loads based on the measurements as discussed in point 4 previously on each of the chips shown in Fig. 3 (C).

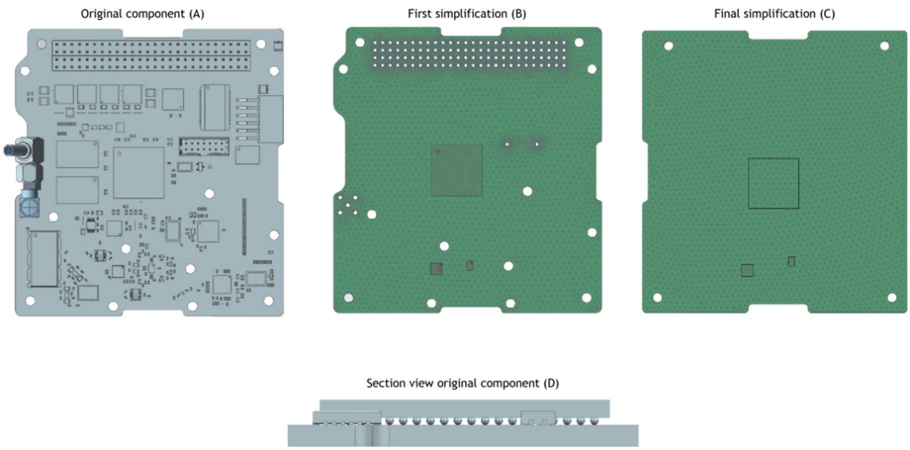


Fig. 4. Simplification process

The computation time of model (B) displayed in Fig. 4 was significantly lower than model A. However, since the whole satellite consists of several PCBs and other components, the model needed to be simplified further. From model (B) to (C) the computation time could be reduced by an additional 94.19%. The computing power of the computer used for the simulation is decisive for the time range of the simulation. Nevertheless, it was an important aspect for this project to keep the time as low as possible so that simple changes can be done to see the impact within appropriate time. Although the second model had less detail, the coarser mesh had a negligible impact on the results while significantly reducing the computation time.

As shown in the section view (D) the components are either connected through solder balls or legs. Assuming a perfect contact between PCB and chips leads to less computing time because the solder joints between PCB and components need to be meshed very fine which leads to many small elements and nodes. Additionally, the header pin holes near the top of the PCB have been closed as the mesh around them would need to be very fine to show accurate results. Those fine meshes then lead to a significantly higher number of equations being solved by the solver.

The PCB in model (C) was furthermore replaced by a standardized PCB. This decision was taken because all the PCBs used in the satellite are variations of the PC/104 standard. The model needed to be meshed once and since the goal of this analysis is to be easy implementable to further satellites it is convenient to have only one model being used for every PCB. Comparing the original shape of the PCB in (A) to the standardized in (B) one can see a slight difference.

The static simulation includes convection, conduction and radiation as means of heat transfer just as the measured sample. After applying the mentioned simplifications to the components, the static results of the analysis were accurate and the computing time was appropriate. Even though the heat transfer in space is quite different, the static analysis was necessary to ensure that the simplifications and material properties, which will be explained in section “3. MATERIAL PROPERTIES AND PCB LAYER SIMPLIFICATION”. The comparison of the measured samples and the result of the analysis being evaluated is shown in section “4. RESULTS AND DISCUSSION”.

3 Material Properties and PCB Layer Simplification

The crucial thermal properties of a material for thermal analysis are:

1. Specific heat
2. Conductivity
3. Density

These properties are well known for standard materials such as Aluminium, Titanium etc. and can be implemented to the simulation. NX provides these properties for some of the analysed materials but since PCBs, as shown in Fig. 5, can consist of multiple different and very thin layers another method was used to define their properties.

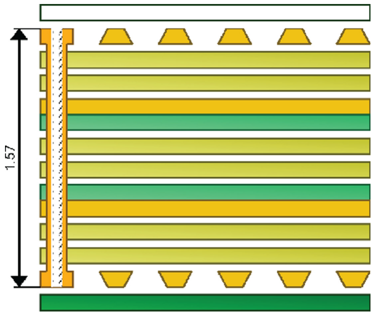


Fig. 5. PCB stack up by POLAR

Instead of modelling and meshing each single layer it is assumed that the PCB is one simplified material. This assumption has been made to keep the computing time on a minimum level because the satellites' inside components mainly consists of PCBs. To achieve accurate results the material properties of the PCBs have been averaged and calculated which will be discussed in the section "Results and discussion".

4 Results and Discussion

As shown in Table 2 an averaged material mix for the PCB has been calculated based on thickness and material of the layers seen in Fig. 5. The material properties for the dielectric was taken from the PCB manufacturers data sheet while that of copper was extracted from the Siemens NX database.

Table 2. Simplified material PCB

Material	Thickness [mm]	Thickness [%]	Specific heat C_p $[\frac{J}{kg*K}]$	Thermal conductivity k $[\frac{W}{m*K}]$	Density ρ $[\frac{g}{cm^3}]$
Dielectric	1.427	91	1200	0.5	1.97
Copper	0.146	9	385	401	8.96
Simplified material	**1.573**	**100**	**1126.65**	**36.54**	**2.6**

These thermal properties were then used to create a custom material for the PCBs in the simulation.

The setup of the physical experiment as explained earlier is demonstrated in Fig. 6. The measurement taken with an infrared camera is seen on the left side and the results of the analysis on the right side.

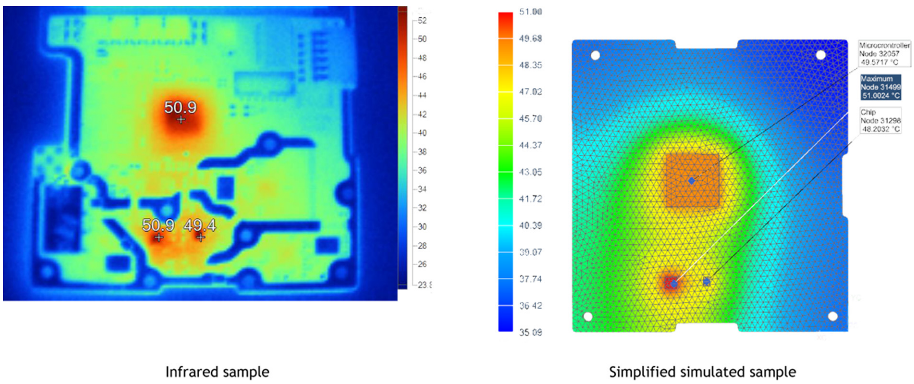


Fig. 6. Comparison measured sample and simulation

The largest component on the board displays a measured maximum temperature of 50.9 °C for the experiment with the additional two components showing temperatures of 50.9 °C and 49.4 °C. After multiple tests, the average maximum temperature for the large component was 50.65 °C.

For the analysis, the large component resulted in a temperature of 49.57 °C and 51 °C and 48.2 °C for the additional two components. These results were acquired by running the simulation setup as explained in Sect. 2 "THERMAL ANALYSIS", making use of the mixed material properties, a ratio of the thermal heat load applied to the PCB components, a simplified finite element model and environmental constraints.

Comparing the results of the physical experiment and the simulated thermal analysis, the large component varies by 1.33 °C while the additional two vary by 0.1 °C and 1.2 °C respectively. This results in an average deviation between the experiment versus simulation of 2.3%. With the accuracy of the properties and methods used to define the PCB boards, the same methods can be used and implemented into a larger finite element model of an entire satellite for further research.

The whole model of the simplified satellite and the materials allocated to each component for the finite element model is shown in Fig. 7.

Copper can be used for the copper coils of the attitude determination and control system, silicon for electronic components on the PCBs which themselves having the PCB material mix applied to. The structure consists of aluminium and the stack rods on which the PCBs and sub-systems are attached to the satellites' structure are made of titanium. The lenses of the camera module are manufactured of anodized aluminium and the fasteners are made of stainless steel. A section view has been chosen to show the interior of the satellite. The thermal, optical and physical properties which are necessary to perform a simulation are shown in the table below (Table 3).

Based on this foundation the orbital analysis will be performed and compared to the real-time data measured by sensors which were implemented into the satellite and available as telemetry data. For implementing the methods and results found in this study, there are other space environmental effects that need to be considered. A major

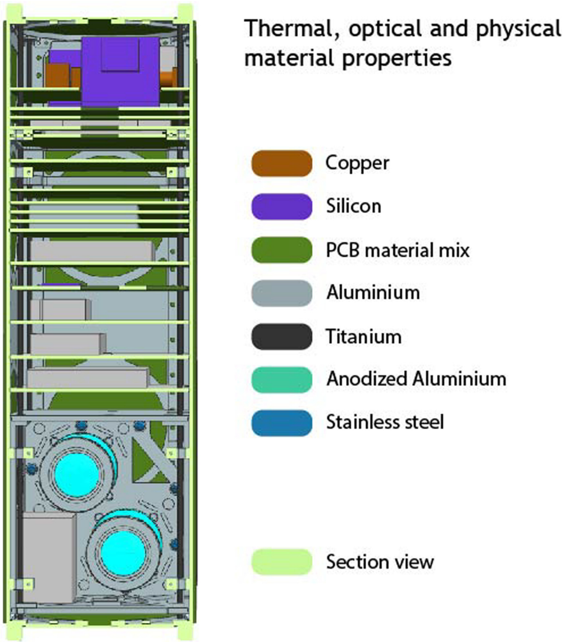


Fig. 7. Thermal model of ZACube-2

Table 3. Thermal material properties [15–17]

Material	Specific heat C_p $\left[\frac{J}{kgK}\right]$	Conductivity k $[\frac{W}{m*K}]$	Density ρ $[\frac{g}{cm^3}]$
Silicon	712	159	2.33
Aluminium_6061	896	NX settings	2.77
Copper	385	387	8.92
AISI_410_SS	460	NX settings	7.73
Titanium _Ti-6Al-4 V	526	6.7	4.43

effect being the thermal radiation experienced in space. When simulating the satellite in orbit, additional radiation constraints need to be added and are visualized in Fig. 8:

Space radiation effects can be broken up into three factors;

1. Direct radiation
2. Albedo radiation
3. Earth infrared radiation

These will increase the computing time immensely. Furthermore, depending on the calculation intervals, it is dependent from time so that the calculations for the whole satellite have to be done repetitively. For the orbital simulation, a calculation interval can be chosen which gives you the “resolution” of the results, how frequently you

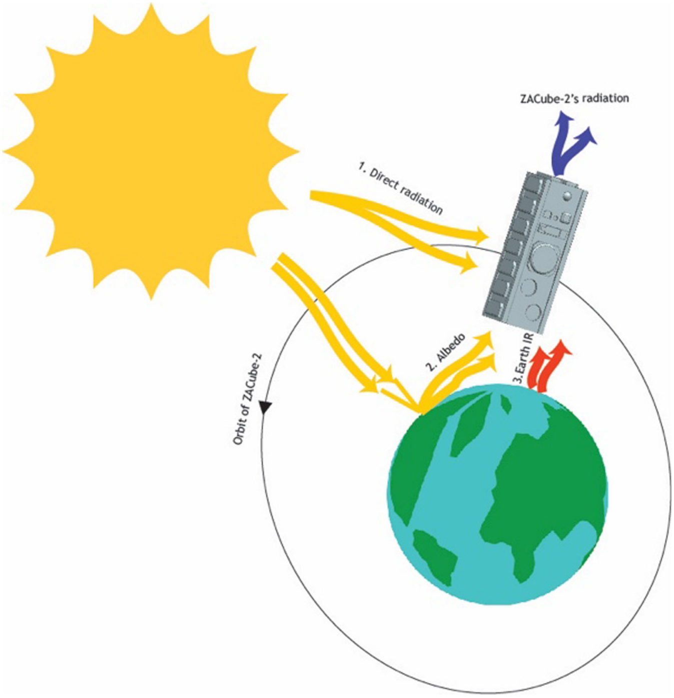


Fig. 8. Heat impact in space

obtain thermal results. Orbital heating can be setup with an orbital inclination of 97.38°, 484.9 km minimum altitude, an eccentricity of 0.0022 and an argument of periapsis of 124.5° as well as an orbital period of 5670 s. This is based on the current orbital parameters supplied by the current ZACube-2 two-line element set (TLE). TLE is a common way of tracking the position of satellites [18].

1. The radiation is setup for the outer surfaces, such as solar panels and the front side of the satellite, as well as for the inside of each of the mentioned components in Fig. 2.
2. Albedo is defined as indirect radiation from the sun reflected by a planet [19], in this case earth and the setup for the satellite's orbit is implemented to the simulation software.
3. Earth emits infrared energy which reaches the satellite, the heat flux depends on the position of the satellite and is calculated by the solver of the simulation software.

> "*Altitudes between 120 and 600 km are within the Earth's thermosphere, the region above 90 km where the absorption of extreme ultraviolet radiation from the Sun results in a very rapid increase in temperature with altitude.*" [19]

Since ZACube-2 orbits in Low Earth Orbit (LEO) in a range between 480 km and 510 km, heat besides the heat loads of the satellite itself, is a very important factor to determine the temperature range of the satellite. The impact of radiation on the satellite depends on the absorptivity and emissivity coefficients of the materials. The coefficients as shown in Table 1 can be used for the analysis. Absorptivity describes the attitude of a body to absorb radiation as thermal energy. Emissivity describes the opposite, it describes how much thermal energy an object emits. The values lay between zero and one and are related a black body which is a perfect absorber and emitter (Absorptivity and emissivity equal one) [20]. Radiation can be setup into the orbital analysis in the same way as in the previous static analysis. In space the impact of radiation will be more noticeable since convection in this environment is negligible.

5 Conclusion

It was seen how Cubesatellite systems can be significantly simplified, while still keeping an accurate representation of its thermal characteristics. This procedure can be transferred to any thermal model. Reducing a fully populated PCB board to a simplified basic PC/104 board with only the significant heat producing components resulted in a model that was efficient to simulate while resulting in simulated thermal conditions that were within 2.3% of that of the experimental test. Conducting the experimental test with the thermal camera allowed for the thermal dissipation across the board to be captured. This was vital in the validating of the simulated results. The accuracy of the simulation was within the objective of 5% which allows these results to be implemented into a larger fully Cubesatellite model such as that of ZACube-2.

By simplifying the model to make use of widely used materials as well as the PCB material mix used for the PCB simulation, a thermal model can be created to be used for an orbital thermal simulation. Implementing the space environment factors such as radiation and orbital parameters, the satellites thermal conditions throughout its orbit can be simulated.

Further research should be conducted using these methodologies to compare the simulated thermal conditions versus that of actual data generated by ZACube-2. This will then validate the accuracy of the model and methods used in simulating Cubesatellite thermal conditions. In accordance with the SPI Manifesto, the process can therefore be reinforced and optimized, leading to a sustainable improvement of the status quo [9]. Additionally, the results can be used to develop detailed thermal control systems, to calculate component lifetime and have a better idea of the thermal budget of future satellites.

While this research focused on a Cubesatellite as the case study for the PCB thermal modellings, these methods and procedures can be implemented in a wider industry that uses electronic components such as the automotive, aerospace and telecommunications industries. For the space industry, an electronic failure can result in mission failure, loosing significant amounts of invested resources. For other industries, such as automotive, an electronic failure could result in a catastrophic failure causing injury or even death to a vehicles occupants. By being able to accurately predict the thermal effects of various heat sources of a system, the electronic component can be optimally designed to prevent malfunctions and failures and improve its efficiency.

References

1. Mabrouk, E.: NASA - What are Cubesats? http://www.nasa.gov/content/what-are-smallsats-and-cubesats. Accessed 01 June 2020
2. Developer Resources — CubeSat. http://www.cubesat.org/resources. Accessed 13 Mar 2019
3. Griffith, D., Cogan, D., Magidimisha, E., Van Zyl, R.: Flight hardware verification and validation of the K-line fire sensor payload on ZACube-2. In: du Plessis, M. (ed.) Fifth Conference on Sensors, MEMS, and Electro-Optic Systems, Skukuza, South Africa, p. 100. SPIE (2019). https://doi.org/10.1117/12.2503094
4. ZACUBE-2 - Satellite Missions - eoPortal Directory. https://directory.eoportal.org/web/eoportal/satellite-missions/content/-/article/zacube-2. Accessed 04 June 2020
5. Lienhard, J.H., Lienhard, J.H.: A Heat Transfer Textbook. Dover Publications Inc, Mineola (2019)
6. Yeh, L.T.: Review of heat transfer technologies in electronic equipment. J. Electron. Packag. **117**, 333–339 (1995). https://doi.org/10.1115/1.2792113
7. Almubarak, A.A.: The effects of heat on electronic components. IJERA **07**, 52–57 (2017). https://doi.org/10.9790/9622-0705055257
8. Fortna, H.C.: Avionics integrity issues presented during NAECON 1984, Ohio (1984)
9. SPI Manifesto. https://2020.eurospi.net/index.php/manifesto. Accessed 12 Dec 2019
10. Schmidt, A., Siebert, K.G.: Design of Adaptive Finite Element Software: the Finite Element Toolbox ALBERTA. Springer, Berlin (2005). https://doi.org/10.1007/b138692
11. Bejan, A.: Convection Heat Transfer. Wiley, New York (1984)
12. Lide, D.R.: CRC Handbook of Chemistry and Physics. CRC Taylor & Francis, Boca Raton (2008)
13. Emissivity - Metals | Fluke Process Instruments. https://www.flukeprocessinstruments.com/en-us/service-and-support/knowledge-center/infrared-technology/emissivity-metals. Accessed 04 Jan 2020
14. Thermal data. http://webserver.dmt.upm.es/~isidoro/dat1/index.html. Accessed 04 June 2020
15. ASM Material Data Sheet. http://asm.matweb.com/search/SpecificMaterial.asp?bassnum=MQ410BF. Accessed 03 Feb 2020
16. ASM Material Data Sheet. http://asm.matweb.com/search/SpecificMaterial.asp?bassnum=MTP641. Accessed 03 Feb 2020
17. Specific Heat of some common Substances. https://www.engineeringtoolbox.com/specific-heat-capacity-d_391.html. Accessed 03 Feb 2020
18. Transilvania University of Braşov, Braşov, Romania, Croitoru, E.-I., Oancea, G., Transilvania University of Braşov, Braşov, Romania: SATELLITE TRACKING USING NORAD TWO-LINE ELEMENT SET FORMAT. AFASES 2016, vol. 18, pp. 423–432 (2016). https://doi.org/10.19062/2247-3173.2016.18.1.58
19. Wertz, J.R., Larson, W.J. (eds.): Space Mission Analysis and Design. Microcosm; Kluwer, El Segundo; Dordrecht; Boston (1999)
20. Çengel, Y.A.: Heat Transfer: A Practical Approach. McGraw-Hill, Boston (2003)

A New Concept for People Movers on Single-Track Secondary Railway Lines

Walter Sebron(✉), Elaheh Gol Hashem, Peter Krebs, and Hans Tschürtz

Vienna University of Applied Sciences, 1100 Vienna, Austria
{walter.sebron, elaheh.gol_hashem, peter.krebs, hans.tschuertz}@fh-campuswien.ac.at

Abstract. This paper introduces a safety concept for autonomous on-demand people mover systems intended for single-track secondary railways lacking grade separation, with the aim of revitalising unprofitable or even abandoned secondary lines. The considerations in this paper are backed by an ongoing research project, which is publicly funded by the MA23 (MA23 is the research funding department of the magistrate of the city of Vienna). The paper discusses the motivation for a non-traditional railway system using stochastic demands instead of fixed timetables and the main challenges that result from this paradigm change. In particular, the impact on the safety and availability of such systems is considered, arguing the need for a tailored safety concept that differs from traditional railway safety approaches.

Lastly, pointers for further research and development are given that need to be addressed in order to deploy on-demand people mover systems in a safe and dependable way.

Keywords: Autonomous driving · Safety · Availability · Railway · On demand · System of systems

1 Background

1.1 Introduction

Today's secondary railway lines draw a very sad picture of railway transportation at the beginning of the 21st century, especially looking at people transport. In most cases, their lack of flexibility and poor performance makes them unattractive for passengers. This in turn makes their operation unprofitable or may even incur a loss, causing those lines to be shut down or to be abandoned completely. Therefore, traffic is shifting to road (automobiles and busses), leading to traffic congestion, and growing environmental pollution and stress. This is a challenge for rural areas, and even more for fast growing suburban zones, which are recently shifting into the focus of many "smart city" initiatives. Thus, new concepts for railways become necessary in order to support the general objective of improving the modal split (i.e. the share or percentage of people using a particular type of transportation) from individual towards public transportation.

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 421–432, 2020.
https://doi.org/10.1007/978-3-030-56441-4_31

According to [1], the major drivers for the cost of railways are:

- maintenance costs for infrastructure and rail network, including costs for human maintenance labour,
- train operation costs, including maintenance of rolling stock, and human operators (such as drivers, dispatchers, or ticket vendors),
- and corporate overhead costs.

Since the first and the last of these cost driver classes cannot be avoided and hardly be reduced, the focus for cost reduction rests on the train operation costs, and within these on the human operators. A number of reports and studies also support this claim. For instance, [2] claims that some studies indicate a halving in operational costs, mainly due to staff costs reduced by abolition of the drivers' function, even in cases of line conversion, when staff is likely to be retrained and deployed to other functions. Additionally, [3] states that the split between infrastructure and operator costs is approximately 30%:70% on an European average. While maintenance costs (human and technically related) are inevitably necessary, human operating staff and therefore its costs can be largely avoided when employing autonomous systems. This would tackle one of the biggest obstacles for cost-effective operation of secondary lines.

Following this line of thinking, the TramD initiative, whose primary objective was to provide a solution for the severe increase of individual traffic in the suburban area of Kaltenleutgeben south of Vienna, proposed the concept of a "Tram on Demand". This concept was based on the idea of a fully autonomous tram employing lightweight vehicles on the track of an abandoned secondary line. These can be called upon demand like the cabin of an elevator, enabling a local, flexible public transport service without downtime - for details refer to [4].

Although the basic principles of this idea are easily understood, it raises a number of issues, especially regarding the safety of such novel systems. In general, it must be shown that an on-demand system can operate with the same or, ideally, an even higher level of safety than traditional railway systems – despite the differences in their operational principles. Consequently, a safety concept tailored to these kinds of systems is indispensable for ensuring the acceptance of this new type of transportation systems, both from a technical as well as societal perspective. The development of such a safety concept for autonomous people movers on single-track secondary lines is the focus of a research project (publicly funded by the MA23[1]) whose preliminary results are summarised in this paper.

The aim of this paper is to introduce and discuss this new concept, and to provide a first overview on the findings of the corresponding project so far. Secondary lines are usually characterised by running through open areas with no grade separation[2] or other protective measures, as opposed to existing people mover and metro systems. Additionally, a report from the Austrian national railway provider, ÖBB Infra [5], shows that 55% of the Austrian lines are single-track only, the majority of which are

[1] MA23 is the research funding department of the magistrate of the city of Vienna.

[2] Operating on closed tracks separated from the environment by walls and fences, tunnels and bridges, underground, or as elevated railway.

secondary lines; this suggests, that there is a large potential for deployment of a specialised railway concept tailored to the characteristics of secondary lines.

The biggest challenge for such a specialised railway concept lies in the single track scenarios, where several autonomous railcars are operated on-demand with a 24/7 service. Due to the on-demand operation, the dispatching of the railcars to the individual stations is nondeterministic which results in movement patterns of the railcars that differ significantly from those in traditional railway operation. This places high demands on both availability and safety of such specialised systems. These properties are reflected in the concept's name "Autonomous Secondary-line on-Demand open-Track Systems" (AuSoDoTS). The corresponding project's focus is on safety for AuSoDoTS from a top-level perspective, the so-called "system of systems". The purpose of this limitation is to consider the safety and availability of the AuSoDoTS concept as a whole, omitting lower level problems or hazards arising from electronic defects, component failures, or sensor shortfalls and outages; these problems are tackled by traditional safety standards and by numerous other projects.

Note that, in [6], Leveson criticises the term "system of systems", arguing that this makes little sense with respect to systems theory and systems engineering, since there is always a meta system level that includes all such systems. Therefore, a system is a set of components, which are interrelated and share a common environment; this abstraction can be easily translated into a railway system and its constituting components, or sub-systems. Even though this perspective has its merit, the argumentation provided by IEEE-RS in [7] justifies the continued use of the term "system of systems" (SoS) when it comes to differentiated and highly complex (sub-) systems, such as the track itself (e.g. rails, points, loops, grade separation), the railcars, the infrastructure (e.g. stations, crossings, signals), the human actors (like passengers, or operators) and the communication in between these (sub-) systems where necessary, including (passenger) user interfaces. Additionally, this SoS perspective provides a way to distinguish the approach followed for the AuSoDoTS concept from related work that deals with details of the single (sub-) systems.

1.2 Related Work

Among the multitude of work on people movers or railway automation, probably the closest one to the AuSoDoTS approach is the project "autoBAHN", covered in [8]. It deals with a "model and development of an autonomous railway system", looking into main routes and secondary lines, focussing on sensor systems, hardware and software rather than on the SoS perspective. It provides a good overview at that level; however, it only cursory investigates on-demand operation, and does not consider the role of the railcars, or safety in necessary detail.

In the field of autonomous trams, first experiments based on tram assistance systems exist, see e.g. [9], as well as safety investigations for collision avoidance assistance systems for trams, see [10]. These experiments and investigations strongly focus on sensor systems and sensor data interpretation.

An overview of all present automated metro lines can be found in [11], or at the homepage of the International Association of Public Transport (UITP); however, these are all driverless closed-track systems, often with a limited or no autonomy (e.g. remotely controlled) in the sense of the AuSoDoTS concept, which also do not offer an on-demand service. A few automated and driverless lines are exemplarily listed here, to highlight the differences with respect to AuSoDoTS:

- In Nuremberg, Germany, the double-track U2 and U3 metro lines run underground, and they operate according to a strict timetable.
- In Yongin, South Korea, the double-track “EverLine” features different types of grade separation, as in town areas its track is elevated, in rural areas its track runs mainly at street level but is fenced off completely, without containing any crossings. It operates at intervals of about 6 min.
- In the U.S., Detroit’s “Downtown People Mover” follows a one-way single-track circular route completely built as elevated railway; it operates at intervals of about 3 min.

2 The Concept of AuSoDoTS

As stated above, the AuSoDoTS’ SoS-view contains several systems:

- the track itself,
- the railcars,
- the infrastructure,
- the human actors,
- and the communication.

Each system is analysed with respect to its specific behaviour (also within the SoS) and its interaction with the other systems. For the sake of brevity, only main points for these systems are discussed here, to offer some more insight. The limitations of the SoS lie of course in the physical boundaries, like track length, number of stations and railcars, number of points and loops in the track, but also in the financial aspects, as probably existing infrastructure should preferably be used and only little additional investments should be made. For communication, there are no new specific needs for this concept, just the same as for every safety-relevant communication: it has to be secure as well. Regarding the human actors, the project focus is on behaviour that might result in hazards. Overall, the AuSoDoTS concept is intended for the highest level of automation according to Grade of Automation 4 (GoA4), also known as Unattended Train Operation (UTO, see [11]).

The AuSoDoTS functional principle can be compared to a “horizontal lift”: Passengers can place a demand at any station of a track in order to call an available railcar, selected by arbitrary dispatching rules. Once the railcar arrives, the destination can be freely chosen, i.e. there is no fixed station order. The resulting system behaviour can be described as stochastic with regard to the arrival and completion of demands, which is a major difference to traditional railway operation using a timetable.

2.1 Tracks

Tracks consist of rails, points, loops, and further elements, on which railcars move between stations, whereby the movements of the railcars are controlled by some dispatching and interlocking logic according to the demands, in order to transport the passengers safely and within a certain time to their destination. The complexity of this task increases with a higher number of the previously mentioned (sub)systems of the SoS, as for instance with the number of stations. To make this possible, the individual movements of the railcars between stations and the tracks linking them together must be defined and then carried out in a timely and coordinated manner. In order to ensure this, all systems must follow certain rules during operation. Therefore, construction rules for tracks have been defined, that deal with the following aspects, among others:

- The number of total railcars supported by a track
- The necessary number of points in stations
- The optimal number and placement of loops between stations
- The number and placement of signalling elements

The set of rules will directly influence the layout of tracks, and therefore the system behaviour, resulting costs (e.g. for modifications), and performance.

2.2 Rolling Stock

The capacity and resulting size of the railcars depend on the needs of a specific secondary line, where it is planned to deploy the AuSoDoTS concept. Two types of railcars were assumed: a mini-railcar for up to 8 passengers, and a small railcar for about 20 passengers. It is obvious that the size of the railcars directly determines some of their properties (e.g. weight, acceleration, brakes, number of seats), and also impacts other factors of the AuSoDoTS SoS, like the minimum necessary platform length for boarding and alighting of passengers, or the minimum length of loops used by railcars to cross each other between stations. Additionally, all railcar types are assumed to be able to move bidirectionally, i.e. the traditional distinction between the "front" and "rear" of a train is no longer meaningful.

From the yearly report by ÖBB Infra [5], it can be seen that only about 53% of the Austrian secondary railway lines that are operated by ÖBB Infra, i.e. not privately owned, are electrified. This implies that railcars need to be designed in a lightweight fashion with minimised power consumption, using an on-board energy supply provided by e.g. batteries or supercaps. The lightweight design of course directly affects above mentioned railcar properties, but also has implications on safety hazards and crash consequences, due to the difference in mass compared to a conventional train.

The AuSoDoTS concept further requires the railcars to be working in a bidirectional way, to omit unnecessary turnarounds and for the sake of operational flexibility. Whether passenger doors should be on one or both sides of a railcar depends on the needs of a specific line, i.e. whether the platforms in different stations exist on both sides of the track (which may differ from station to station) or just on a specific side, as in many people mover systems. Distance and track monitoring by railcar-mounted sensors is mandatory in both directions of a railcar.

2.3 Infrastructure

The topology and layout of stations (e.g. number of tracks, platforms and stopping points) are important, as they directly influence the performance of the AuSoDoTS SoS. For example, the number of stopping points for the railcars at one single-platform station can limit the total number of railcars on the track; if there is only one stopping point, this can raise the danger for deadlocks in the station area, especially during high traffic periods. Different topologies and their pros and cons for AuSoDoTS are currently under investigation in the project. Crossings should be enhanced with automatic surveillance and communication in order to inform oncoming railcars about the crossing's status. In order to lower the deployment cost virtual signalling elements will be used.

3 Applied Methods

3.1 Analysis

Regarding safety, analysis methods are extremely important. For the systems of the AuSoDoTS SoS, the Extended Shell Model Method [12] was utilised to get a clear picture of the respective system's properties and components. Subsystems and functions which were identified as critical were analysed, by using the "Hazard and operability study" (HAZOP, [13]) and the "Functional Failure Analysis" (FFA, [14]) methods. With these classical methods, however, only hazards can be identified that result from single parameter or functional deviations.

Due to the SoS perspective, the need for a specific safety analysis method for critical situations and scenarios within a more complex SoS became apparent. A method based on the concepts explained in [15] and [16] is currently used and further adapted and refined in the project. The application of this method is still in progress. The method starts with a use case. A use case describes the service to be performed by an autonomous system within a defined super system, i.e. an SoS. A use case has a starting point and one or more defined goals, and can be described as a tree-like structure, where each branch symbolises a path to a single goal. A scenario describes a sequence of actions/steps within a use case that lead to a defined goal (i.e. a single path in a use case). The individual actions/steps are called scenes. A scene describes a snapshot of the environment, including its landscape, the dynamic elements (e.g. vehicles, or living beings) as well as the self-representation of the ego vehicle (machine performance) and the relationships between these entities. Within a scene, multiple events can occur even simultaneously, which can alter the outcome of the scene, resulting in different successor scenes. Via this method, specific scenarios (e.g. a railway crossing) are examined for critical situations in a scene-by-scene approach. Conditional modifiers (e.g. different weather conditions) are applied to identify worst-case situations. The aim is to identify so-called "SoS hazards" within the scenarios.

Finally, all analysis results were and are subjected to final reviews by railway and safety experts.

3.2 Simulation and Demonstration

In order to assess the dynamic properties of tracks deploying the AuSoDoTS concept, simulation of train operation is performed, using industry-grade simulation tools modified to support the unique characteristics of on-demand operation. The focus lies on identifying possible areas of improvement concerning performance metrics, such as waiting times and passenger throughput under different use profiles of a track (e.g. sporadic traffic vs. rush hour).

In order to additionally demonstrate the on-demand behaviour of AuSoDoTS, an automated model railway was built up to assess specific scenarios and user interaction.

3.3 Telephone Survey

In order to better understand basic passenger requirements and expectations, and to learn the acceptance criteria of the concept, a telephone survey was conducted. In total, 413 people living along existing or abandoned branch or secondary lines in Austria have been introduced to the AuSoDoTS concept and interviewed according to preselected questions. Among others, questions on intended usage purpose and usage frequency, on acceptable demand response times, and generally on likeliness of acceptance of such an AuSoDoTS system and the expected level of safety, were raised in the survey. Results of the analysis have been incorporated into the safety concept as basic assumptions and requirements from the perspective of passengers.

4 Safety Versus Availability

One of the first steps of the project was to follow a philosophy of balance between safety and availability, or more precisely: availability gets in conflict with safety if a safe state "standstill" exists, as proposed by traditional railway safety standards.

The telephone survey showed that although a majority of persons put a strong reliance in the safety of AuSoDoTS (approx. 63% believed that due to AuSoDoTS the number of railway accidents affecting humans would stay the same or even decrease, while only about 10% believed it would increase), safety seems to be expected self-evidently. This is why first the definition and analysis of different concepts for safety was done, and then one safety concept was selected. This way, the remaining problems in balancing safety and availability at SoS-level are coming from availability limitations resulting from the chosen safety concept, which puts efforts on resolving them via availability improvements.

4.1 Foundation of the Safety Concept

One major railway hazard, the "train-train collision", is prevented as with traditional systems by using signalled fixed block operation. Hereby, the track is divided by signals into non-overlapping sections called blocks. The signalling and interlocking system regulates access into blocks for trains, so that each block contains a maximum of one train at any one time, and that trains are prevented from entering an occupied block (i.e. at the corresponding signals). If a train wants to traverse a section of the track, the interlocking system must allocate the necessary blocks in the correct order at the required time, taking into consideration that a block might be requested by more than one train at the same time due to overlapping routes. The selection strategy employed in this case can be of arbitrary nature (e.g. at random or using priorities) as long as the basic requirement of one train per block is ensured.

However, a consequence of signalled fixed block operation on single-track routes is the possibility of deadlocks between trains, i.e. a situation where no train can progress on the track due to circular dependencies between block requests (e.g. two opposing trains need to enter the same block at the same time), as is also described in [17]. Deadlocks are rare occurrences in traditional railway operation, as the timetable of the trains can usually be pre-checked or at least adapted to prevent circular dependencies between block requests. In case of on-demand train operation with stochastic demands, however, deadlocks occur with much higher probability due to the possibility (and necessity) for extended train movements in comparison to traditional trains. Furthermore, no pre-check of train movements is possible, as no timetable exists. This makes it necessary to define an extended scheduling model that guarantees freedom from deadlocks in addition to the freedom from collisions provided by traditional signalled fixed block operation.

Note that this specifically refers to the major “train-train collision” hazard; for other hazards, other considerations and countermeasures have been investigated, using an analysis method for critical SoS situations.

4.2 Alone in the Dark

Deadlocks are primarily an availability problem, since when a deadlock situation occurs, no more train movements are possible, but also no (more) collisions can happen in the area affected by the deadlock. However, a combination of unfavourable conditions can nevertheless lead to safety problems even during a deadlock. Solutions for avoiding deadlocks are therefore also relevant for safety.

For instance, the situation when travelling by AuSoDoTS at night can become critical when a deadlock occurs and the system comes to a standstill; the railcar will most probably come to a stop somewhere in the middle of a track section, and additionally, since AuSoDoTS is a GoA4 concept, there is no driver or train attendant available. Considering the human factor in this situation, this can produce effects like fear, panic, aggression, or illness; additional adverse factors like fire or smoke (e.g. from battery or motor) can amplify the human reaction in such a situation, leading to possible safety risks and hazards.

5 General Hazard Classification for AuSoDoTS

Since the AuSoDoTS concept represents a novel approach different from traditional railway or people mover systems, the “traditional” railway hazards need to be re-evaluated. The corresponding investigation of the influences of AuSoDoTS characteristics on these “traditional” railway hazards resulted in the following general classification:

- Hazard largely unaffected (e.g. train-to-train collision, derailment).
  This class of hazards equivalently exist for both traditional railway systems and AuSoDoTS.
- Hazard no longer relevant (e.g. catenary failures, driver mistakes).
  This class of hazards is eliminated via the AuSoDoTS concept.
- Hazard strongly influenced (e.g. collision train-to-traffic/pedestrian/obstacle on track).
  The risks within this class of hazards are amplified by AuSoDoTS.
- New hazard resulting from the unique characteristics of an AuSoDoTS-system (e.g. considering critical SoS situations).
  An identification of new hazards or even new hazard patterns needs to be conducted, in order to be able to derive countermeasures for implementation.

For example, the hazard "collision of a train with a passenger in the station area" is strongly dependent on human behaviour: A passenger estimates the train behaviour according to certain assumptions (e.g. only one direction of movement). In traditional systems, the driver is essential to avoid accidents in such situations by visually ensuring clearance of the track; this could probably still be handled by some sensor-logic system.

However, the characteristics of AuSoDoTS influence the risk of this hazard significantly; altered or extended movement patterns of AuSoDoTS railcars invalidate traditional assumptions about train behaviour. An example is the following scenario as depicted in Fig. 1 for a station layout where passengers need to cross part of the track via railway crossing in order to reach the station platform:

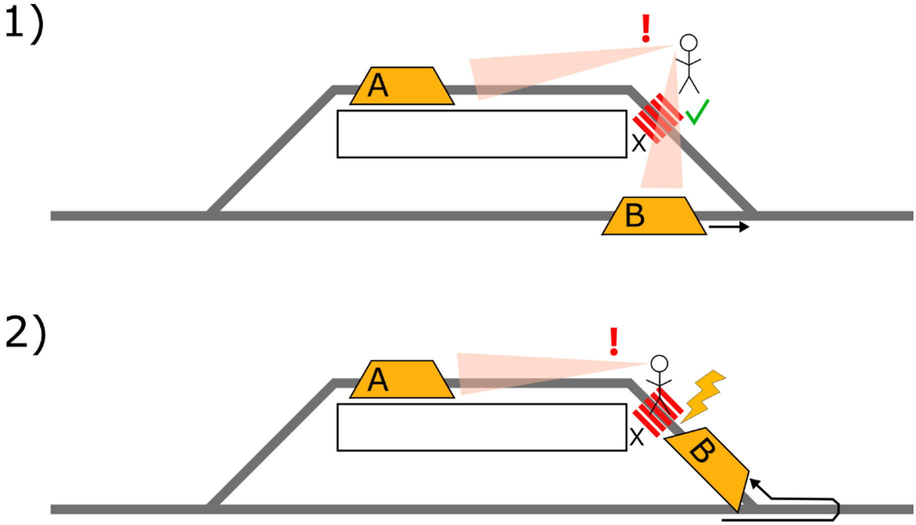


Fig. 1. Example scenario showing hazardous situation resulting from unexpected movement of a railcar

- In scene 1), the passenger notices a railcar (A) ready for departure at the station platform and visually ensures that it is safe to cross the railway crossing (X) by ascertaining the current and probable future positions of other railcars in the vicinity. Despite the proximity of another railcar (B), the passenger decides that the crossing is safe, as railcar (B) appears to pass through the station. The passenger hereby judges based on their experience with traditional railway movement patterns, with which they are familiar, and is also influenced by their motivation to reach the departing railcar (A) in time.
- In scene 2), however, a hazardous situation occurs as railcar (B) actually reverses its course in order to stop at the platform and to allow its passengers to disembark, while the passenger focuses exclusively on the departing railcar (A). It can be clearly seen, that the combination of non-traditional movement by the railcar and lack of attentiveness on the passenger's part due to flawed assumptions on the railcar's behaviour can increase the probability of an accident.

This example shows, that the lack of a driver must be compensated not only by advanced clearance monitoring solutions on the railcar itself, but also via improved train-passenger communication that can tell passengers when behavioural assumptions are wrong, or even hazardous. This hazard example belongs to the third category of the above general hazard classification list.

The hazard given in the previous chapter ("alone in the dark") would fall into the fourth category. Again, the human reaction plays a significant role in dealing with this hazard. Therefore, avoidance of this hazard, and consequently avoidance of deadlocks, is the best countermeasure, as opposed to reactive safety measures.

6 Summary and Outlook

In order to revitalise abandoned secondary lines, and increase attractiveness of still operating ones, a new railway concept is needed. This paper describes such a new concept for people movers deployed on secondary lines, the AuSoDoTS concept, as well as its corresponding research project. The concept follows a "horizontal lift" principle that processes travelling demands occurring in a stochastic fashion instead of following a fixed timetable as in traditional railway operation.

The AuSoDoTS concept is considered as a "SoS", with several systems constituting it, such as the rolling stock, the track itself, the infrastructure, the human actors, and the communication between these systems. This approach is modelled in the project into a concept focussed on safety while deliberately improving availability, e.g. via avoidance of deadlocks.

6.1 Further Work in Preparation

One part of additional consideration is dealing with the different possibilities of balancing safety and availability within the AuSoDoTS concept. In addition, the different implementation possibilities and details of the different systems within the AuSoDoTS concept and their impact are investigated more in depth, e.g. the railcars, different station topologies, and different user interface designs.

There is another paper already in preparation on the set of basic and construction rules for the AuSoDoTS concept; based on this, avoidance of deadlocks (unavailability) within the concept will be discussed. On this basis, ideas on increasing the maximum permissible number of railcars moving in parallel without producing deadlocks will be investigated.

6.2 Open Issues

Security is, although partially considered, not covered in the AuSoDoTS project's definition. However, the importance and necessity to further investigate this field is obvious when looking at transportation of people and communication between system parts.

Additionally, new methods for analysing SoSs become necessary (including tool support) when looking at the complexity of the systems involved, as well as the complexity of their possible interactions and influences from a common environment. Such methods should help to better balance the different needs of safety, security and availability. Furthermore, it has not been considered yet how the use of specific learning algorithms or even artificial intelligence techniques could further help to improve the AuSoDoTS concept.

Finally, the concept is currently intended for passenger transportation only, but its adaptation for small or light freight transport (e.g. single containers) should be explored. This would require additional infrastructure for loading and unloading, different railcars, and even adapted or additional operational rules.

References

1. RRT: Railway Economics and Pricing. In: Railway Reform: Toolkit for Improving Rail Sector Performance, chap. 3, pp. 33–36 (2017)
2. UITP: Metro Automation Facts, Figures and Trends. https://www.uitp.org/sites/default/files/Metro%20automation%20-%20facts%20and%20figures.pdf. Accessed 22 Apr 2020
3. ECDGMT: Study on the Cost and Contribution of the Rail Sector, Final Report September 2015, European Commission Directorate General for Mobility and Transport, pp. 16–17 (2015)
4. TramD: Homepage of the "Tram on Demand" initiative. This initiative deals with traffic congestion in a suburban region south of Vienna. https://tramondemand.wordpress.com/category/konzept/. Accessed 17 Apr 2020
5. ÖBB Infra: "Facts and Figures", Report by the Austrian Railway Infrastructure AG, ÖBB-Infrastruktur, Edition 2020, p. 6. https://infrastruktur.oebb.at/en/company/facts-and-figures/facts-and-figures.pdf. Accessed 18 Apr 2020
6. Leveson, N.: Drawbacks in using the term "system of systems", Technical report, System Safety Research Lab, MIT (2012)
7. IEEE-RS: IEEE-Reliability Society: Technical Committee on 'Systems of Systems'-White Paper (2014)

8. Gebauer, O.: Modell und Entwicklung einer autonom fahrenden Eisenbahn. Doctoral thesis, University of Salzburg (2015)
9. Lechleitner, C., Newesely, G., Zinner, Ch.: Die Straßenbahn lernt Sehen - Innovationen im Bereich Straßen- und Stadtbahnen. In: ZEVrail Sonderheft Tagungsband 42. Tagung Moderne Schienenfahrzeuge, pp. 105–111. Georg Siemens Verlag (2014)
10. Gruber, T., Zinner, C.: Approach for demonstrating safety for a collision avoidance system. In: Koornneef, F., van Gulijk, C. (eds.) SAFECOMP 2015. LNCS, vol. 9338, pp. 167–177. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24249-1_15
11. UITP: Statistics Brief, World Report on Metro Automation 2018. http://metroautomation.org/wp-content/uploads/2019/05/Statistics-Brief-Metro-automation_final_web03.pdf. Accessed 17 Apr 2020
12. Sebron, W., Tschürtz, H., Krebs, P.: Extending the shell model via cause/consequence analysis of component failures. In: Walker, A., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 70–82. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_6
13. Kletz, T.: HAZOP and HAZAN: Identifying and Assessing Process Industry Standards, 3rd edn. Hemisphere, Washington, DC (1992)
14. Society of Automotive Engineers: ARP-4761: Aerospace recommended practice: guidelines and methods for conducting the safety assessment process on civil airborne systems and equipment. 12th edn., SAE, Warrendale (1996)
15. Ulbrich, S., Menzel, T., Reschka, A., Schuldt, F., Mauer, M.: Defining and substantiating the terms scene, situation, and scenario for automated driving. In: IEEE 18th International Conference on Intelligent Transportation Systems (ITSC), Las Palmas de Gran Canaria, Spain (2015)
16. Watzenig, D., Horn, M. (eds.): Automated Driving: Safer and More Efficient Future Driving. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-31895-0
17. Pachl, J.: Deadlock Avoidance in Railroad Operations Simulations. Institute of Railway Systems Engineering and Traffic Safety, Braunschweig (2011)

Experimental Set-up for Evaluation of Algorithms for Simultaneous Localization and Mapping

Stefan Hensel[1], Marin B. Marinov[2(✉)], Christoph Kehret[1], and Maria Stefanova-Pavlova[3]

[1] Department for Electrical Engineering, University of Applied Sciences Offenburg, Badstraße 24, 77652 Offenburg, Germany
stefan.hensel@hs-offenburg.de
[2] Department of Electronics, Technical University of Sofia, 8, Kliment Ohridski Blvd., 1756 Sofia, Bulgaria
mbm@tu-sofia.bg
[3] Center for Innovation and Technology Transfer-Global, Jahroslav Veshin Str. Bl.32/B, 1408 Sofia, Bulgaria
stefanova-pavlova@citt-global.net

Abstract. The precise positioning of mobile systems is a prerequisite for any autonomous behavior, in an industrial environment as well as for field robotics. The paper describes the set up for an experimental platform and its use for the evaluation of simultaneous localization and mapping (SLAM) algorithms. Two approaches are compared. First, a local method based on point cloud matching and integration of inertial measurement units is evaluated. Subsequent matching makes it possible to create a three-dimensional point cloud that can be used as a map in subsequent runs. The second approach is a full SLAM algorithm, based on graph relaxation models, incorporating the full sensor suite of odometry, inertial sensors, and 3D laser scan data.

Keywords: LiDAR · LOAM · Mobile robotics · Mapping · Navigation · ROS · SLAM

1 Introduction and Motivation

Since the beginning of mobile robotics, navigation, and positioning in an unknown environment have been the fundamental challenge for the use of autonomous systems in practical applications. While the first approaches were reserved for military systems and cost-intensive flight applications, the continuous further development of algorithms and sensor systems used now allows for their application in civil use. Certain applications are of paramount importance in the development of these systems. There is a constant increase in detail for environmental monitoring and precise map creation, which is required in inner-city autonomous driving [1, 2]. Besides, the required flexibility in applications of mobile platforms, like surveillance, emergency aid in catastrophe scenarios, and use at home or in hospitals for medical and comfort themes, has vastly increased. In all of these applications, a certain suite of sensor systems and their

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 433–444, 2020.
https://doi.org/10.1007/978-3-030-56441-4_32

combination have proven to be extremely successful [3, 4]. The sensor systems include precise three-dimensional laser scanners (LiDAR), cameras, and inertial navigation systems.

The high information density of optical systems and the required quality for high precision maps are also demanding in theoretical and computational terms for the algorithms used. The family of simultaneous localization and mapping algorithms (SLAM) has proven to be particularly successful for these systems [5].

The main contribution of this work is the design and realization of an experimental platform, which allows for detailed investigation and comparison of different localization and mapping algorithms. The remainder of the paper is organized as follows: Sect. 2 gives an overview of the hardware of the experimental platform. The basis, sensor systems, and computational resources are depicted in detail. Section 3 explains the evaluated SLAM algorithms and the software set-up upon the platform to implement them. In Sect. 4 the experimental evaluation is done and Sect. 5 closes with a summary and conclusions.

2 System Setup

This chapter provides an overview of the platform set up in hard- and software. The robot used as an experimental vehicle, the sensor technology, and necessary adjustments made are explained, as well as the distributed computing components and the software used for communication and control [6]. The evaluated SLAM algorithms and their implementation are dealt with in detail in subsequent chapters.

2.1 Related Systems and State of the Art

Mobile robotic systems can be subdivided into land, air, and water-based. This contribution deals with landbound vehicles. Modern robots make use of an abundance of sensors for mapping and localization. The main problem of most sensors is their restricted use to indoor applications – indoor vacuum cleaners for example can be seen as small robots but are restricted to a very narrow set of environments and sensors.

Good examples of a sophisticated outdoor platform are autonomous cars. They commonly make use of several sensor types and numbers, an overview can be found in [4]. Those systems are highly engineered but due to the needs of security the sensors and software are very expensive and out of scope for more usable applications.

We describe a platform to evaluate modern localization and mapping (SLAM) systems for indoor and outdoor use. This greatly reduces possible mobile bases (minimum wheel diameter for outdoor systems, no small range systems for the LiDARS, no additional ultrasonic for mapping). As most modern logistic systems in industrial scenarios and many field robots do not use Ackermann steering for ease of navigation and robustness of parts, this contribution is also based on a skid-steer platform capable of covering all considered scenarios.

2.2 Mobile Platform Hardware

The Husky mobile robot platform from Clearpath Robotics is used as the sensor carrier. It is a mobile base that can be equipped with a variety of sensors, computers, and batteries for power supply due to its 75 kg payload. Its compact dimensions, wheelbase, given wheel height, and skid-steer drive allow for it to be used both indoors and outdoors, and make it possible to evaluate mapping for structured and natural environments. The maximum speed is 1 m/s, the runtime of one battery charge is 3 h. The unit is controlled by a computer with x64 architecture, equipped with an Intel Core-i5 CPU, 8 GB RAM.

2.3 Additional IT-Infrastructure

Besides the internal platform computer, an additional processing unit with a Core-i5 CPU is built upon the system. We employ a passive unit, which is battery powered by an extra power source. An additional Laptop computer is used for communication, control, and visualization purposes. All computers are connected via a battery-powered communication hub attached to the system. The computer system setup is depicted in Fig. 1.

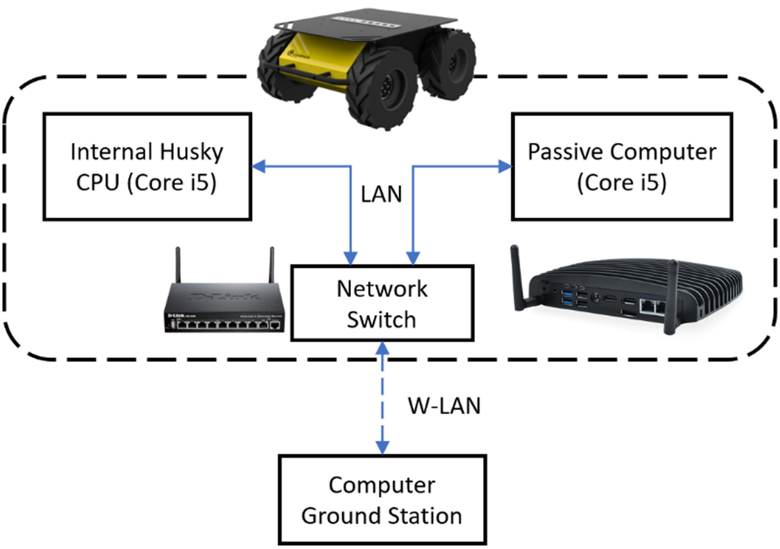


Fig. 1. IT-Setup. Husky platform with an additional processing unit and base station, commonly a Laptop CPU

2.4 Sensor Hardware

Satellite navigation is not suitable for use in industrial buildings, mining, forests, or densely built-up urban areas. As a large number of autonomous applications are situated in these environments, modern mobile systems are equipped with a combination of external sensors for environmental perception and internal sensors for determining the current status. A map created or already existing at runtime is also interpreted as an additional sensor, often called a virtual sensor. The sensor data fusion of all systems finally results in the precise positioning of the system.

External Sensors

At present, laser scanners have prevailed over optical and radar-based systems in the precise and spatially accurate measurement of objects at a high measuring rate [4]. The 3D laser scanners used in demanding tasks are called LiDAR (Light Detection And Ranging).

In the given system the VLP-16 of Velodyne is used. It consists of 16 vertically arranged lasers that rotate at 10 Hz and gain 300 000 laser points per second. The sensor has a range of 100 meters and is mainly used in mobile robotics and autonomous driving due to its compact dimensions [7, 9].

Internal Sensors

While optical sensor systems such as lasers and cameras allow for the determination of the position with respect to the environment, the so-called inertial sensors are used to determine the relative movement of a platform about a reference or inertial system based on occurring forces and moments.

An inertial measurement unit (IMU) consists of three acceleration sensors, three angular rate sensors (gyros), and usually three magnetic field sensors, which are arranged orthogonally to each other and determine the current orientation in space based on the inertia [9].

The combination of the individual sensors allows for the compensation of existing drift, which inevitably occurs due to the integration of the rotation rates and accelerations [10]. The final estimation of the position is performed either using nonlinear state observers, e.g. the Extended Kalman Filter, or the so-called orthogonal filters.

For the Husky robot, the IMU MTi-300 from XSens was chosen. It consists of the described sensor triads and estimates the position with the help of a Kalman filter. The uncertainty of the yaw angle is below 1°, the pitch and roll angles are determined with an accuracy of 0.2°. The advantages of the system are the IP67 protection class and the availability of software interfaces that facilitate integration on the robot platform and allow for outdoor use.

Figure 2 shows an overall view of the platform, enhanced with the described sensor systems and computer hardware.

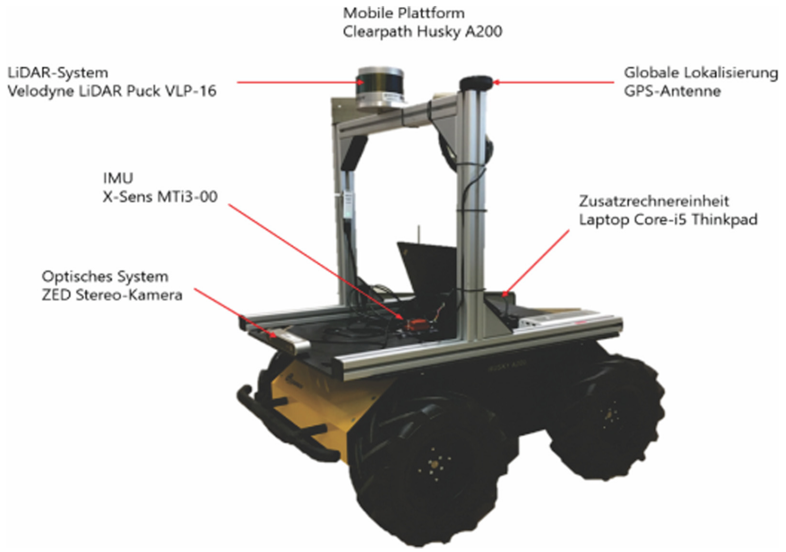


Fig. 2. Experimental platform. Husky enhanced with rigging, sensors, and computers.

2.5 Software System Setup

All computers are running on a Linux Ubuntu operation system. The communication is realized by a local group and TCP/IP protocols.

Middleware – Robot Operation System (ROS)
The communication sensor integration and algorithm implementation are done with the help of the open-source middleware ROS [11]. A middleware is a distribution software, that allows and controls the communication between processes in an abstract sense. In addition to a large number of existing sensor interfaces, ROS offers native program libraries, which include the areas of visualization, data recording, simulation, and system monitoring. Self-made modules can be created in the C++ and Python programming languages or via a software-bridge in Matlab/Simulink.

3 Localization and Mapping

Within the given setup two localization and mapping algorithms were implemented. Both make use of the given sensors differently. For this contribution, we considered several scenarios and applications of the given mapping systems ranging from a field robotic application, autonomous driving, to industrial applications fur automated logistics. As the environments differ, distinct features have to be considered. Two approaches are evaluated. The first one is a local mapping system that infers the traveled trajectory incrementally, whereas the second one is a full SLAM system, which works on global optimization of all given data. The second algorithm is again subdivided into three and two-dimensional maps for the applications. The results are two three dimensional maps with local and global optimization and a two-dimensional map globally optimized. The implemented algorithms are described in the following paragraphs.

Laser Odometry and Mapping – LOAM

Laser Odometry and Mapping is a real-time method capable to estimate the current position using 3D-LiDAR data. The traveled distance between two consecutive laser scans is determined by a point matching method and the current position is calculated from the integration from the starting point. The algorithm works in two steps and consists of fast local position estimation and a slower correction step, which creates a point cloud from the recorded measurements using block bundle adjustment. The point cloud acts as a map, which in turn can be used as a reference system for the renewed position determination. This parallel structure allows for a high calculation speed with high precision. Detailed information can be found in [10] and [12].

The system explicitly does not use any additional sensors, so that the method belongs to the class of scan-matching based SLAM methods. It was selected for evaluation because of its low computational load, which allows real-time capability on commercially available CPUs. Besides, variants of the algorithm can be found in the top ranks of neutral performance benchmarks, e.g. the KITTI database [13].

Cartographer SLAM

Methods of simultaneous localization and mapping have a long history in robotics [3] and can be roughly divided into two relevant classes: Probabilistic methods based on state estimators (Extended Kalman Filter SLAM) and methods that generate a graph from the existing poses (position and orientation) and sensor observations of the environment and optimize it (Pose Graph SLAM). Cartographer is a pose graph SLAM method based on a combination of LiDAR and inertial sensors [14]. Wheel odometry can also be integrated but is too inaccurate due to the skid-steer steering used by Husky. Cartographer SLAM can process both 3D and 2D data. The latter uses single LiDAR beams and directly outputs a two-dimensional probability grid map, the 3D data is converted into the plane by a projection step.

An extension to the classical SLAM methods is the subdivision into a local real-time SLAM process, which is linked to a global optimization step [15]. The global optimization is based on a block bundle adjustment process that is activated as soon as the robot arrives again at a position that has already been passed. This optimization of sparse matrices is called loop closing and is based on photogrammetric algorithms. It corrects the orientation of the origin points of the local maps, which in turn affects the estimated position of the robot via these maps.

System Architecture

The system components described are connected via a distributed computer architecture as shown in Fig. 1. The communication of the individual components is handled by ROS topics and services. The Husky's control computer receives the controller data and passes it on to the robot's actuators via a controller. The passive PC receives the sensor data and estimates the position and map with the respective SLAM algorithms. The robot is controlled manually with the remote PC and a gamepad. This means no autonomous exploration is required, and the data can be acquired in a systemized way to only evaluate the mapping and localization process.

ROS is used to record all positioning commands and sensor data during the movement phase. In a subsequent step, the data is processed offline with the respective

SLAM-algorithm, and trajectory and map are calculated based on identical sensor data for each method, which allows for the comparison of the methods.

4 Data Acquisition and Experimental Results

The evaluation of the SLAM-Algorithms was done qualitatively and quantitatively. The most important qualitative difference is the type of maps produced. While LOAM outputs a three-dimensional point cloud, Cartographer SLAM generates a two-dimensional occupancy grid map, which represents free spaces, obstacles, and unknown areas based on conditional probabilities. Both maps can be easily interpreted by humans, whereas the point cloud appears aesthetically more sophisticated. The grid map on the other hand can be used directly for subsequent navigation and path planning by the mobile system.

Further criteria were computational complexity: Although both algorithms run on a passive mini PC in real-time, growing scale makes the use of cartographer SLAM more prohibitive, due to the memory need for loop closures. LOAM shows an even lower computational burden, as it was specially designed with low processing power in mind. The additional lack of a loop closure procedure further reduces calculations and the use of memory. Another important quality feature is the robustness concerning environmental influences: The capability to incorporate low-quality measurements and to deal with outliers, e.g. from false reflections of the LiDAR. Required and used sensor technology is the last criterion, assessing the costs and complexity of incorporating additional readings.

We will call the use of three-dimensional LiDAR data and algorithms LOAM 3D and Cartographer 3D (i.e. three-dimensional data is used and a corresponding map created; the two-dimensional map can be created by a projection of the data), the SLAM approach, that produces directly two-dimensional occupancy grid maps is called Cartographer 2D.

A summary of the different qualitative criteria and properties is presented in Table 1.

No reference measuring system is available for the quantitative evaluation so that the difference between a known start and endpoint of the journey was selected as a quality criterion.

Table 1. Qualitative criteria for SLAM algorithms

Criteria	LOAM 3D SLAM	Cartographer 2D SLAM	Cartographer 3D SLAM
Map type	3D point cloud	2D occupancy grid	2D occupancy grid
Computational complexity	**Low**	High	Very high
Robustness	**Very high**	**Very high**	High
Sensor system complexity	LiDAR	LiDAR, IMU, odometry	LiDAR, IMU
Subsequent map usability	No	**Yes**	**Yes**

First, typical scenarios for the application of mobile robot systems were selected: The drive within the university building reflects the use in office buildings or support in hospitals and shows a variety of repetitive characteristics. Scan matching algorithms and feature-based optical approaches clearly have problems in dealing with these kinds of scenarios. The university laboratories for machine tools and automation can be used as representative for an industry 4.0 scenario. Here, the mobile platform has narrow pathways but a lot of unique features that can be used to localize and build up maps. One error scenario here is the change of larger map areas, e.g. raw material being used for production, which violates the implicit assumption of both systems for having a static map. Last but not least, a drive was made in the outdoor area, which offers only a few landmarks for 2D systems and simulates the use in the unstructured environment of field robotics. These scenarios are typical for agricultural applications and scenarios like reconnaissance, help in catastrophe theatres or mining operations. In contrast to industrial indoor facilities, only a few features can be used to match sensor readings and relocalize. Tree stems and plants are sparse and ambiguous if only scanned in 2D, whereas inertial systems and three-dimensional sensors clearly are advantageous. In each scenario journeys with different lengths were carried out. Besides, the return or multiple passages of the same points of interest were carried out occasionally. These journeys are additionally called loop, because Cartographer SLAM can use loop closing, but not the LOAM method. The quantitative results are shown in Table 2. The relative deviation of start and endpoint in X and Y direction of the map is shown.

Table 2. Quantitative evaluation

Scenario	Traveled distance, m	LOAM 3D SLAM		Cartographer 2D SLAM		Cartographer 3D	
		Error X, %	Error Y, %	Error X, %	Error Y, %	Error X, %	Error Y, %
Indoor short	61	0,08	0,43	**0,03**	**0,02**	0,54	0,13
Indoor long	475	**0,18**	**0,24**	0,48	0,18	**0,23**	**0,14**
Indoor long loop	500	0,07	0,27	0,41	0,06	**0,15**	**0,2**
Industrial short loop	50	1,22	0,14	**0,6**	**0,04**	0,24	0,01
Industrial middle 2 loop	115	0,24	0,04	**0,18**	**0,02**	**0,01**	**0,16**
Outdoor short	75	**0,31**	**0,11**	0,64	0,65	**0,11**	**0,21**
Outdoor long	*500*	**0,03**	**0,1**	0,24	0,14	0,18	0,33

Example Mapping Results

Example results are shown in Fig. 3 and Fig. 4 for the industrial and outdoor scenario. An example image, the generated LOAM point cloud map, and the raster map generated in Cartographer-2D are shown.

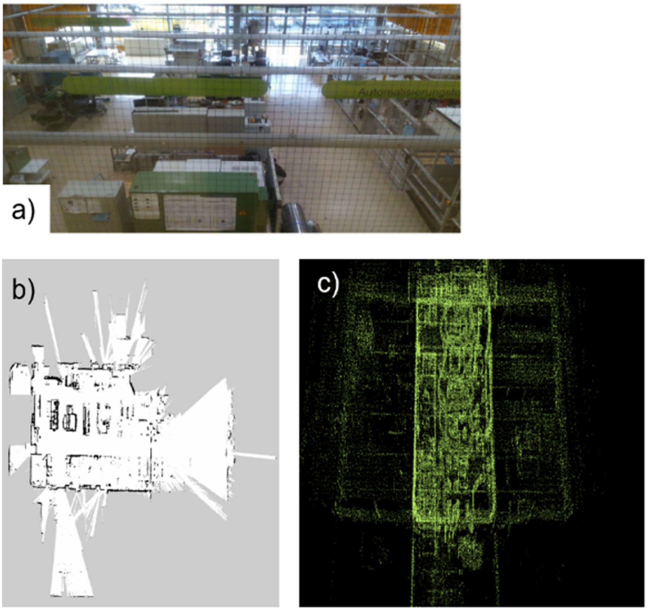


Fig. 3. Example result of the evaluation of the Industry 4.0 scenario with a) video image, b) Cartographer 2D map and c) LOAM 3D point cloud map.

The results indicate that Cartographer SLAM is slightly superior to the LOAM method in terms of accuracy. This is mainly due to the capability of adding the loop closure when revisiting areas. As a result, all cartographer deviations are bounded, independent of scenario and length, as long as a loop is closed. This can be seen in Table 2 on all middle measurements indicated by loop. The advantage of the LOAM method, on the other hand, is the high processing speed and the possibility to obtain a relative localization easily. Even without loop closures, LOAM has a highly efficient point matching procedure, that keeps the error at very low values. All three algorithms show excellent and comparable results, whereas the LOAM approach often slightly outperforms the more sophisticated Cartographer localization.

In unstructured areas with few landmarks, the 3D methods establish themselves as methods of choice in comparison to 2D methods. Laser scan lines have not sufficient information to compensate for drift in IMU, the larger amount of LiDAR points is therefore a major improvement. A good example can be seen in Fig. 4, where the stems of trees are online short line segments, whereas the LOAM map completely covers the outdoor façade and treetops to increase point matching accuracy.

The systematically greatest advantage of Cartographer SLAM is the creation of the raster map, which serves as a direct input for the later robot navigation and where established methods can be used. The 3D point cloud of the LOAM method, on the other hand, provides a visually appealing representation of the navigated area, which, however, has to be saved for later localization. Navigation is also made more difficult for moving systems due to the lack of indicators such as drivable terrain but is sufficient as obstacle detection for flying robots such as drones.

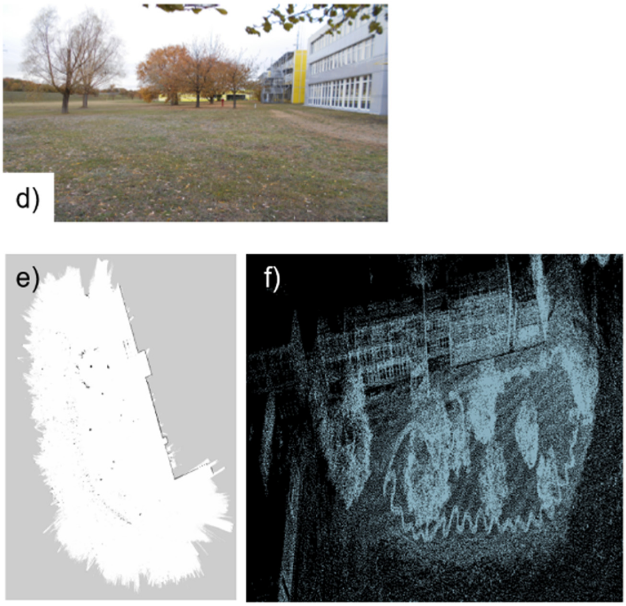


Fig. 4. Example result of the evaluation of the outdoor scenario with d) video image, e) Cartographer 2D map and f) LOAM 3D point cloud map.

5 Conclusion

In this paper, a system for the implementation and evaluation of complex localization and mapping procedures was presented. The methods considered are fundamentally suitable for robust localization in demanding scenarios, such as logistics applications in industry 4.0, exploration, and monitoring in field robotics or applications in service robotics. Global SLAM approaches with optimization by loop-closing show themselves to be systemically slightly superior in the application, since global optimization is performed and all information from a large number of sensors can be combined and used as desired.

The most recent interpretation of the SPI term is System, Software, Services, Safety, and Security Process and Product Improvement, Innovation, and Infrastructure [16]. EuroAsiaSPI has developed as experience and knowledge exchange platform for Europe and worldwide, where Software Process Improvement (SPI) practices and knowledge can be gathered and shared. The connected SPI Manifesto [17] defines the required values and principles for most efficient SPI research and industry projects and is a Strategy Paper for further development on the base of those principles. One main goal is to support changes by innovation and include all stakeholders. There are two important values, which support the whole development process namely "SPI is inherently linked with change" means that change is part of the normal development process which leads to new business opportunities [18]. A current major change is the precise positioning of mobile systems as a prerequisite for any autonomous behavior, in an industrial environment as well as for field robotics. The presented paper describes the set up for an experimental platform and its use for the evaluation of simultaneous localization and mapping (SLAM) algorithms. “Create a learning organization” is an important principle for the exchange of best practices on the researched topics [19].

Acknowledgments. This paper has been produced within the framework of the ERASMUS + project Geothermal & Solar Skills Vocational Education and Training (GSS-VET).

References

1. Corke, P.: Robotics, Vision and Control, 2nd edn. Springer, Edition (2017). https://doi.org/10.1007/978-3-642-20144-8
2. Yagfarov, R., Ivanov, M., Afanasyev, I.: Map comparison of lidar-based 2D SLAM algorithms using precise ground truth. In: 2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV), Singapore (2018)
3. Cadena, C., et al.: Past, present, and future of simultaneous localization and mapping: towards the robust-perception age, vol. 32, no. 6, pp. 1309–1332, December 2016
4. Hensel, S., Marinov, M., Kehret, C.: Design of a mobile platform for the evaluation of localization and mapping algorithms. In: Proceedings of the 8th FDIBA Conference - Technologies and Education for a Smart World, Sofia, 30 November 2018
5. Thrun, S., Burgard, W., Fox, D.: Probabilistic Robotics. MIT Press, Cambridge (2005). ISBN 978-0-262-20162-9
6. Hensel, S., Marinov, M.B., Schmitt, M.: Experimental setup for investigation and evaluation of a mapping and localization system. In: Proceedings of the 9th FDIBA Conference - Challenges of the Digital World, Sofia, Bulgaria, 28–29 November 2019
7. Bergelt, R., Khan, O., Hardt, W.: Improving the intrinsic calibration of a Velodyne LiDAR sensor. In: 2017 IEEE SENSORS, Glasgow (2017)
8. Nüchter, A.: 3D Robotic Mapping - The SLAM Problem with Six Degrees of Freedom. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-540-89884-9
9. Wendel, J.: Integrierte Navigationssysteme: Sensordatenfusion, GPS und Inertiale Navigation (2009)
10. Stachniss, C., Leonard, J.J., Thrun, S.: Simultaneous localization and mapping. In: Siciliano, B., Khatib, O. (eds.) Springer Handbook of Robotics, pp. 1153–1176. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-32552-1_46
11. Quigley, M., et al.: ROS: an open-source robot operating system. In: ICRA Workshop on Open Source Software (2009)
12. Pomerleau, F., Colas, F., Siegwart, R.: A review of point cloud registration algorithms for mobile robotics. Found. Trends Robot. **4**(1), 1–104 (2015)
13. Geiger, A., Lenz, P., Urtasun, R.: Are we ready for autonomous driving? The Kitti vision benchmark suite. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2011)
14. Zhang, J., Singh, S.: LOAM: lidar odometry and mapping in real-time. In: Robotics Science and Systems (2014)
15. Hess, W., Kohler, D., Rapp, H., Andor, D.: Real-time loop closure in 2D LIDAR SLAM. In: Robotics and Automation (ICRA) (2016)
16. Manifesto. http://2018.eurospi.net/index.php/manifesto. Accessed 2 Apr 2020

17. The SPI Manifesto, Alcala, EuroSPI 2009. http://2019.eurospi.net/images/eurospi/spi_manifesto.pdf
18. Messnarz, R., Ekert, D., Grunert, F., Blume, A.: Cross-cutting approach to integrate functional and material design in a system architectural design – example of an electric powertrain. In: Walker, A., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 322–338. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_25
19. Korsaa, M., et al.: The people aspects in modern process improvement management approaches. J. Softw. Evol. Process **25**(4), 381–391 (2013)

Good and Bad Practices in Improvement

A Cognitive Map of Maturity: Symptom-Based Improvement Recommendations

Jan Pries-Heje[1], Jørn Johansen[2(✉)], and Morten Korsaa[2]

[1] Department of People and Technology, Roskilde University, Roskilde, Denmark
janph@ruc.dk
[2] Whitebox, Hørsholm, Denmark
{jj,mk}@whitebox.dk

Abstract. When an experienced assessor enters a company there are certain characteristic symptoms that reveals the maturity of the company even before the assessment. In this paper we start out from a list of 32 characteristic symptomatic problems generated by two experienced assessors after maturity assessments in more than 300 companies. We then use cognitive mapping asking five times why to get behind the symptomatic problems and reveal the underlying problems or causes of the problems. Our mapping revealed a number of interesting relationships between the symptomatic problems. We then evaluated our findings by building a tool in the form of a website where users could score statements - formulated from the symptoms - and from that we could point them to areas where they probably needed to improve. We improved the tool in three learning cycles of design evaluation and ended up in a summative evaluation where we compared the outcome of using the website tool with a CMMI maturity assessment. We conclude that looking for symptoms, and as an individual scoring statements build on the symptoms, can point to improvement areas. However, doing so is not a replacement for a maturity assessment, the scoring of statements cannot necessarily reveal the maturity of the organization.

Keywords: Cognitive map · Process improvement · Maturity · Improvement · CMMI

1 Introduction

Imagine that you are a manager in a product development company or department, in distress over some bad performance. Somewhere. It is your responsibility, but where is the real problem? Where should you direct your attention? What should you do?

Start looking, and you will find an overwhelming amount of processes, problems, causes of problems, symptoms of problems, tools, practices, people, organisational structures. Each of them could potentially be the problem. And they seem to relate to each other in many ways.

The authors of this paper are looking for a way to help you out. From our experience, there is a relatively limited number of symptoms that is of real importance in most organisations. There are also a limited number of relevant relationships between

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 447–461, 2020.
https://doi.org/10.1007/978-3-030-56441-4_33

the symptoms, the experienced problems, and the causes of the problems. In this paper we start our work from a list of 32 symptoms that we have meet when we have been doing assessments in more than 300 different product development organisations. We take each symptom and ask two things; (1) What are the causes of this symptom? – the problem behind so to say; (2) What does it cause or lead to or has as effect?

An example could be that the documentation for a product is not updated. That could be caused by lack of time for making documentation, and that again could be caused by a very tight development schedule or budget for product development. Further, when documentation is not updated it could lead to difficulties in maintaining and further developing the product, and that again could lead to the problem that it becomes excessively costly to maintain and further develop the product.

The example above has five levels of problems; a kind of problem-hierarchy that is. An old Japanese improvement technique called "5 times Why" operates exactly with five levels. The point being that you should never go with the first symptom of a problem but instead look for the root cause(s). The Japanese car company Toyota developed the "5 times Why" technique in the 1930s. It became popular in the 1970s, and Toyota still uses it to solve problems today [1].

The research question we aim at answering in this paper is: "How can we use a mapping of the symptomatic problems, causes and effects, and the relationships identified between them, to design a tool that can help determine what the most urgent improvement areas are in your company?"

Our approach for answering that research question is to use the "5 times Why" techniques on the 32 symptoms that we regularly have met in companies.

In this work we have used the CMMI [2] maturity model as our basis for discussions because two of the authors are very experienced in using this model. However, it could have equally relevant to use SPICE or Automotive SPICE [3] and the addressed practice capabilities in these models.

2 Existing Research on Problems

A problem can be defined as a perceived difference between what is and what ought to be [4]. There can be many aspects of a problem [5]. It can be a consciousness of a gap, a desire, or a need. On the other hand, an aspect can also be that something is undesirable and therefor implies the imperative for change. A third aspect can be that it is difficult as opposed to trivial. A fourth aspect that it is solvable as opposed to impossible to solve. And finally, there can be different perceptions by different stakeholders. One may see it as undesirable not to have updated documentation whereas another stakeholder may see it as no problem at all.

In the literature you find many problem analyses and problem-solving techniques. A well-known method by Peter Checkland [6–8] is called Soft Systems Methodology (SSM). Checkland had experience as a consultant for big international companies like Shell before coming back to academia as a professor in systems thinking. He distinguishes between hard and soft problems. A hard problem is a well-formed problem that can be solved with well-known engineering techniques. The problem presents itself so that it is easy to see what type of problem it is. A soft problem on the other hand may have many aspects, many humans involved, and many different stakeholder perspectives. Thus, it needs work and discussion to understand the problem – if it is a problem!

Checkland has two important points that we will use here. First, it pays off to distinguish between the real world and systems thinking at a meta- level about the real world. Second, it pays off to produce models of purposeful activity in the real situation and use the models as devices to explore the situations and structure a discussion.

Rittel and Webber [9] defined some problems as being 'wicked' in that you cannot solve the problem unless you have some knowledge that you can only get from solving the problem; the problem cannot be understood until after the formulation of a solution. Hence, the only viable strategy is to start solving the problem and learning in the process. Further, wicked problems can be considered to be a symptom of another problem. They are linked together.

Many years later Snowden and Boone [10] presented a framework for sense making to be used by leaders' for decision-making where they look at the relationship between cause and effect. If the relationship between cause and effect is known, it is a 'simple problem' that we can use best practice for. If the relationship is potentially knowable – through hard work by experts – then it is a complicated problem. And if the problem is wicked and thus the relationship between cause and effect only retrospectively coherent then it is a complex problem.

Looking at the relationships between problems and the causal relationships Colin Eden [11] came up with cognitive maps of problems or constructs as he calls them where the link between two problem constructs is in the form of an arrow to show the nature of the linkage; "an arrow out of a construct shows a consequence and an arrow into a construct an explanation" [11, p. 5]. Some years later Eden and Ackerman [12] developed cognitive or causal maps into a techniques [13] that could be used for making strategy. Finally, Venable [14] refined cognitive maps into coloured cognitive maps that can be used for creating a design of a solution. His idea was that each problem should be formulated with its opposing node. E.g. 'high employee turnover' has the opposing node 'low employee turnover'. And when you switch around a whole cognitive map – from the original nodes to the opposing nodes - you will end up with a potential design solution.

3 Research Method

To answer our research question, what are the most urgent improvement areas in your company?, we decided to apply DSR - Design Science Research [15]. DSR is a research approach where you build something and then learn from it (when evaluating). Thus, in order to answer our research question, we decided to build cognitive maps showing the linkages and relationships between problems and symptoms, causes and effects, to better understand what the most urgent improvement areas in a company may be. And we decided not to do it for a specific company but instead do it at a systems-oriented meta-level.

A main reason for choosing DSR as our research methodology is that it combines the need for practical relevance and utility. DSR emphasizes that a design should address a need or a problem and at the same time should "stand on the shoulders" of existing research within the problem area [15]. Besides having a 'relevance iteration cycle' where you start by identifying a need or a problem you also have a 'rigor iteration cycle' where you identify all relevant academic literature; what do we actually know by now? The artefact that you are building in order to learn can be a product artefact or a process [16].

Hence, we developed causal cognitive maps. We started out from symptoms that can be seen in companies and asked what can this cause? (= Consequences), and what is causing this? As said above we decided to use the "5 times Why" techniques so we developed each symptom in five levels typically by starting with the symptom in the middle of the map and then eliciting two levels of consequences and two levels of causes.

What data should we use for eliciting the maps? Here we took advantage of the more than 600 assessments in more than 300 companies that the group of authors together have carried out. That has given us extensive knowledge of how things are related. So, we simply used the cognitive mapping techniques to make explicit what was in our cognitive mind. At first, we split the symptoms in three groups and mapped a group each. Then to avoid bias and give some inter-coder reliability we swapped the maps around among us until all three authors agreed of the linkages and relationships. We decided only to represent the most important linkage(s), one, two or three. So, a symptom can be caused by many things, but we decided to prioritize the causes and only represent the most important ones. In doing so we are following the principle of organizational learning from the SPI Manifesto [17].

This paper primarily supports the following principles in the SPI Manifesto: Create a learning organization; Support the organization's vision and business objectives; Use dynamic and adaptable models as needed – in the sense to bring insight for the principles.

4 Symptoms Observed

In this paper we start out from a list of 32 characteristic symptoms generated by two of the authors of this paper being experienced assessors after maturity assessments in more than 300 companies. They were acknowledged in the following way.

Over the last 25 years as many as 600 assessments are performed in over 300 companies. During this work, the assessors got more and more trained to collect signals related to how "clever" the company are to develop new products or deliver projects for customers. It came so far, they the assessors started to guess on the maturity after looking at some of the main documents, development model and be welcomed in the reception.

Last year the assessors started to identify the most common symptoms. After several brainstorms and discussions, it ended in 32 symptoms, which was formulated as statements of the symptoms. Examples of symptoms were:

- We cannot tell how much effort an individual has used on an activity (#1)
- We do not know who and how many that have a say in the project and the results of the project (#8)
- Unfortunately, we do not find defects until the product is in operation with the customer or end-user (#13)
- We are often correcting the same mistake again and again (#19)
- Employees experience "bad" deliveries from colleagues (#28)

The 32 symptoms were grouped into 5 natural categories, based on the type of the symptoms. The five examples above are one from each of the following five categories:

1. To be in control of the projects across the organization
2. Knowing what do develop and deliver?
3. Projects having the adequate competences to run projects
4. Having project insight and status
5. Quality in work and work products

Having these symptoms defined, we had to qualify them and find a way to make this operational.

5 Cognitive Maps

To obtain a better understanding of the symptoms we decided to apply cognitive maps. We started out from a symptom. Then we asked, what is causing this? And then we asked again; what is the cause of the cause?, thereby identifying the underlying problem. Further we asked, what is the effect or result of the symptom? And then; what is the effect of the effect? For some of the problems we could probably have continued further back to an even more underlying problem or further forward to an effect of an effect of an effect. However, we had decided to apply the "5 times Why" heuristic so we ended our mapping of each symptom with a map that had five layers.

We did this mapping for all 32 symptoms. Furthermore, we circulated the maps among the three authors thereby neutralising any bias that any one of us may have had. We also had some discussions about certain problems; is this a cause or an effect? That was not always easy to answer.

An interesting observation that we made while going through the symptoms one by one was that some (causal) problems or effects started to reappear. We noted that and discuss it later in this paper.

In Fig. 1 you find an example where we have mapped symptom #6: "We have difficulties correcting defects in something delivered".

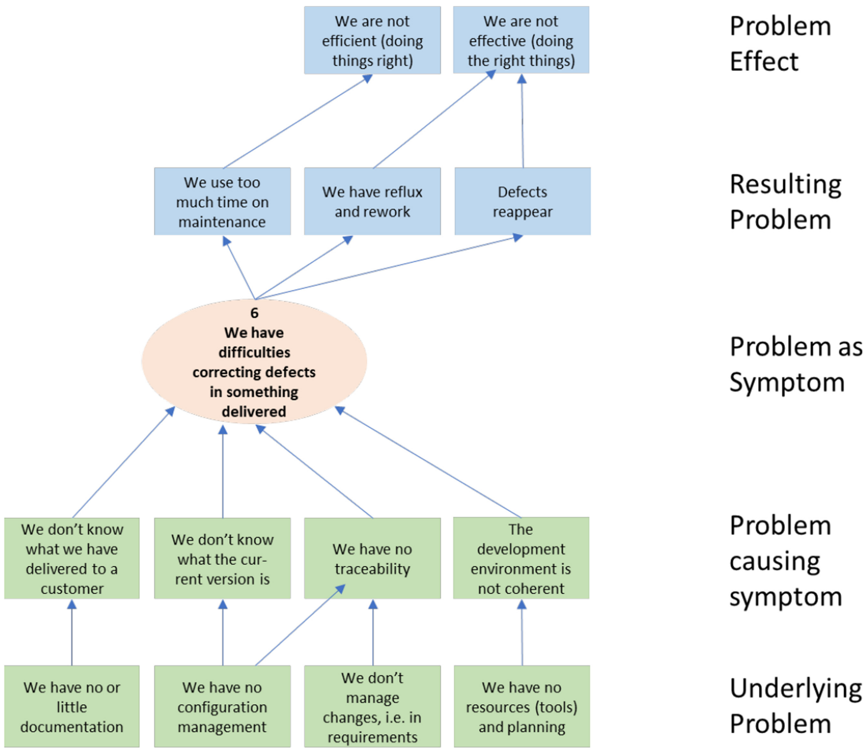


Fig. 1. An example of a cognitive map – here of symptom #6. As you can see the map has five levels of problems corresponding to the "5 times Why" heuristic that we have used.

One "causality track" that we find in Fig. 1 is the following: We have no configuration management => We don't know what the current version is => We have difficulties correcting defects in something delivered => We use too much time on maintenance => We are not efficient (doing things right).

Another causality track found in Fig. 1 is: We don't manage changes, i.e. in requirements => We have no traceability => We have difficulties correcting defects in something delivered => We have reflux and rework => We are not effective (doing the right things).

When we were eliciting and creating the cognitive maps, we also found that some symptoms are related to other symptoms. That was easily seen when two symptoms resulted in (or caused) the same effect or when two symptoms were caused by the same underlying problem(s). In Fig. 2 we have shown the same symptom #6 as we presented in Fig. 1 – But now with two closely related symptoms represented as well.

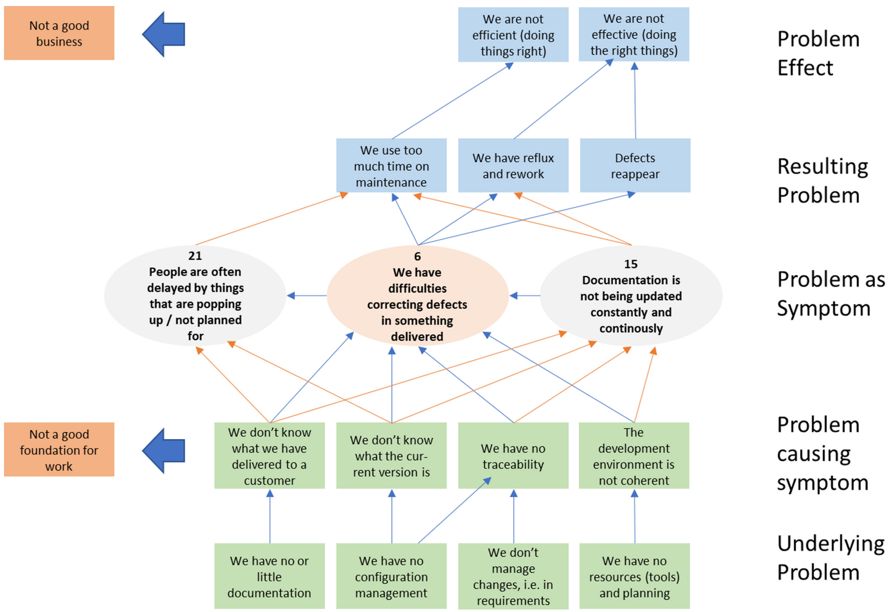


Fig. 2. A cognitive map of symptom #6 and the relationships to symptom #15 and symptom #21. Furthermore, two business-oriented effects of the whole is shown to the left

Thus, Fig. 2 shows that symptom #15 "Documentation is not being updated constantly and continuously" and symptom #21 "People are often delayed by things that are popping up/not planned for" are closely related to symptom #6.

Another thing we identified was that there were problems at different levels. We have mapped things at the most concrete level – the problem instantiations. However, there is also an effect of the whole map or network of problems. For example, the two effects – not doing the right things and not doing things right – together will cause the business as a whole to be bad. And the four problems that together are causing the three symptoms in Fig. 2 will altogether form a not so good foundation for work. In Fig. 2 we have mapped these effects related to the 'whole' map in orange to the left of the cognitive map.

Let us take another example. Let us look at symptom #15. See Fig. 3.

For the 15th symptom; "Documentation is not being updated constantly and continuously", in Fig. 3, we find again that the symptom leads to rework - but starting from different underlying problems. And again looking at the meta-level or the 'whole' we can conclude that the overall effect is that it will lead to bad business and that the level causing the symptom(s) as a whole will make it nearly impossible to maintain the product.

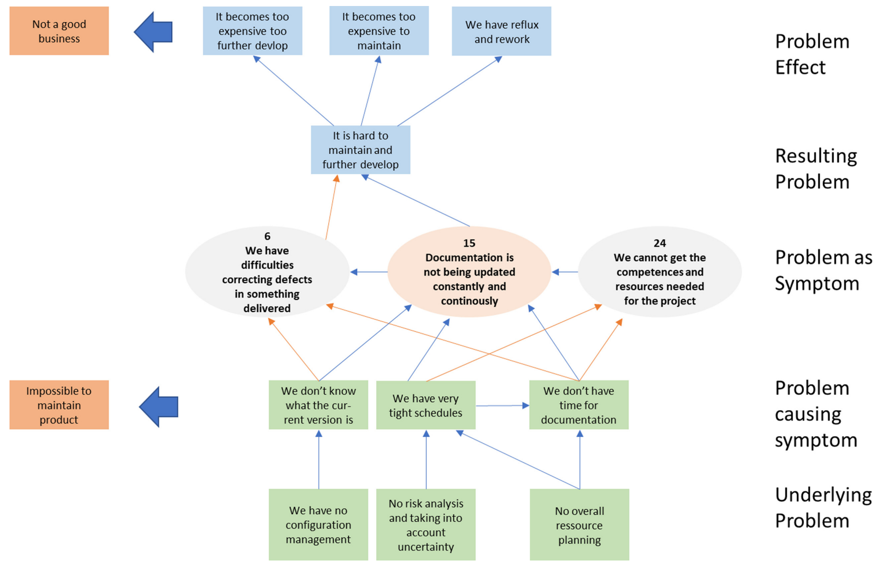


Fig. 3. Example of a cognitive map – here of symptom #15. And how it is related to symptom #6 and symptom #24.

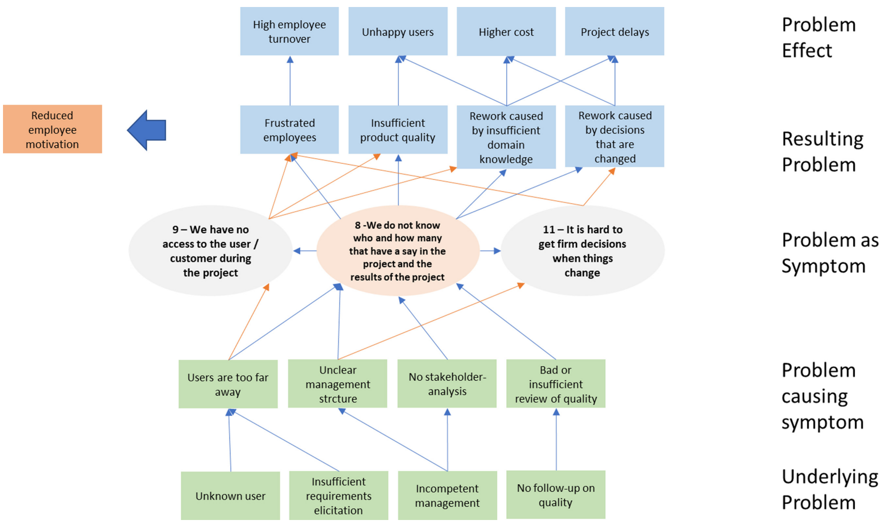


Fig. 4. An example of a cognitive map – here: Symptom #8

Another interesting observation in Fig. 3 is that a problem can cause another problem directly or through another – more indirectly so to say. An example "We have very tight schedules". That can lead directly to the symptom #15 but it can also lead to "We don't have time for documentation" that then again can lead to symptom #15 "Documentation is not being updated constantly and continuously".

In the next example shown in Fig. 4 it is even more obvious that the higher-level problems can be the same as some of the lower-level underlying problems.

Furthermore, for symptom #8 there is again a relationship to two other symptoms, namely #9 "We have no access to the user/customer during the project", and to #11 "It is hard to get firm decisions when things change". Finally, the 'whole' of Fig. 4 will lead to reduced employee motivation.

Let us now take a closer look at Symptom #9. See Fig. 5

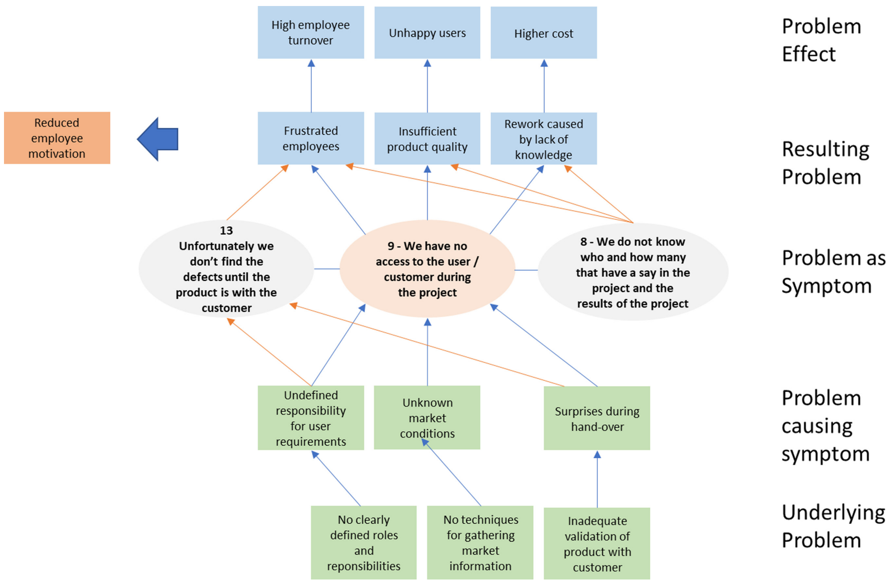


Fig. 5. An example of a cognitive map – here: Symptom #9

When one compares Fig. 4 and Fig. 5, we find that the layers above the symptoms are nearly the same for Symptom #8 and Symptom #9. Whereas the layers below are quite different for the two symptoms.

So to conclude this mapping section of the paper we have found that it was very useful to use cognitive maps to obtain an overview of the symptoms and the relationships either directly or through other symptoms.

6 Evaluation

An important part of using DSR as the research method is that you have to evaluate the design [18]. Typically, you start by having one or more formative evaluations where the result is used to 'form' the artefact. And you at the end are having a summative evaluation that 'sums up' or conclude your research. In the beginning the first formative evaluations may be in artificial settings but later you aim at having the real users in the real context with the real problem – real, real, real – a so-called naturalistic evaluation [18].

We have used a framework for evaluation called FEDS by John Venable et al. [18] to plan and carry out our evaluation. For our first formative evaluation we developed a tool that we made available on a website. The tool presents a person in an organization for each of the 32 symptoms to "score" it, seen for the persons perspective (Fig. 6).

Nr	Assertions / Statement:	Totally agree	Partly agree	Neither nor	Partly disagree	Totally disagree
	In control					
9	We have no access to the user / customer during the project					

Fig. 6. An example of "scoring" a symptom – here of symptom #9

We asked 9 persons from different companies to "score" all 32 symptoms and give their opinion about how easy it was to understand the symptom statements.

After having evaluated the website with 9 users we analysed the outcome and changed some of our statements and the build-in relationship between problems. The tool was updated, so a mail with the recommendations automatically was returned, when the "scoring" was finished. The website was then again tested also with 9 users. This time the focus was to analyse the algorithm in the tool to strengthen the recommendations. And again, this led to changes in and some small changes in the formulations of the statements.

Finally as our summative evaluation we decided to evaluate the web-site tool up against a classic CMMI [2] assessment of maturity. The two assessors fill out the questionnaire separately after a performed CMMI assessment to check if they had the same answer on the questions (Table 1).

Table 1. Differences in the score of one assessed company

Same score	48%
1 next to (25% disagree)	37%
2 next to (50% disagree)	16%
3 next to (75% disagree)	0%
4 next to (100% disagree)	0%

During the exercise and following discussions and analysis, it became clear, that it is not possible to derive a maturity level for an organisation from the symptoms alone – but only pinpoint the main weaknesses important to address.

We started to identify which overall problem related capabilities the symptoms addressed in an organization showing the ability to run projects successfully and ended up with ten important themes: (1) Project management; (2) Control across the organization; (3) Proactive management; (4) Clear goals for the projects; (5) Capable development organisation; (6) Quality of results and products; (7) Control of work products and product parts; (8) Management insight and involvement; (9) Tools and support; (10) Continuous improvement and learning. The algorithm was adjusted to calculate a score for the 10 themes as basis for recommendations: What is most important to focus on and to improve.

There is obvious links between the ten defined main themes related to the overall ability to successfully run projects and the processes. This link can be used to pinpoint the recommendations, but not the maturity level. The reason is, that the themes each are based on several processes. But it seems to be clear which themes are aggregated from which symptoms. This exercise was used to strengthen the questions.

Figure 7 shows the scoring of the themes presented in one of the company trials. In the left column you can see that "Control of work products and product parts" and "Project management" are scoring highest meaning that the example company have good control these two themes. Likewise, "Continuous improvement and learning" and "Management insight and involvement" are the two lowest scoring themes meaning that this is where the example company has to focus and improve.

Rank	Score	Ability to run projects successfully	No.
2	43%	Project management	1
8	22%	Control across the organization	2
4	31%	Proactive management	3
6	28%	Clear goals for the projects	4
5	30%	Capable development organisation	5
3	33%	Quality of results and products	6
1	47%	Control of work products and product parts	7
10	14%	Management insight and involvement	8
7	25%	Tools and support	9
9	21%	Continuous improvement and learning	10

Fig. 7. Example of a result scoring the symptoms for an organization

We also started to discuss a mapping between the lowest levels of problems and the specific practices in the CMMI model. We will continue this work, because we believe we can find connection between the 10 themes and specific practices at the processes in CMMI. We also believe this work will strengthen the symptom-based model, e.g. if we find some practices in CMMI not addressed at the lowest level of problems – then an important symptom may be missing.

7 Findings and Discussion

We have two set of findings to report. First, we validate the use of cognitive maps.

Based on the links from symptoms to problems below and above we validated, that some symptoms where related – because the symptoms linked to the same problems/causes. It proved to us that the development of cognitive maps gave an insight in how symptoms are related seen from a problem/cause point of view. We even identified clusters of problems, which made good sense – and gained the understanding of the symptoms.

E.g. looking at symptom #6 in Fig. 1, it is strongly related to symptom #15, as both symptoms share many problems and causes. Symptom #21 is seen to be slightly weaker coupled. As a sanity check we formulated the relationship between problems, causes and symptoms as:

> *"If you do not have a good foundation for work (the problem level just below the symptoms) then it is very difficult to keep the documentation up to date. And if you do not have an updated documentation, it is difficult to correct something that has been delivered. If maintenance is difficult, it typically generates ad hoc rework, which disturb people."*

If we take another example. In Fig. 3, symptom #8 strongly links to both symptom #9 and #11at the upper problem level. Sanity check: Does it make sense to explain the relationship?

> "*If we do not know "who decides what" in the project, it is difficult to make clear decisions and will typically also include a lack of access to users or during the project (since we do not know who decides). All three symptoms lead to lack of motivation, initiated by frustration, failure to fulfill quality and a lot of rework."*

We find the cognitive maps helps a lot to clarify and structure explanations of the main reasons for main problems. They also help to qualify and strengthen the symptoms, as well as the model. We believe, that over time the working with the model will identify "weak" symptoms, which then will be updated.

The second finding came out of our summative evaluation up against an organisation where two of the authors had performed a CMMI assessment. Here it was clear, that it is not possible from symptoms to derive a maturity level for an organisation – but only pinpoint the main weaknesses important to address. But we find there is a link between the 10 themes and the processes in CMMI. We will continue that research.

As a whole our approach of looking at symptoms are somewhat similar to the SPINACH method [19] developed by Information Promotion Agency in Japan. It has a checklist with about 150 potential symptoms that could trigger further analysis to produce affinity diagram of causal system.

We also tried to 'switch over' a map to the opposite following Venable [14]. In Fig. 8 we have switched around the map that was shown in Fig. 5.

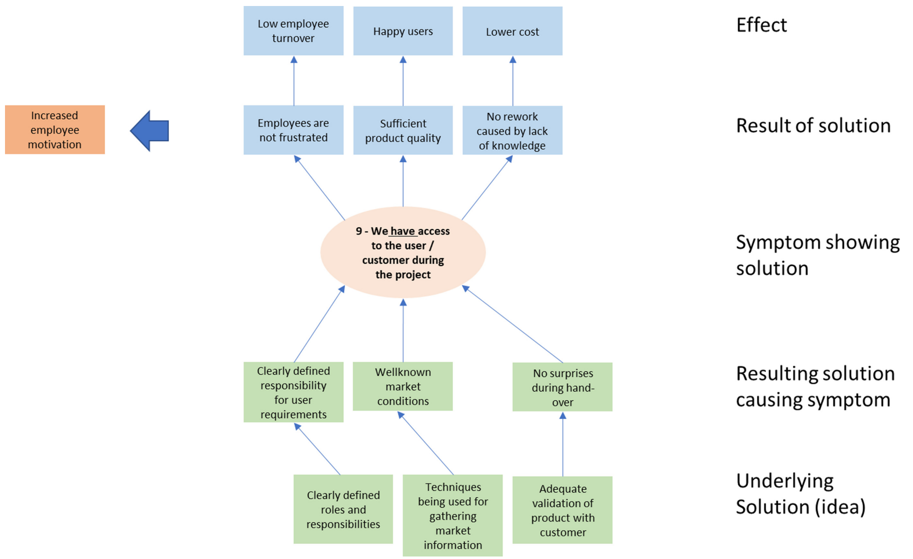


Fig. 8. An example of a design solution map – here: Symptom #9 from Fig. 5 has been switched around

In Fig. 8 we can then see that instead of having root causes at the bottom we have potential core solutions such as ‘clearly defined roles and responsibilities’, ‘techniques being used for gathering market information’ and ‘adequate validation of product with customer’. Interestingly enough these core design solutions are very similar to practices in CMMI [2] like respectively Project Planning Specific Practice 2.4 Plan the Projects Resources, Requirement Development Specific Practice 1.2 Elicit Needs and Validation Specific Practice 2.1 Perform Validation. So, there is a clear connection to CMMI at practice level.

8 Conclusion

We have answered our research question; how can we use a mapping of the symptomatic problems, causes and effects, and the relationships identified between them, to design a tool that can help determine what the most urgent improvement areas are in your company?

We have presented a sub-set of the 32 cognitive maps we have created, one for each of the symptoms we started out from. We have shown how both causes, underlying problems and effects can be related and go across the symptoms. We have also found that a group of problems or effects seen as a ‘whole’ can lead to bad business, reduced employee motivation or other major meta-level effects for a company.

To evaluate our 32 cognitive maps, we have built a tool where people can “score” all 32 symptoms and get overall recommendations. We will ask the users about their experiences with the tool. We will use these data to strengthen the model and the recommendations.

Based on our research in this paper we have realized at least two interesting topics to further research.

The first thing is to use a tool, which enable the possibility to combine all the cognitive maps to one map, with all symptoms and different types of problems and causes. We believe this map bring new knowledge on how the symptoms are related and how problems and causes are related. It will give new possibilities to improve the model and strengthen the symptoms foundation in problems and causes. We will continue this work and write a new paper addressing the findings.

The second thing it to continue the investigation of how the 10 themes are related to the processes in CMMI. We can see that the lowest level of problems in the cognitive maps for each symptom has a relationship with the specific practices in CMMI processes. For example, in Fig. 3 the underlying problem “No overall resource planning” is the caused by not performing the process Project Planning Specific Practice 2.4 Plan the project’s resources. We will map these relationships and see how the themes and symptoms are related to CMMI and use that knowledge to see if some CMMI practices are missed, which could indicate a missed problem, cause of symptom. With these relationships in place, we can guide a distressed manager towards the specific practices where improvements are highly likely to be the most beneficial.

References

1. Liker, J.: The toyota way. Esensi (2004)
2. Team, C.P.: CMMI for Development, Version 1.3, Pittsburgh, 2010, vol. 2. Software Engineering Institute, Carnegie Melon University (2018)
3. Hoermann, K., Hoermann, K., Mueller, M., Dittmann, L., Zimmer, J.: Automotive Spice in Practice, p. 75. Rocky Nook (2008)
4. Kroenke, D.: Using MIS 2013. Pearson Education UK (2013)
5. Dumdum Jr, U.R.B.: An approach to problem formulation in ill-structured situations in information systems development (1993)
6. Checkland, P.B.: Soft systems methodology. Hum. Syst. Manag. **8**(4), 273–289 (1989)
7. Checkland, P., Poulter, J.: Learning for action: a short definitive account of soft systems methodology and its use for practitioner, teachers, and students, vol. 26. Wiley Chichester (2006)
8. Checkland, P.: Soft systems methodology: a thirty year retrospective. Syst. Res. Behav. Sci. **17**(S1), S11–S58 (2000)
9. Rittel, H.W., Webber, M.M.: 2.3 planning problems are wicked. Polity **4**(155), e169 (1973)
10. Snowden, D.J., Boone, M.E.: A leader’s framework for decision making. Harvard Bus. Rev. **85**(11), 68 (2007)
11. Eden, C.: Cognitive mapping. Eur. J. Oper. Res. **36**(1), 1–13 (1988)
12. Eden, C., Ackerman, F.: Making Strategy. The Journey of Strategic Management. Sage Publications, Londres (1998)
13. Ackerman, F., Eden, C., Cropper, S.: Getting started with cognitive mapping: tutorial notes. Strathclyde: Strategic Decision Support Research Unit, Strathclyde University, vol. 18, p. 2008, November 1992
14. Venable, J.R.: Using coloured cognitive mapping (CCM) for design science research. In: Tremblay, M.C., VanderMeer, D., Rothenberger, M., Gupta, A., Yoon, V. (eds.) DESRIST 2014. LNCS, vol. 8463, pp. 345–359. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-06701-8_25

15. Hevner, A.R.: A three cycle view of design science research. Scand. J. Inf. Syst. **19**(2), 4 (2007)
16. Walls, J.G., Widmeyer, G.R., El Sawy, O.A.: Building an information system design theory for vigilant EIS. Inf. Syst. Res. **3**(1), 36–59 (1992)
17. Pries-Heje, J., Johansen, J.: SPI manifesto. European system & software process improvement and innovation (2010)
18. Venable, J., Pries-Heje, J., Baskerville, R.: FEDS: a framework for evaluation in design science research. Eur. J. Inf. Syst. **25**(1), 77–89 (2016)
19. Norimatsu, S., Kishi, T., Wada, N.: Development of "SPI strategy framework" and its application. In: Larrucea, X., Santamaria, I., O'Connor, Rory V., Messnarz, R. (eds.) EuroSPI 2018. CCIS, vol. 896, pp. 307–317. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-97925-0_25

Enough Assessment Guidance, It's Time for Improvement – A Proposal for Extending the VDA Guidelines

Thomas Wegner[1(✉)], Richard Messnarz[2], Damjan Ekert[2], Bernhardt Steger[2], Ralf Mayer[3], Rainer Dreves[4], Bernhard Sechser[5], Christian Schlager[6], and Christoph Karner[7]

[1] ZF Friedrichshafen AG, Schweinfurt, Germany
thomas.wegner@zf.com
[2] ISCN GesmbH, Graz, Austria
rmess@iscn.com
[3] BOSCH Engineering GmbH, Abstatt, Germany
Ralf.Mayer@de.bosch.com
[4] Continental, Hanover, Germany
dr.spice@web.de
[5] Process Fellows, Ebermannstadt, Germany
bernhard.sechser@processfellows.de
[6] Magna ECS, Saint Valentin, Austria
christian.schlager@magna.com
[7] KTM AG, Mattighofen, Austria
christoph.karner@ktm.com

Abstract. The VDA Guidelines used for the interpretation of ASPICE 3.1 contain many rules and recommendations and allow to check the consistency of ratings in an assessment. The VDA Guidelines contain detailed procedures and templates to be used in planning, performing and reporting an assessment. However, an assessment is only a small part compared to the support and effort needed to improve a project (level 2) and to establish a companywide process (level 3). While the ISO norm ISO/IEC TR 33014:2013 (Information technology – Process assessment – Guide for process improvement) defines a process improvement process, this has not been considered so far in the VDA Guidelines. In this paper we will (1) outline a proposal submitted by the SOQRATES working group (www.soqrates.de) to iNTACS and VDA AK 13, (2) describe the improvement guide from ISO, and (3) discuss Dos and Don'ts per principle of the SPI manifesto based on the experience from improvement projects.

Keywords: Best practices · Improvement · Process improvement · Process development · VDA guidelines · ASPICE 3.1 · ISO/IEC TR 33014

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 462–476, 2020.
https://doi.org/10.1007/978-3-030-56441-4_34

1 The Need to Focus on Improvement

In the working group meeting of SOQRATES a group of suppliers outlined a proposal to extend ASPICE 3.1 [1] and the VDA guidelines [2] to include more improvement guidance. Below you can find the proposal formulated.

1.1 Sustainable Assessments and Improvement

The experience shows that not all assessments lead to a sustainable improvement. Capability assessments are used to analyse the current status and with a status information which is valid only a certain period. However, the findings of such assessments are not considered in internal improvement programs, or in case of improvement assessments the proposed solutions are not implemented due to missing practical solutions or missing support from management.

Also, too often companies achieve a good result in one project instead of trying to achieve an improvement that can help to improve all projects in the organisation.

Additionally, not all assessments result in a list of practical improvement actions and the latter are not tracked until closure.

The VDA Guidelines only define the duties of assessors and their task ends with delivering a rating and report but the assessors have no responsibility at all for the improvement actions to be derived, the improvement program to be set up, the improvement actions to be tracked and the improved results to be achieved.

This improvement topic is not the duty of an assessor but of an improvement coordinator (which in many cases is also a local assessor in the company, acting as an improvement manager).

While such an improvement guide (also with rules and recommendations for improvement and not assessment) is probably not to be part of ASPICE 3.1 assessment model, it could become an additional chapter in the next version of the VDA Guidelines.

In ASPICE 3.1 there is the process PIM.3 "Process Improvement" and the SOQRATES group also recommends including this process in future in the VDA scope.

According to ASPICE 3.1 PIM.3 and other norms (e.g. IATF 16949) require the quality strategy in the organisation describes an improvement process and program, and how assessment results are used to establish improvement actions, how they are tracked, and how re-assessments need to be planned. For instance, if the time between assessments is very little, maybe the positive results do not have to be re-assessed and the work is focussed on the weak areas that have been improved (to validate the improvement). SOQRATES proposes 3 months.

It would make sense to extend the ASPICE 3.1 assessment model to reference an improvement program in existing base practices. For instance, in ASPICE 2.5 it was mentioned that an organisation wide improvement program should be derived (linked to) an organisation wide quality strategy. E.g. all ASPICE models till version 2.5 had the following BP1 definition for SUP.1: "A project level strategy for conducting quality assurance is developed. This strategy is consistent with the organisational quality management strategy", and this strategy aspect is not there anymore in ASPICE 3.1.

Other norms like ISO 9001:2015 or IATF 16949:2016 define and require a continuous improvement cycle as part of the norm, the VDA Guidelines and ASPICE 3.1 currently just focus on the assessment.

The latest version of IATF 16949:2016 requires also embedded (software) assessments to check the compliance with development processes.

There are exemplary improvement steps described in the norms ISO 19011:2018 and ISO TR 33014:2013 and this shall be used to establish an automotive specific proposal for an improvement process.

2 The Process Improvement Guide

The technical report "ISO/IEC TR 33014 Information technology—Process assessment —Guide for process improvement" [10, 39, 41] defines an improvement framework of activities at strategic level, tactical level, and operational level (see Fig. 1).

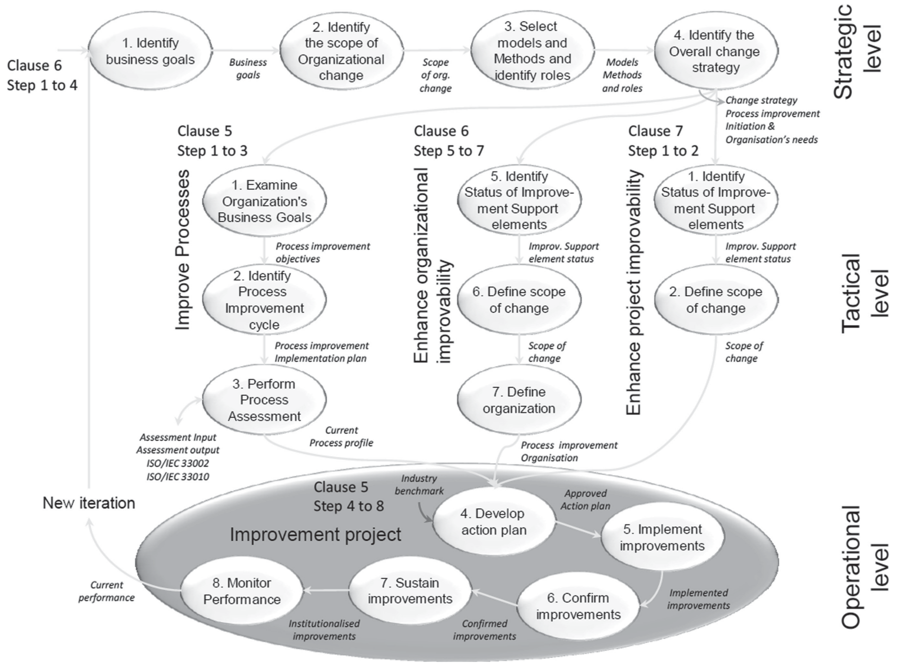


Fig. 1. Improvement framework based on ISO/IEC TR 33014

The VDA Guidelines cover assessments of the VDA scope up to capability level 3. Assessments evaluate the current status of a project. This is only one bubble in the above shown improvement framework according to Fig. 1. The improvement starts with "1. Identify Business Goals". For instance, this means to combine an existing goal of the organisation with the improvement program, such as increase the re-use of a platform in all products and assure that the platform has above 90% configurable requirements (generic) and test coverage which can be applied in all variant projects. Or increasing test automation and test coverage to reduce potential faults at a late stage, and then analyse problems found and where they could have been found, and concluding about time and money saved due to early finding of faults.

The improvement framework describes an improvement process (clause 5 steps 1 to 3, see Fig. 1) which focuses on the change of processes, and which is similar to the assessment, improvement and re-assessment strategy of processes. However, while in the VDA Guidelines a list of expectations is established, the improvement framework in ISO/IEC TR 33014 defines a detailed action plan and a set of metrics to measure the progress of the improvement.

The improvement framework describes an "enhance organisational improvability" process (clause 6 steps 5 to 7, see Fig. 1) which asks for an implementation plan at organisational level, a monitoring of the implementation and also a sustainability strategy, so that the improvements are launched and maintained across the projects in the organisation.

The improvement framework describes an "enhance project improvability" process (clause 7 steps 1 to 2, see Fig. 1) which is the project level improvement which focuses on the change of the project, and which is similar to the assessment, improvement and re-assessment strategy of the project (levels 1 to 2) in ASPICE 3.1 and the VDA Guidelines.

Also, the technical report "ISO/IEC TR 33014 Information technology—Process assessment—Guide for process improvement" describes typical roles to be involved in an improvement program (see Fig. 2).

	Supplier			Performer			User
Role / Viewpoint	Management Organization	Project Organization Project	Expert Individual	PI Unit Organization	PI Manager Project	PI team member Individual	Customer Project Mngr. Product Devl.
Horizon	Vision and strategy	Mission & objectives	Tasks	Mission	Objectives	Tasks	Tasks
Focus	Improvement Economy and business development	Overview and follow-up - scope and management	Deliver value -The knowledge the organization is missing	Consistency - structure and quality	Plan, budget and results - control	Solutions, methods, techniques and career	Solutions, efficiency, quality and security
Problem	Missing basis for decision making	Missing control and security	Missing continuously involvement	Missing trenchancy and accept	Missing stability - moving target - poor capacity	Missing time and quality - rework and overtime	Missing or too poor results on a scanty basis
Solution	**More insight:** The right Information, data and dialogue	**Better predictability:** Higher maturity	**More usable services and in demand:** More intervention	**Process improvement:** Organizational qualify	**Operational processes:** Methods and techniques	**Qualify:** Competence rewarding education	**Product improvement:** Efficient solutions

Fig. 2. Process improvement roles mapped to related elements in ISO/IEC TR 33014

It is interesting to note that the framework in ISO/IEC TR 33014 includes the roles PI (Process Improvement) unit, PI manager, and PI team member.

iNTACS could beside the role of an assessor establish a syllabus for a training of process improvement managers and PI team members.

These roles are missing in the iNTACS training program and the VDA Guidelines so far.

3 The SPI Manifesto and Experiences

3.1 The SPI Manifesto

As a manifesto for the establishment of improvement programs in organisations an SPI manifesto with values and principles has been created at EuroSPI 2009 in Alcala de Hernares, Spain [4, 11, 12, 19, 20, 23, 28, 29, 34, 40].

Members of that SPI (System, Software, Services Process Improvement) manifesto editorial team are also members of the editorial team of ISO/IEC TR 33014. The SPI manifesto created a framework to which improvement case studies can be linked. Since the foundation of the SPI manifesto an annual SPI manifesto workshop to exchange good and bad practices for SPI has been organised at EuroSPI conferences.

The SPI Manifesto defines the required values and principles for a most efficient SPI work. One main goal is to support changes by continuous improvement and innovation and include all affected stakeholders.

The current version of the SPI Manifesto describes three values and 10 principles. Each value is described with a context, a value, and hints how to achieve it (Fig. 3).

In this paper **Dos and Don'ts** to successfully implement the principles are proposed.

The authors describe **Dos and Don'ts** based on more than 25 years process improvement experiences. These **Dos and Don'ts** are suitable for assessors or PI manager and cover most of the principles.

3.2 Dos and Don'ts Suitable for ASPICE Assessors

Principle "Know the culture and focus on needs"

Do consider the differences in a global world [3, 6–9, 21, 22, 32, 33, 35]. Some cultures need more time to explain and express their materials.

Example: Make yourself familiar with the cultural specialities of countries e.g. Cultural dimensions by Geert Hofstede [9] *(Power Distance, Individualism, Masculinity, Uncertainty Avoidance, Long Term Orientation, and Indulgence)*

Do make yourself familiar with the cultural specialities, get the necessary trainings or do self-studies.

Example: If you discuss thinks with colleagues from other nationalities/level of experience, discussions may take longer, be patient it will pay off later. Appreciation of the team is a success factor, do your best to earn it.

Principle "Motivate all people involved"

Do act integer in your daily work. Judge and report always based on objectives. **Don't** say only what the audience (employee, supervisor, senior manager, director, vice president, etc.) likes to hear. **Don't** promise what you cannot keep [37].

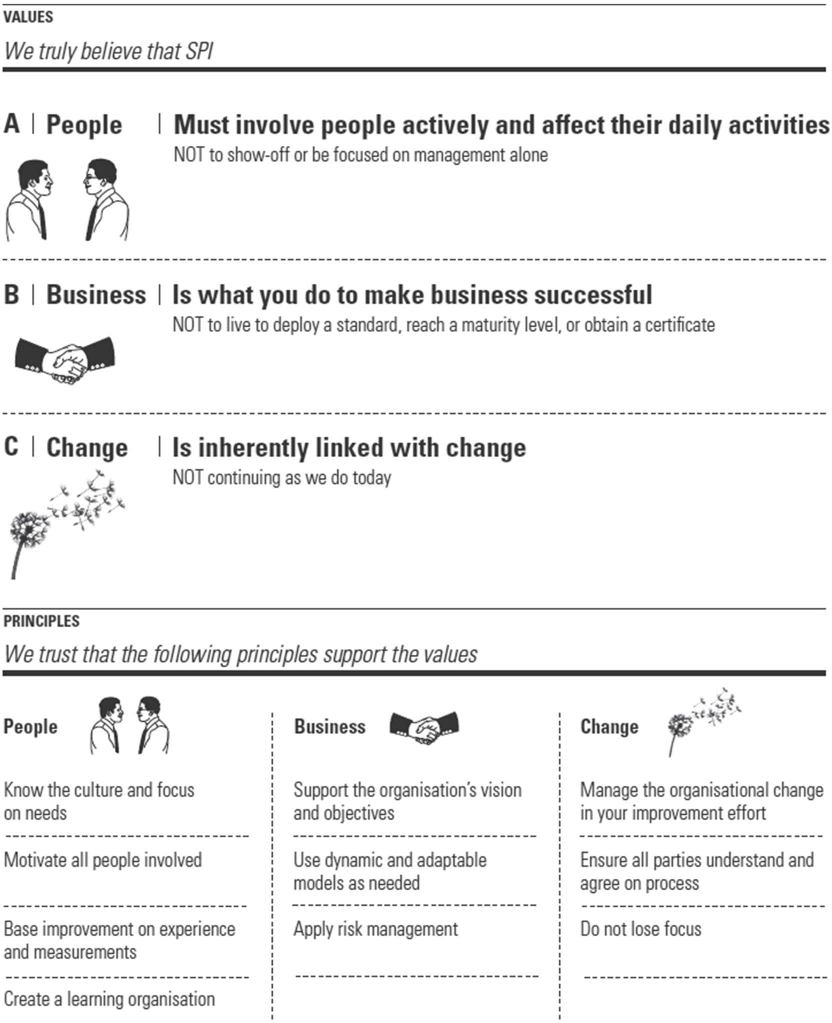


Fig. 3. SPI manifesto value and principles (EuroSPI 2009)

Don't ever use the argument, you do that because ASPICE wants it. **Do** use the real meaning as an argument.

Don't just act as an evaluator and rate the people down without openness and listening first, since you are an initiator of improvement as well. **Don't** focus only on the bad things. **Don't** please the audience. **Don't** talk out of both sides of your mouth.

Do take always into account who your recipients are - engineers. **Do** consider the target group demands in your reports and presentations. A presentation for the experts is most likely not fitting for the management. Spend time and effort to ensure that management understands where and how they can support in the improvement project.

Do ensure that the final report reflects the real situation and issues in an understandable way (consider that experts and semi-experts are reading the report).

Example: Use the process names instead of the process IDs (like SYS 1, SUP1, etc.). In some cases, assessment results can lead to unintended misunderstandings and unnecessary, severe consequences initiated by the management. If e.g. the Quality Manager gets a bad ASPICE result because his escalations are not efficient, he alone is lost. In most cases this is not caused by him, it is an organization and management issue. If this is not understood, the wrong persons get unfair consequences which leads to severe demotivation.

Principle "Base improvement on experience and measurements"

Do provide an explanation for the rating and an outlook what is missing exactly [5, 18, 31] as beside assessment the goal is to achieve real improvements.

Principle "Do not lose focus"

Do stick to the assessment schedule. **Don't** overrun the interview sessions.

Example: If the assessors don't follow the schedule, the project team becomes flustered and misinterpret the delay as their fault.

3.3 Dos and Don'ts Suitable for PI Manager

Principle "Know the culture and focus on needs"

Don't just copy an approach from one culture to another one without adapting it to the target culture as cultures are different.

Especially in Asia try to ensure that your partner never loses his face. Simply using the same expressions/wording adequate for Germans will damage or even destroy emotional level in some cultures. A sentence like "What you are doing is completely wrong" might be still fine for a German, but it might be a disaster in other cultures.

Principle "Motivate all people involved"

Do work hard on a strategy to build-up stainable intrinsic motivation. **Do** 'break the silos', be open and create a constructive, positive and trustful environment. Apply motivation technics. **Don't** use the argument, you do that because ASPICE wants it.

Example: Being positive infects the entire team, leave your personal and private issues at the company gate. A positive tenor of key persons gets more and more important and essential for motivation!

Never forget: "If you want to build a ship, don't drum up people to collect wood and don't assign them tasks and work, but rather teach them to long for the endless immensity of the sea." - Antoine de Saint Exupéry (1900–1944)

Do empower the involved people, strengthen their own responsibility and keep tool workflows simple. **Do** assure experienced engineers take care about the improvement.

Example: At a Tier-1 supplier, a new employee was called in to completely redesign the development processes. Numerous misunderstandings arose, specifications were not target-oriented and too little attention was paid to the established *approach. As a result, the acceptance of the developers towards the new development process decreased.*

Do make yourself familiar with the Belbin Team Roles (Resource Investigator, Team worker, Co-ordinator, Plant, Monitor Evaluator, Specialist, Shaper, Implementer, and Completer Finisher) and adapt your way of acting and leading. **Don't** think that all people are all the same, each person in the team is an individual with its specific characteristic and cultural/educational background.

Do organize team buildings, consider academical studies and articles.

Example: In our modern digital age driven by cost reduction measure activities that promote team building (face to face kick-off workshops, team building activities, etc.) are sometimes cancelled or replaced by Skype and other digital conferences. Do not only focus on the explicate costs. If the team building (Tuckman studies) is incomplete a negative impact on the team performance is most likely. Tuckman team development consider the 4 steps "forming" "storming" "norming" "performing".

Don't define improvements which lead to a big amount of work but show positive results very late in the project or organization. **Do** distribute the task to manageable packages. **Don't** plan to big work packages and too many changes at once.

Example: For a Tier-1 supplier, a work package ("Development of the system specification") was defined based on a deviation, which took several hundred hours. Since the responsible person saw no chance of success with this size, this work package was never started, but the activity was pushed further and further back. Only by a new approach, where this task could be divided into manageable packages (e.g. "Create the requirements for the system feature XY"), the employee could be motivated again accordingly.

Do act integer in your daily work. **Don't** be the arrogant who always knows the best solution. **Don't** show off and pretend that you know everything.

Example: Sometimes using inappropriate wording in combination with playing the hero blocks creativity. It generates fear and decreases the efficiency. A newcomer might have the best idea to solve the problem.

Don't force others to always walk in your shoes.

Example: Even if it is getting less, there are still a few who think that pressure and focusing on the negative are an efficient approach to promote success and improvement ('pain and fear as motivation agent'). Their solution is always the best fitting, there is nothing left or right.

Don't continuously remind the team on the huge amount of work ahead, this makes them feel inefficient, even if this was not intended by you.

Principle "Base improvement on experience and measurements"

Do involve experts if you have not enough experience in that specific field; **don't** be afraid to lose your face as no one can cover all topics alone. **Do** involve internal experts, if you do not have them, find external support. **Don't** define improvements only with quality engineers and/or process owners.

Do look always for existing (maybe not documented) processes and solutions and build processes on these actual working methods. **Do** be open to evaluate the presented solutions whether it fulfils the needs (and ASPICE requirements). **Do** take yourself time to really understand the actual working methods and the restrictions and needs they are built on. **Don't** start the improvement process by developing a new process from the scratch and by ignoring the existing work and processes. **Don't** define improvement actions solely based on specialist literature, using google search engine or asking someone with "dangerous half knowledge" who feels that he is the expert but isn't.

Do consider which activities are useful, not only which activities are required by ASPICE or customer requirements.

Example: SYS.2.BP5 requires verification criteria for each system requirement. But sometimes the system requirement is already written in a very meaningful way that the verification criteria can be derived easily by the text itself. So, why to write an additional verification criterion?

Don't believe that you can use the same material without adaptation to an organisation specific way of work. **Don't** just copy an approach from one organisation to another one without adapting it to the company's vision and objectives [5, 18, 28, 33] as companies are different. **Don't** apply solutions from other organizations to the present one. Each organisation is unique and so are the people. Copy paste a process which worked in another organisation will most likely fail.

Do conduct a gap analysis based on objectives. **Do** trust in what you can see as a result [18, 26] and metrics.

Principle "Create a learning organisation"

Do develop and implement an improvement process (based on the ISO/IEC TR 33014). **Don't** build your improvement on just one project. **Do** plan more than one iteration for the improvement project. Document, review and approve the new process. Collect regularly feedback from the users of the process. **Don't** expect that the results will be perfect and that no changes will be required in the future. **Do** continuously learn from assessment results and establish joint improvement across all projects in a domain [27, 30].

Don't set a too rigid framework for the process improvement project. Expect that in the future changes will be necessary. **Do** rethink whether, a solution is over engineered or not and check if there is a leaner way to archive the desired target, before implementing proposed improvements.

Example: For a Tier-1 supplier it was agreed that not only the requirements but also the graphical representations of the architecture should be documented in the same requirements management tool to achieve full traceability supported by the tool. However, the architecture was developed in other tools better suited for this purpose. This meant that the architect always had to ensure that, no matter how small a change was made, new images of the architecture were generated and integrated into the requirements management tool. This took time without bringing any benefits. Only the introduction of the acceptance of a link by name concept led to an improvement showing the same results.

Do launch an organization staffed with experts and budget and management support. **Don't** believe in miracles and think that improvement is for free, you need to spend effort and budget. **Don't** just believe that managers will give you enough budget and priority for your improvement project. **Do** get management commitment before implementing improvements where a management feedback is needed.

Example: Ensure that the persons involved in the improvement can make decisions so that decisions are not rejected later by others. Ensure management commitment on resources and tool support needed for the improvement.

Do define authorities at the beginning. **Do** ensure that the improvement team leader has the necessary authorities to approve the improvements and decisions.

Example: In an improvement workshop a new process has been defined and agreed with all involved parties. The process is prepared and documented but rejected by the department manager or another party. The improvement team is demotivated, future improvements activities will most likely fail with the argument "our work will be rejected anyhow".

Don't slavishly follow the given process and don't forget to give feedback. **Do** adapt the guidelines to the individual project or organization in a meaningful way.

Example: For a Tier-1 supplier it was a requirement that in the first phase of a customer project the specifications had to be completely reviewed. In customer projects, which were largely based on a platform or had decades of positive experience with the existing specifications, this led to unnecessary work that could not be handled by the small project teams.

Do launch web-trainings where possible, to ensure that everyone in need of a training can get at least an instant basic training or a first warm-up training prior to the onsite training. A comprehension at the end of the web-training is a plus. **Don't** rely only on the face to face or on-site training strategy.

Don't use monetary incentives as a single motivation for improvement.

Principle "Support the organisation's vision and objectives"

Do implement the process (and process landscape) as understood by the organization. **Don't** implement your personal view of a process.

Do consider the terminology [6, 7, 38] of the organisation. **Do** use the domain specific language.

Principle "Use dynamic and adaptable models as needed"

Do integrate different norms into one V-model so that with one assessment and improvement all norms can be handled [13–17, 24, 25, 36]. **Don't** create thousands of templates, and one template per different norm, letting projects drown in formalism.

Do define a release cycle (e.g. each year a major release, each four months a minor release) for the new process. **Do** install a change management process in order to analyse proposed changes from users and their impact to the existing process. Implement these changes and release a new version of the process.

Example: Defining a new process can be a long process itself collecting all information, proposing solutions, changing solutions etc. At some point you need to release a version of the process as this could otherwise lead to a never-ending improvement cycle with no results for the end user. Therefore, if you are working on a new development process, set milestones for the release and at some points move some none critical or minor issues to a following release rather than trying to consider everything.

Principle "Apply risk management"

Do carry out a risk assessment when tailoring the applied processes, show the risks resulting from not performing activities required by specifications, standards or processes. Document the deviations and the reasons. **Do** take mitigation measures into account.

Example: In today's world, there is increasing pressure on suppliers to reduce time-to-market and still use even smaller development teams. As a result, in some customer projects not all specifications from the given processes can always be fully implemented. A risk analysis can be used to determine which specifications can be reduced or even eliminated without endangering the quality of the product to be delivered. E.g. SYS.2.BP5 requires verification criteria for each system requirement. But sometimes the system requirement is already written in a very meaningful way that the verification criteria can be derived easily by the text itself. So, why to write an additional verification criterion?

Principle "Ensure all parties understand and agree on process"

Do involve all affected parties when defining improvement packages, including the one's impacted by the changes.

Example: If you start to improve your system requirements analysis process and apply changes to the process and attributes of the system requirements, don't forget to involve the testing department as the changes could also impact them (e.g. changing testing attributes or verification criteria).

Principle "Do not lose focus"

Do prepare an improvement plan.

Example: Define realistic goals and milestones which can be reached with the available resources. Define metrics to track the achievement. Perform small steps and ensure that results are accepted by the project team before performing the next step.

Do add new change requests to a backlog, analyse their priority (and impact) and start implementation according to the release plan.

Example: During implementation of changes new change requests will arise (e.g. usage of another tool, implementation of an automated script to decrease effort, etc.). Although these new change requests will be connected to other change requests already in implementation, do not start implementation of these new change requests. Start with analysis and prioritize them. Only if a new change requests get a high priority forward it.

Do describe the only necessary and needed content. **Don't** overdo it with documentation**. Don't** overdo process documentation or the handling of a tool or template

Example: Throw off ballast, keep the documentation as lean as possible, and work with pictures as "a picture is worth a thousand words".

Do take always into account who your recipients are - engineers. **Don't** overstress engineers with information that they already know from their daily work and education.

Example: Based on the education and experience, the documentation shall be kept as lean as possible. Engineers are well educated and "trained by the university *education" to understand complex situations very fast. This does not mean that you should not check e.g. if a strategy contains all what is asked by the VDA Guidelines!*

Don't implement too much secured and thereby inflated tool workflow.

Example: Because there were always deductions in assessments, as the given process did not cover or exclude all eventualities, a Tier-1 supplier developed a correspondingly complex workflow for the central SW development tool. As a result, the acceptance of the developers towards this tool decreased dramatically, which eventually affected the use of the tool and finally resulted in the development of workarounds. (This happens very often when an agile approach is implemented.)

Don't expect that everyone will be satisfied with the results of the improvement project. Focus on the relevant affected parties.

Acknowledgements. We are grateful to a working party of Automotive suppliers SOQRATES (www.soqrates.de) who exchange knowledge about such assessment strategies. This includes: Böhner Martin (Elektrobit), Brasse Michael (HELLA), Bressau Ernst (BBraun), Dallinger Martin (ZF), Dorociak Rafal (HELLA), Dreves Rainer (Continental Automotive), Ekert Damjan (ISCN), Forster Martin (ZKW), Geipel Thomas (BOSCH), Grave Rudolf (Elektrobit), Griessnig Gerhard (AVL), Gruber Andreas (ZKW), Habel Stephan (Continental Automotive), Hällmayer Frank (Software Factory), Haunert Lutz (Giesecke & Devrient), Karner Christoph (KTM), Kinalzyk Dietmar (AVL), König Frank (ZF), Lichtenberger Christoph (MAGNA ECS), Lindermuth Peter (Magna Powertrain), Macher Georg (TU Graz & ISCN), Mandic Irenka (Magna Powertrain), Maric Dijas (Lorit Consultancy), Mayer Ralf (BOSCH Engineering), Mergen Silvana (TDK/EPCOS), Messnarz Richard (ISCN), Much Alexander (Elektrobit), Nikolov Borislav (msg Plaut), Oehler Couso Daniel (Magna Powertrain), Riel Andreas (Grenoble INP & ISCN), Rieß Armin (BBraun), Santer Christian (AVL), Schlager Christian (Magna ECS), Schmittner Christoph (Austrian Institute of Technology AIT), Schubert Marion (ZKW), Sechser Bernhard (Process Fellows), Sokic Ivan (Continental Automotive), Sporer Harald (Infineon), Stahl Florian (AVL), Wachter Stefan (msg Plaut), Walker Alastair (Lorit Consultancy), Wegner Thomas (ZF).

References

1. Automotive SPICE © 3.1, Process Assessment Model, VDA QMC Working Group 1 November 2017
2. Automotive SPICE © Guidelines, 2nd edn, November 2017, VDA QMC Working Group 13, November 2017
3. Biró, M., Messnarz, R., Davison, A.G.: The impact of national cultural factors on the effectiveness of process improvement methods: the 3rd dimension. In: Proceedings of the 11th ICSQ Conference, ASQ, USA (2001)
4. Biró, M., Messnarz, R., Colomo-Palacios, R.: Software process improvement leveraged in various application domains. J. Softw. Evol. Process **26**(5), 465–467 (2014). Special Issue: Software process improvement leveraged in various application domains
5. Biró, M., Messnarz, R.: Key success factors for business based improvement. In: Proceedings of the EuroSPI '1999 Conference, Pori (1999). Pori School of Technology and Economics. Ser. A., 25
6. Clarke, P., et al.: An investigation of software development process terminology. In: Clarke, P.M., O'Connor, R.V., Rout, T., Dorling, A. (eds.) SPICE 2016. CCIS, vol. 609, pp. 351–361. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-38980-6_25

7. Clarke, P.M., et al.: Refactoring software development process terminology through the use of ontology. In: Kreiner, C., O'Connor, R.V., Poth, A., Messnarz, R. (eds.) EuroSPI 2016. CCIS, vol. 633, pp. 47–57. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-44817-6_4
8. Feuer, É., Messnarz, R., Wittenbrink, H.: Experiences with managing social patterns in defined distributed working processes. In: European Software Process Improvement. EUROSPI 2003, Graz (2003)
9. Hofstede Cultural Dimensions. https://www.hofstede-insights.com/product/compare-countries/. Accessed 7 Apr 2019
10. ISO/IEC TR 33014 Information technology — Process assessment — Guide for process improvement (2013)
11. Korsaa, M., et al.: The SPI manifesto and the ECQA SPI manager certification scheme. J. Softw. Evol. Process **24**(5), 525–540 (2012)
12. Korsaa, M., et al.: The people aspects in modern process improvement management approaches. J. Softw. Evol. Process **25**(4), 381–391 (2013). Special Issue: Selected Industrial Experience Papers of EuroSPI 2010
13. Macher, G., Sporer, H., Brenner, E., Kreiner, C.: Supporting cyber-security based on hardware-software interface definition. In: Kreiner, C., O'Connor, R.V., Poth, A., Messnarz, R. (eds.) EuroSPI 2016. CCIS, vol. 633, pp. 148–159. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-44817-6_12
14. Macher, G., Messnarz, R., Kreiner, C., et al.: Integrated safety and security development in the automotive domain. Working Group 17AE-0252/2017-01-1661, SAE International, June 2017
15. Macher, G., Much, A., Riel, A., Messnarz, R., Kreiner, C.: Automotive SPICE, Safety and Cybersecurity Integration. In: Tonetta, S., Schoitsch, E., Bitsch, F. (eds.) SAFECOMP 2017. LNCS, vol. 10489, pp. 273–285. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66284-8_23
16. Macher, G., Armengaud, E., Brenner, E., Kreiner, C.: A review of threat analysis and risk assessment methods in the automotive context. In: Skavhaug, A., Guiochet, J., Bitsch, F. (eds.) SAFECOMP 2016. LNCS, vol. 9922, pp. 130–141. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-45477-1_11
17. Macher, G., Höller, A., Sporer, H., Armengaud, E., Kreiner, C.: A comprehensive safety, security, and serviceability assessment method. In: Koornneef, F., van Gulijk, C. (eds.) SAFECOMP 2015. LNCS, vol. 9337, pp. 410–424. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24255-2_30
18. Messnarz, R., Tully, C.J., Biro, M.: Better Software Practice for Business Benefit: Principles and Experiences. IEEE Computer Society Press, Los Alamitos (1999)
19. Messnarz, R., O'Suilleabhain, G., Coughlan, R.: From process improvement to learning organisations. J. Softw. Evol. Process **11**(3), 287–294 (2006). Special Issue: Special Issue on SPI Industrial Experience
20. Messnarz, R., Ekert, D., Reiner, M., O'Suilleabhain, G.: Human resources based improvement strategies—the learning factor. J. Softw. Evol. Process **13**(4), 355–362 (2008)
21. Messnarz, R., Nadasi, G., O'Leary, E., Foley, B.: Experience with teamwork in distributed work environments. In: Stanford-Smith, B., Chiozza, E. (eds.) Proceedings of the E2001 Conference, E-Work and E-commerce, Novel Solutions for a Global Networked Economy. IOS Press, Amsterdam, Berlin, Oxford, Tokyo, Washington (2001)
22. Messnarz, R., et al.: Social responsibility aspects supporting the success of SPI. J. Softw. Evol. Process **26**(3), 284–294 (2014). Special Issue: Software Process Assessment and Improvement (EuroSPI 2011)

23. Messnarz, R., Spork, G., Riel, A., Tichkiewitch, S.: Dynamic learning organisations supporting knowledge creation for competitive and integrated product design. In: Proceedings of the 19th CIRP Design Conference – Competitive Design, Cranfield University, 30–31 March 2009, pp. 104 (2009)
24. Messnarz, R., Kreiner, C., Macher, G., Walker, A.: Extending automotive SPICE 30 for the use in ADAS and future self-driving service architectures. J. Softw. Evol. Process **30**(5), e1948 (2018)
25. Messnarz, R., Kreiner, C., Riel, A.: Integrating automotive SPICE, functional safety, and cybersecurity concepts: a cybersecurity layer model. Softw. Qual. Prof. **18**(4), 13 (2016)
26. Messnarz, R., Sehr, M., Wüstemann, I., Humpohl, J., Ekert, D.: Experiences with SQIL – SW quality improvement leadership approach from volkswagen. In: Stolfa, J., Stolfa, S., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2017. CCIS, vol. 748, pp. 421–435. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64218-5_35
27. Messnarz, R., Stöckler, C., Velasco, G., O'Suilleabhain, G., Biro, M., Remszö, T.: A learning organisation approach for process improvement in the service sector. In: Proceedings of the EuroSPI 1999 Conference (1999)
28. Messnarz, R., Horvat, R.V., Harej, K., Feuer, E.: ORGANIC-continuous organisational learning in innovation and companies. In: Stanford-Smith, B., Chiozza, E. (eds.) Proceedings of the E2005 Conference, E-Work and E-commerce, Novel Solutions for a Global Networked Economy. IOS Press, Amsterdam, Berlin, Oxford, Tokyo, Washington (2005)
29. Messnarz, R., Much, A., Kreiner, C., Biro, M., Gorner, J.: Need for the continuous evolution of systems engineering practices for modern vehicle engineering. In: Stolfa, J., Stolfa, S., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2017. CCIS, vol. 748, pp. 439–452. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64218-5_36
30. Messnarz, R., Ekert, D.: Assessment-based learning systems—learning from best projects. J. Softw. Improv. Pract. **12**(6), 569–577 (2007). Special Issue: Special Issue on Industrial Experiences in SPI
31. Sauberer, G., Riel, A., Messnarz, R.: Diversity and PERMA-nent positive leadership to benefit from industry 4.0 and kondratieff 6.0. In: Stolfa, J., Stolfa, S., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2017. CCIS, vol. 748, pp. 642–652. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64218-5_53
32. Siakas, K., Messnarz, R., Georgiadou, E., Naaranoja, M.: Launching innovation in the market requires competences in dissemination and exploitation. In: Winkler, D., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2012. CCIS, vol. 301, pp. 241–252. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-31199-4_21
33. Siakas, K., Balstrup, B.: A field-study of cultural influences on software process improvement in a global organisation. In: EuroSPI 2000 Proceedings (2000)
34. Stolfa, J., et al.: DRIVES—EU blueprint project for the automotive sector—a literature review of drivers of change in automotive industry. J. Softw. Evol. Process **32**(3), e2222 (2020). Special Issue: Addressing Evolving Requirements Faced by the Software Industry
35. O'Suilleabhain, G., Jordan, M., Messnarz, R., Bíró, M., Street, K.: The perception of quality based on different cultural factors and value systems. In: Proceedings of the EuroSPI 2000 Conference. European Software Process Improvement, Copenhagen (2000)
36. Schlager, C., Messnarz, R., Sporer, H., Riess, A., Mayer, R., Bernhardt, S.: Hardware SPICE extension for automotive SPICE 3.1. In: Larrucea, X., Santamaria, I., O'Connor, Rory V., Messnarz, R. (eds.) EuroSPI 2018. CCIS, vol. 896, pp. 480–491. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-97925-0_41
37. Key Note 2 'Integrity in Software Process Improvement' from EuroAsiaSPI 2012 in Vienna by Taz Daughtrey, James Madison University and ASQ, USA (Video: https://www.youtube.com/watch?v=SMqieU6Zi4M&feature=youtu.be)

38. Messnarz, R., Ekert, D., Zehetner, T., Aschbacher, L.: Experiences with ASPICE 3.1 and the VDA automotive SPICE guidelines – using advanced assessment systems. In: Walker, A., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 549–562. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_42
39. Pries-Heje, J., Johansen, J.: Change strategy for ISO/IEC 33014: a multi-case study on which change strategies were chosen. In: Barafort, B., O'Connor, R.V., Poth, A., Messnarz, R. (eds.) EuroSPI 2014. CCIS, vol. 425, pp. 317–330. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-662-43896-1_29
40. SPI manifesto. European system & software process improvement and innovation (2010). https://2020.eurospi.net/images/eurospi/spi_manifesto.pdf
41. Pries-Heje, J., Johansen, J.: ImprovAbility Success with process improvement, 3rd edn (2015). Whitebox, ISBN 978-87-998116-0-1

The Kanban Maturity Model in the Context of Common PI Models

Tomas Schweigert[1(✉)], Serkan Bilecen[2], and Halime Eda Bitlisli Erdivan[2]

[1] Expleo Technology Germany GmbH, Görlitz, Germany
Tomas.schweigert@expleogroup.com
[2] Turkish Standards Institution, Ankara, Turkey
{sbilecen,hebitlisli}@tse.org.tr

Abstract. The paper analyses potential impacts of the Kanban Maturity Model (KMM) on the current Software Process Improvement (SPI) Approaches as represented in the SPI Manifesto. Additionally, the paper presents and compares ISO/IEC 33014 (2013), the PEMM Model, ImprovAbility Model and Test Process Improvement (TPI) with KMM. It also contains a short comparison between the agile extension of TestSPICE and the KMM. In the last section the paper explores where the SPI community can benefit from ideas of the KMM.

Keywords: Kanban Maturity Model · KMM · ImprovAbility · SPICE · ISO 33014 · TestSPICE · PEMM · SPI Manifesto · TPI · Test Process Improvement

1 Introduction

2010 The SPI Manifesto [1] was finalized at the EuroSPI Conference 2019 in Alcala (Spain) At this time it represented the common thinking of the avant-garde of process improvers opposing bureaucratic structures commonly known as Software Engineering Process Group (SEPG). It was accompanied by an European Training and Certification Scheme [2]. The SPI Manifesto rapidly became a cornerstone of SPI thinking with only a little need to debate about the appropriateness and the integrity of the content. There was one change idea presented in a paper which was more addressing ethical issues [3] and during a workshop at the EuroSPI Conference 2019 in Edinburgh, no additional topics about it were discovered. But the SPI Manifesto was set aside by the rush of the agile Manifesto [4] and the Scrum Method [5]. In the decade after the agile manifesto was published, lots of so called agile maturity models were published [6, 7], but turned out as more or less poor rewritings of the Capability Maturity Model Integration or ISO/IEC 15504 Part 5 [8].

In 2010 David Anderson published his book about Kanban [9], positioning it as a method for evolutionary change. The Kanban Journey had its next milestone with the publication of the Kanban Maturity Model 2018 [10].

The paper will try to give an answer to the following questions:

- How is the model positioned in the context of other PI Models?
- Is the model clear and its architecture understandable?

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 477–489, 2020.
https://doi.org/10.1007/978-3-030-56441-4_35

- Is the model consistent in the use of its elements?
- Does the model provide Innovation in the field of Maturity, Capability and/or improvement models?

In order to check the degree of innovation an overview over the currently available improvement models will be given, starting with the SPI Manifesto.

2 The SPI Manifesto

The SPI Manifesto consists of 3 Values and 10 Principles. They describe aspects that need attention during an SPI Project rather than to describe what to do in a given SPI Project on a technical level or how to run an improvement campaign (Table 1).

Table 1. Values and principles of the SPI Manifesto

People	Business	Change
Value: **SPI must involve people actively and affect their daily activities**	**Value** **SPI is what you do to make business successful**	**Value** **SPI is inherently linked with change**
Principle Know the culture and focus on needs	Principle Support the organizations vision and objective	Principle Manage the organizational change in your improvement effort
Principle Motivate all people involved	Principle Use dynamic and adaptable models as needed	Principle Ensure all parties understand and agree on process
Principle Base improvement on experience and measurement	Principle Apply risk management	Principle Do not loose focus
Principle Create a learning organization		

There is no formal detailing of the principles but there are explanations for each principle in the manifesto.

2019 Geogiadu et al. argued for an extension [3]:

New Value: Society

New Principles:

- To fulfill ethical duties
- To comply with legislation

3 The ImprovAbility Model

The ImprovAbility Model [11] of process improvement was mainly developed by IT-University, Copenhagen, Denmark, DELTA and 4 companies in a 3-year project. It created a view on organizations and projects ability to improve. It also contains a section about change strategies [12]. The analysis of change strategies helps to find out the right approach to implement necessary changes (Table 2).

Table 2. Parameters of ImprovAbility model

Foundation parameters	Initiation parameters	Project parameters	In use parameters	Change strategies
Vision and strategy	Sensing urgency	Project team	Deployment strategy	Commanding
Organisational culture	Idea creation	Project process	Product quality	Employee driven
Expectation management	Idea processing	Project competence and knowledge	Deployment means	Exploration
Knowledge management		Project prioritizing	Roles and responsibility	Learning driven
Management competence		Project goals and requirements	Maintenance	Metrics driven
		Management support		Optionality
		Involvement of others		Production organized
				Reengineering
				Socializing
				Specialist driven

For assessment purposes these parameters are detailed into specific questions. These questions help to understand the organization under assessment, but they can also be used as building blocks for parameter implementation during the improvement project.

4 ISO/IEC 33014 - The ISO Standard for Software Process Improvement

ISO/IEC JTC1 SC7 WG10 (Process Assessment) developed the ISO/IEC 33014 [13] standard, based partially on ImprovAbility model. It organizes Software Process Improvement in 3 views:

- Steps of process improvement
- Organizational support for process improvement
- Organization of process improvement work

The following table shows a rough overview about the next level of detail (Table 3):

Table 3. The structure of ISO/IEC 33014

Steps of process improvement	Organizational support for process improvement	Organization of process improvement work
Examine organization's business goals	Identify business goals	Enhance project improvability
Initiate process improvement cycle	Identify the scope of organizational change	Identify status of improvement support elements
Perform a process assessment	Select models and methods and identify roles in process improvement	
Develop action plan	Identify the overall change strategy	
Implement improvements	Identify status of improvement support elements	
Confirm improvements	Define scope of change – and what to change	
Sustain improvements	Define organization for enhancement and process improvement	
Monitor performance		

5 Process Enterprise Maturity Model (PEMM)

PEMM [14] defines separate maturity levels for organization and process, stating that the maturity of the organization limits the maturity of processes (Table 4).

PEMM also contains maturity topics, which help to assess the status and to organize improvement programs similar to the questionnaires of ImprovAbility (Tables 5 and 6).

Table 4. The structure of two different maturity level dimensions of PEMM

Organisational maturity characteristics	Process maturity characteristics
E-0 Immature organization	P-0 Process functions by random
E-1 Organization with basic capability of process management	P-1 Process is reliable, predictable and stable
E-2 Organization with the ability to handle projects with contribution from several units	P-2 Process is designed E2E along a value chain
E-3 Organization with an established process and project framework	P-3 Process can be integrated with other processes to optimize the performance and output of the organization
E-4 Organizational ability to integrate processes of customer and supplier	P-4 Process is defined beyond the borders of the organization integrating customer and supplier

Table 5. The topics of PEMM on enterprise level

Leadership	Culture	Expertise	Governance
Awareness	Teamwork	People	Process model
Alignment	Customer focus	Methodology	Accountability
Behavior	Responsibility		Integration
Style	Attitude towards change		

Table 6. The topics of PEMM on process level

Design	Performers	Owner	Infrastructure	Metrics
Purpose	Knowledge	Identity	Information systems	Definition
Context	Skills	Activities	Human resource systems	Uses
Documentation	Behavior	Authority		

6 Test Process Improvement (TPI) Next

Contrary to ISO/IEC 33020 that requires Capability levels based on generic practices, levels in TPI NEXT [15] are based on detailing and improving basic testing practices.

By doing so, the TPI NEXT Model provides an alternative architecture for maturity models (Fig. 1).

7 The Agile Extension of TestSPICE 4.0

In order to provide a SPICE compliant Process Assessment Model (PAM) for agile processes the TestSPICE SIG developed an agile Process Reference Model (PRM) [16] (Table 7):

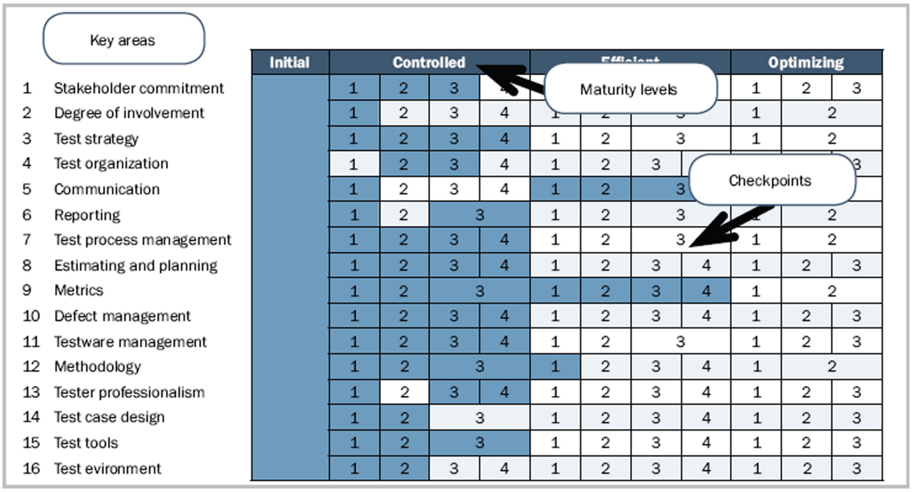


Fig. 1. The maturity matrix of TPI NEXT [15]

Table 7. Agile extension of TestSPICE 4.0

Agile management process group (AMP)	Agile transition process group (ATP)	Agile governance process group (AGP)
AMP.1 Backlog management	ATP.1 Organizational capacity management	AGP.1 Align with strategic directions
AMP.2 Sprint/Iteration management	ATP.2 Agile framework adaption	AGP.2 Support organizational effectiveness
AMP.3 Impediment management	ATP.3 Agile framework monitoring	AGP.3 Support performance reliability
AMP.4 Service class and WIP limit management	ATP.4 Agile sourcing	AGP.4 Assure operational performance
AMP.5 Technical debt management	ATP.5 Agile environment management	
AMP.6 Knowledge debt management		
AMP.7 Definition of done (DoD) management		

This PRM deals with

- Agile Management
- Agile Transition
- Agile Governance

8 The Kanban Maturity Model

In order to explain the Kanban Maturity Model [10], this chapter investigates the following topics:

- The Kanban Maturity Model at a glance
- Kanban Values, Principles, and General Practices
- The Target of the Kanban Maturity Model
- The maturity levels of the Kanban Maturity Model

8.1 The Kanban Maturity Model at a Glance

Similar to the Approach of Jacobsen [17] the KMM is not based on process but on practices. Practices are divided in 2 groups: General and specific practices (a short overview given by Srinath Ramakrishnan [18]. However, it is not free of process awareness. The descriptions of the Maturity levels are linked to typical process capabilities.

Specific Practices deepen and guide the implementation of the general practices. This approach is similar to the approach of the TPI Next model.

There are 2 types of specific practices:

- Core practices
- Transition practices

Core practices describe typical activities associated to the level.

Transition practices are practices that open for the achievement of the next level (Fig. 2).

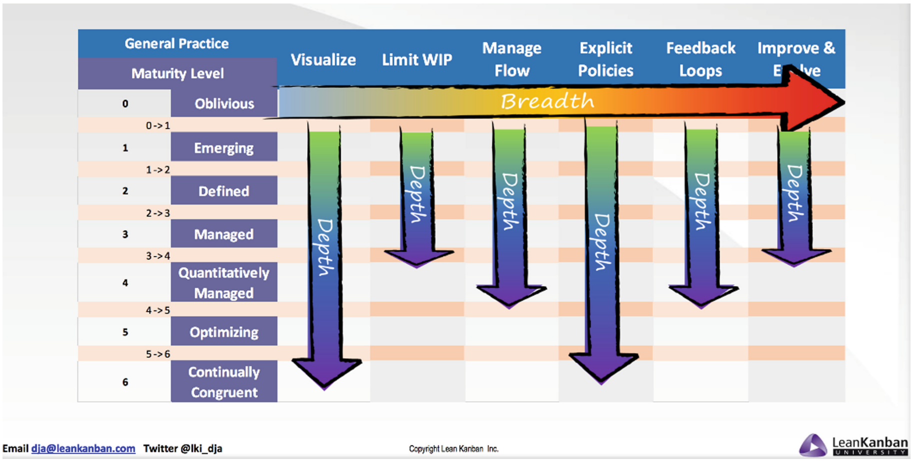


Fig. 2. Kanban practices determine the degree of maturity [19]

8.2 Kanban Values, Principles and General Practices

The Kanban Maturity model is designed as an improvement path for organizations that want to improve their Kanban capabilities.

Based on the work of Burrows [20], Anderson and Carmichael [21] describe their view on Kanban also as a set of values and principles (Table 8):

Table 8. Relations between Kanban values, principles and practices

Kanban values	Kanban principles	Kanban general practices
Transparency	Start with what you do now	Visualize
Balance	Agree to pursue improvement through evolutionary change	Limit work in progress
Collaboration	Encourage acts of leadership at every level, from the individual contributor to senior management	Manage flow
Customer Focus	Understand and focus on your customers' needs and expectations	Make policies explicit
Flow	Manage the work; let people self-organize around it	Implement feedback loops
Leadership	Evolve policies to improve customer and business outcomes	Improve collaboratively, evolve experimentally
Understanding		
Agreement		
Respect		

8.3 The Target of the Kanban Maturity Model

The target of the KMM is to develop the "Fit for Purpose" organization. There isn't any specific description of what is meant with the term "fit for purpose organization" but there are some hints in the model that give an assumption about the underlying thinking.

Organizational Capabilities: The KMM aims to establish the following organizational capabilities:

- Relief from overburdening
- Delivering on customer expectations
- Organizational agility
- Predictable financial outcomes and organizational robustness
- Survivability

Level 6 Key Questions. The KMM describes a level 6 company (which should be fit for purpose) as having the capability to ask the following questions:

- Is the way we are doing things still competitive? Are new technologies processes methods or means becoming available that we should be investigating or adopting?
- Do we offer the right products or services? And if not, how should we change?
- Are we serving the right markets? Do we have the capability to serve our chosen markets adequately?
- Who are we as a company? Is our current identity relevant or appropriate? Do we need to reinvent ourselves?

Advanced Shared Purpose on Level 6
At level 6 an organization is capable to find answers to How, What, Why and Who but in a sense of agility. This type of culture includes

- Challenging established norms and finding better ways
- Challenging established conventions and behaviors
- Innovation
- Change

8.4 The Maturity Levels of the Kanban Maturity Model

This section refers to process or capability related parts of the description:

ML 0 Oblivious: The organization is oblivious to the need to follow a process. There is no awareness of management values, organizational processes or policies.
ML 1 Emerging: This ML no or initial process definitions are available. Given processes are not followed consistently.
ML 2 Defined: This ML there is a basic definition of processes which is followed consistently. But there is a lack of consistency in the desired outcomes. The process definition describes "the way we do things" no or initial process definitions are available. Given processes are not followed consistently.
ML 3 Managed: Processes are followed consistently, and outcomes are consistently achieved. But decision making remains mostly qualitative and emotionally driven.
ML 4 Quantitatively Managed: Processes are followed consistently and outcomes are consistently achieved. There is a shift to quantitative decision making and a culture of underpinning decisions with data.
ML 5 Optimizing: There is extensive process instrumentalisation. Workers have a sense of ownership over their own processes and a pride in their capabilities and outcomes. There is an organizational focus on improvement.
ML 6 Congruent: Capabilities are aligned with strategy.

8.5 Comparison of KMM and the Other Models Presented

8.5.1 Comparison of KMM Against PEMM

In the KMM, processes are more seen as a necessary part of the business infrastructure. The maturity of the organization describes the capability of the organization to implement, change and improve processes as part of the implementation, change and improvement of business.

In the PEMM, high capability processes are the outcome of a mature organization. The transformation of Leadership is a transformation to process thinking. While Business is the main driver of KMM it can be debated if the PEMM has the same clear focus on fit for purpose. The Article in which Michael Hammer published the PEMM contains a little article which implies the assumption that the PEMM is not capable to support the business driven adjustment of resources.

8.5.2 Comparing KMM Against TPI Next

The main difference is that TPI NEXT is focusing only on testing. But the similarity is that Key Topics (in KMM General Practices) are deboned in specific Practices and that along this path Maturity levels are defined.

8.5.3 Comparing KMM with the ImprovAbility Model and ISO/IEC 33014

While ISO/IEC 33014 follows the traditional approach of Process improvement using defined, planned steps, the KMM follows a more evolutionary learning by doing approach (Start where you are now and use evolutionary steps). Also, the KMM requires to encourage consistent and competent leadership on all levels of an organisation while leadership does not play a big role in the ImprovAbility Model and ISO/IEC 33014.

8.5.4 Comparing the KMM with the Agile Extension of TestSPICE

The Agile Extension of TestSPICE following the COSO Model sees the agile target organization as an agile organization that achieves a reliable performance, which nearly is an equivalent to the L5 and/or L6 organization of Kanban. The difference is that the Agile Extension of TestSPICE does not measure the agile maturity of an organization but measures the capability of agile key processes using Levels, Process Attributes and generic practices as a framework to describe these capabilities.

8.6 A Short Comment on the Structure of the Model

Starting 2012 a lot of so-called agile maturity models were published which had more or less severe flaws in their structure and not many added value to ISO 15504 Part 5.

Checking the KMM the following quality criteria may be discussed

Is the model clear and its architecture understandable?

The KMM has clearly defined levels and every level is composed on the same type of elements.

Is the model consistent in the use of its elements?

The KMM uses its elements in a constant and congruent manner. This is supported by the ID of the specific practices and the structure of the level description.

What's missing?

The KMM does not clearly deal with dependencies across specific general practices. Other than TPI next that defines Maturity levels from different levels of capability, it is unclear if the implementation of i.e. a L4 specific practice of Flow requires complete implementation of all L3 practices, or a mixed subset or just the L3 practices of Flow.

9 What is New and/or Different in the KMM?

See Table 9.

Table 9. Comparison table of models and their elements

	Target organisation	Values & principles	Improvement process	Change strategies	Defined practices
SPI manifesto	–	Yes	–	–	–
ImprovAbility	–	–	–	Yes	–
ISO/IEC 33014	–	Yes (4.3/4,4)	Yes	Yes	–
PEMM	Yes	–	Yes	–	Yes
TPI next	Yes (testing only)	–	–	–	Yes
Agile extension of TestSPICE	–	–	–	–	Yes (processes and base practices)
KMM	Yes	Refers to Kanban values	Implicit (core and transition practices)	–	Yes (general and specific practices)

10 Can the SPI Manifesto Benefit from the KMM?

Comparing the KMM to the other models presented in this paper the Kanban general practices and their detailing make the KMM unique, delivering fresh ideas for PI professionals. As Kanban literature often focuses on the visualization practices in the Kanban board [22–26]. There are some Limited WIP societies active, which might lead to the assumption that the Limit WIP general practice is the key feature. Leopold [27] explicitly addresses the need for Leadership in the implementation of Kanban.

If there is something in the model to enhance the SPI Manifesto, it should not be at technical level like Visualisation, Manage Flow or Limit WIP. Also the general practice "Implement Feedback Loops" is observed in various principles of other models. Improve collaboratively, Evolve Experimentally is more known from Business transition. But could be useful also as a PI principle.

In conclusion, the general practice "Make Policies Explicit" should be taken into account when updating the SPI Manifesto. This General Practice is not addressed as a principle in the SPI Manifesto but has a potential to create significant value for the SPI community.

References

1. Pries-Heje, J., Johansen, J. (eds.): MANIFESTO Software Process Improvement (SPI Manifest) eurospi.net, Alcala, Spain, 1010. http://2019.eurospi.net/images/eurospi/DownloadCenter/spi_manifesto.pdf. Accessed 02 May 2019
2. Korsaa, M., et al.: The SPI Manifesto and the ECQA SPI manager certification scheme. J. Softw. Evol. Process **24**(5), 525–540 (2012)
3. Georgiadou, E., Siakas, K., Berki, E., Estdale, J., Rahanu, H., Ross, M.: A STEEPLED analysis of the SPI Manifesto. In: Walker, A., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 209–221. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_16
4. Beck, K., et al.: Manifesto for Agile Software Development. Snowbird Ski Resort (2001). http://agilemanifesto.org/
5. Schwaber, K., Sutherland, J.: The official Scrum Guide. https://www.scrumguides.org/scrum-guide.html. Accessed Apr 2020
6. Biro, M., Korsaa, M., Nevalainen, R., Vohwinkel, D., Schweigert, T.: Agile maturity model. Go back to the start of the cycle. In: Industrial Proceedings of the 2012 EuroSPI Conference, pp. 5.9–5.30 (2012)
7. Schweigert, T., Korsaa, M., Nevalainen, R., Vohwinkel, D., Biro, M.: Agile maturity model: oxymoron or the next level of understanding. In: Proceedings of the SPICE 2012 Conference, pp. 289–294 (2012)
8. Schweigert, T., Vohwinkel, D., Korsaa, M., Nevalainen, R., Biro, M.: Agile maturity model: analysing agile maturity characteristics from the SPICE perspective. J. Softw. Evol. Process **26**(5), 513–520 (2013). https://doi.org/10.1002/smr.1617. Wiley Online Library (wiley-onlinelibrary.com)
9. Anderson, D.J.: KANBAN – Successful Evolutionary Change for Your Technology Business. Blue Hole Press, Sequim (2010)
10. Anderson, D.J., Bozheva, T.: Kanban Maturity Model – Evolving Fit for Purpose Organisations. Lean Kanban University Press, Seattle (2018). ISBN 978-0-9853051-5-4
11. Pries-Heje, J., Johansen, J. (eds.): ImprovAbility™ - Success with Process Improvement. Delta, Horsholm (2013). ISBN 978-87-7398-139-9
12. Pries-Heje, J., Vinter, O.: Selecting change strategies in IT organisations – proof of strategy. In: Pries-Heje, J., Johansen, J. (eds.) ImprovAbility™ - Success with Process Improvement. Delta, Horsholm (2013). ISBN 978-87-7398-139-9
13. ISO/IEC TR 33014 (2013). Information technology—Process Assessment—Guide for process improvement (2013)
14. Hammer, M.: The Process Audit - A new framework, as comprehensive as it is easy to apply, is helping companies plan and execute process-based transformations. Harvard Business Review, April 2007. P 111 ff. http://class.svuca.edu/~laurauden/class/Session-11-materials/11-03%20Handouts%20on%20assessment/11-03b-Process%20capability-PEMM/1-Hammer-Process%20Audit.pdf
15. Sogeti TPI Next – Geschäftsbasierte Verbesserung des Testprozesses. dpunkt, Heidelberg (2011)
16. TestSPICE PAM v. 3.1. www.testspice.info
17. Jacobson, I., Ng, P.W., Spence, I.: Enough of processes - lets do practices. J. Object Technol. **6**(6), 41–66 (2007)
18. Ramakrishnan, S.: Brief Summary of Kanban Maturity Model Evolving Fit for Purpose Organizations. https://srinathramakrishnan.files.wordpress.com/2018/05/brief-summary-of-kanban-maturity-model.pdf. Accessed Apr 2020

19. Schultheiss, F.: The Kanban Maturity Model (KMM) - A Framework That Makes Organizations More Agile and Adaptable (2017)
20. Burrows, M.: Kanban – Kanban from the inside Sequim. Blue Hole Press (2015)
21. Anderson, D.J., Carmichel, A.: Essential Kanban Condensed. Lean Kanban University Press, Seattle (2016)
22. Lewis, C.: Kanban Confusion, An Introduction of Context and Use. Alderbank House (2019)
23. Agarval, A.: The Basics of KanbanAditi Agarwal books (2018)
24. Leopold, K.: Agilität neu denken. LEANability, Wien (2018). ISBN 978-3-903205-50-5
25. Skarin, M.: Real World Kanban – do less accomplish more Dallas, the pragmatic Programmers (2015). ISBN 13:978-1-68050-077-6
26. Styeart, P.: Essential Upstream Kanban. Lean Kanban University Press (2017). ISBN 978-0-9845214-7-0
27. Leopold, K., Kaltenegger, S.: Kanban Change Leadership – Creating a culture of Continuous Improvement. Wiley, Hoboken (2015). ISBN 978-1-119-01970-1

Derivation of SPI Manifesto Principle Example from an SPI Story

So Norimatsu(✉) and Noriko Wada

Japan Software Process Improvement Consortium (JASPIC),
Toshima-Ku, Tokyo, Japan
contactB@jaspic.org

Abstract. The authors have been developing and using a framework for SPI strategy in one of the working groups at "Japan SPI Consortium (JASPIC)." Our objective is to understand important (i.e., influential) strategic factors when promoting SPI, and to use them in the actual SPI cases as guidelines.

In response to the request for new examples for principles in SPI Manifesto, this paper describes how we derived such examples from one SPI case. We started with one of the SPI cases we supported. Key strategic features were extracted and transformed into three "principle examples" for the SPI Manifesto. The paper discusses other potential derivation methods to promote the creation and sharing of examples.

Keywords: Software Process Improvement · SPI strategy · SPI Manifesto · Derivation of example

1 Introduction

The authors have been providing strategic analysis, diagnosis, and support to various Software Process Improvement (SPI) cases among the members of Japan SPI Consortium (JASPIC) since 2012 [1, 2]. We describe the SPI case using the "7 components of SPI Case" template [3], analyze strategic factors using the "SPI Strategy Framework" [4], and provide strategic guidance to the organization so that it can achieve its goals effectively.

In response to the request for new examples for principles in the SPI Manifesto [5], we started with one of the SPI cases we supported. Key aspects of its SPI strategy were extracted first to highlight the case. Then, the aspects were mapped to the principles in the Manifesto. Finally, an example description of the principle was created by combining the relevant aspects and additional information.

This paper is organized as follows. Section 2 describes the SPI case that we analyzed. Its background and activities are described as an overview in Sect. 2.1. Important strategic features identified by the promotor are described in Sect. 2.2. Section 3 describes the examples for the Manifesto. Identified strategic features are recomposed into the scope of each principle. Section 4 opens a discussion on the derivation methods.

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 490–496, 2020.
https://doi.org/10.1007/978-3-030-56441-4_36

2 SPI Case Description

In this section, we describe the SPI case we used to create examples for the Manifesto. Due to the confidentiality agreement, the name of the company is not provided.

2.1 Case Overview

The organization has been continuing company-wide SPI to reduce the Cost of Poor Quality. It has established standard processes, enacted process QA groups, and these standard processes are widely acknowledged.

Despite their efforts, the organization experienced a significant product quality issue on the market in 2017. The quality department performed a causal analysis and found that the development project was performed without any notable problems, at least on the surface. Standard processes were followed, and product specifications and designs were documented according to the standard. It was determined that the cause of the product defect was the quality of the contents of the documents. There were missing items in the specification, which lead to design errors and missing test items. Traceability among specification, design, and verification documents was not adequate, either. It was concluded that these documents were primarily created to show compliance with the development process, but they were not adequate to guide engineering activities efficiently.

Based on the finding, improvement activities were initiated. The goal was to present methods to create the engineering documents properly and to establish integrity among specification, design, code, and verification over the development lifecycle. The quality department adopted the UML method for documentation and established criteria for specifications and designs. It also provided training, consulting support, and design reviews so that the development projects can satisfy these criteria. The organization is planning to deploy this UML based documentation process to all the product categories by the end of 2021. As of early 2020, deployment has been successful as planned, and positive effects (such as significantly lower defect density at the system test and lower efforts required for the test case development) have been observed in the deployed projects.

2.2 Important Strategic Features

In this SPI case, the promotor in the quality department has been aware that the following four features are essential:

- A: Meaning and value of the new processes should be understood before standard processes are updated.
- B: Early adopters are identified, supported and become successful
- C: Appeal to engineer's pride and competitive spirit to accelerate adoption
- D: Update standard processes based on a success story

The following sections explain each feature.

Feature A: Understanding of the Meaning and Value of the New Processes
The promotor was aware that if the method is deployed as updates of the standard processes, the focus would be the compliance like before, and the new processes would become another formality before the practitioners would understand their meanings and values. Therefore, the initial step was focusing on promoting the understanding of the new processes, not updating the organization's standard.

These steps include the following:

- Delivering seminars and training courses by external experts
- Establishing guidelines for the contents of proper engineering documentation
- Communicating the status quo on the level of cost of poor quality and its causes

Feature B: Early Adopters
When an initial awareness campaign was conducted through seminars and training courses, individuals and groups who showed strong interests in the new method were identified as potential early adopters. Tool support, consultation, and review were provided intensively for these early adopters. Product categories that have more substantial business impacts were also targeted. As a result, these early adopters started producing successful outcomes at the development project level.

Feature C: Engineer's Pride and Competitive Spirit
Appealing to the engineer's pride and competitive spirit was considered to help accelerate the adoption of the new method. Success stories were widely spread through the engineer's forum to stimulate their minds.

Feature D: Standardization Based on Success Stories
Instead of updating the standard processes by the Engineering Process Group and stipulating them as company-wide process rules as the first step, success stories in the actual projects were converted into the process requirements at the company level, then reviewed and approved by the subject matter experts. These process requirements were then incorporated into the divisional standard processes. This standardization method ensures that the changes are incorporated into the standard when meaning and values have been confirmed by the developers, thus helping smooth adoption of the process change.

3 Derivation of SPI Manifesto Principle Example

3.1 Mapping from Strategic Feature to SPI Manifesto Principle

The four strategic features explained in the previous section were mapped to the SPI Manifesto based on their relevance to the explanation section of the principle. We identified the following three principles with a higher level of relevance, for which principle example will be produced.

- People: Know the culture and focus on needs
- People: Motivate all people involved
- People: Create a learning organization

The level of relevance between strategic features and principles that we observed is shown in Table 1. "High-Medium-Low-None" denotes the level of relevance.

Table 1. A mapping between Strategic Feature and Manifesto (People value)

Manifesto value	People		
Manifesto principle	Know the culture and focus on needs	Motivate all people involved	Create a learning organisation
A. Understanding of Meaning & Value	High	Medium	Medium
B. Early Adopters	None	Low	Medium
C. Engineer's Pride & Competitive Spirit	High	None	None
D. Standardization based on Success	Low	High	Low

We also identified the following two principles with lower relevance.

- Change: Manage the organisational change in your improvement effort
- Change: Ensure all parties understand and agree on process

3.2 Manifesto Principle Example

For the three principles with higher relevance, descriptions of the strategic features were recomposed to match the scope and intent of each principle. Additional information was added from the SPI case for better understandability.

Minimum contextual information may need to be shown so that the reader of the SPI Manifesto can understand the example better. In the Manifesto, it would be better this contextual information is shown along with each example. In this paper, it is shown only once as "Context".

Context

(Context A) An organization conducted company-wide improvement activity, triggered by a significant product quality issue on the market. In order to improve the quality of the engineering documents (e.g., requirements specification, design, and verification specification), a new documentation method was introduced.

(Context B) Stakeholders of the improvement activity included the quality department as a sponsor, several members in the company level engineering process group as promotors, several business units as subgroups, and several hundred developers as practitioners.

People: Know the Culture and Focus on Needs

(Context A)

The promotor of the SPI initiative was aware that there is a culture of "pride as an engineer" in this organization and that the top-down standardization approach would not work for a large group of people with a strong personality. If the new method is initially deployed as part of standard processes, there would be strong resistance, or the focus would be the compliance, and the new process would become another formality before the practitioners would understand its meanings and values. Therefore, the focus was for the engineers to "work on my initiative," and the quality department provided various promotions and support before process standardization.

These promotions and support included seminars by external experts to communicate the engineering values, training, and guidelines to understand the new methods, and forums to share other engineer's experiences and results. These forums also helped to stimulate the engineer's pride and competitive spirit to accelerate adoption.

People: Motivate all People Involved
(Context A) (Context B)

The promotor of the SPI initiative was aware that motivating engineers is a key to the successful adoption of the new method, and the top-down standardization approach should be avoided. First, developers were provided opportunities to understand the values of the method through seminars and training courses. Individuals and groups who showed strong interests were identified as potential early adopters who received intensive support. Success stories and experience reports from these early adopters were disseminated through periodical engineer's forums and publications to motivate a larger group of developers. Test groups also participated in these forums and published the effects of the new method to motivate other test engineers. Project managers who observed these positive results started adopting the method. Senior management was also involved through regular reports about adoption status and outcome of the new method, and the chief technical/quality officer communicated the need for the new method to the organization.

People: Create a Learning Organization
(Context A)

The promotor of the SPI initiative planned learning opportunities of the new method in various stages and sources.

The first step was to introduce the values and concepts to the engineers by providing technical seminars by external experts. The need for the new method was communicated by showing the current level of cost of poor quality.

The second step was to provide an understanding of the method. Individuals and groups who participated in the method training were identified as potential early adopters, and they received further support (i.e., tools, consultation, and review) to use the method in their actual projects. This additional support created not only a practical understanding of the method application but also success stories.

The third step was to exchange and disseminate project level experiences through the regular engineer's forum and technical publication.

Finally, these practical experiences were codified in the standard processes to form a basis for continuous improvement.

4 Future Directions and Discussions

There are three possible methods to create Manifesto principle example:

1. Direct Extraction

   A mini-story as part of the SPI case is extracted to represent a specific principle.

2. Indirect Extraction.

   First, features of the SPI case are extracted and mapped to the relevant principles. Then for a single target principle, a mini-story is recomposed by combining relevant features in order to represent the scope and intent of the principle.

3. Mapping through Description Model

   First, an SPI case is described using the "SPI case description model," such as the "SPI Strategy Framework." Using the mapping between the model element and the Manifesto principle, a mini-story for the target principle can be created by combining information in the corresponding model elements.

Selection of appropriate method depends primarily on the availability of knowledge of each SPI case and understanding of SPI Manifesto. For a person who possesses both knowledge and understanding, the first method of "Direct Extraction" would be straightforward. For someone like the authors who are supporting other organizations, the second method of "Indirect Extraction" is more efficient. The third method of "Mapping through Descriptive Model" takes more effort, but it gives a comprehensive understanding of the SPI case, and it enables various other applications. It can also produce more examples.

In this paper, the authors primarily used the second method in the context of our activities at JASPIC. We also piloted the third method using the "SPI Strategy Framework" that we developed. The details, along with its application, will be reported in the future.

5 Conclusions

In this paper, we presented an SPI case and its strategic features. They were mapped to relevant SPI Manifesto principles, resulting in principle examples.

Acknowledgments. This paper was written as a result of activities performed by a working group for "process improvement strategy" in Japan SPI consortium (JASPIC). We are thankful to the JASPIC steering committee and staff for supporting our activities.

References

1. Japan Software Process Improvement Consortium: http://www.jaspic.org
2. Wada, N., Norimatsu, S., Kishi, T.: Analysis of SPI strategies and their modeling – Exploring the art of SPI promotion. In: Industrial Proceedings, EuroSPI2012 (European Systems Software & Service Process Improvement & Innovation) (2012)
3. Norimatsu, S., Endo, K., Usugi, M., Niwa, A., Tange, K.: Formulation of process improvement knowledge: 7 components of a good PI story. J. Softw. Evol. Proc. **31**, e2107 (2018)
4. Norimatsu, S., Kishi, T., Wada, N.: Development of "SPI Strategy Framework" and its application. In: Industrial Proceedings, EuroSPI2018 (European Systems Software & Service Process Improvement & Innovation) (2018)
5. The SPI Manifesto. http://2019.eurospi.net/images/eurospi/spi_manifesto.pdf. Alcala EuroSPI 2009

Functional Safety and Cybersecurity

Automotive Cybersecurity Engineering Job Roles and Best Practices – Developed for the EU Blueprint Project DRIVES

Richard Messnarz[1(✉)], Georg Macher[2], Florian Stahl[3], Stefan Wachter[4], Damjan Ekert[1], Jakub Stolfa[5], and Svatopluk Stolfa[5]

[1] ISCN GesmbH, Graz, Austria
rmess@iscn.com
[2] University of Technology Graz, Graz, Austria
[3] AVL, Regensburg, Germany
[4] msg Plaut, Graz, Austria
[5] VSB TUO, Ostrava, Czech Republic
jakub.stolfa@vsb.cz

Abstract. Vehicle functions are 90% controlled by electronic and software thus, modern vehicles need to comply with cybersecurity standards and data security law. Under the EU Blueprint for Automotive - DRIVES project (www.project-drives.eu), key skills and job roles for future Automotive development are designed. These also include cybersecurity management, cybersecurity engineering, and cybersecurity testing. The working party SOQRATES (www.soqrates.de) agreed to develop the pattern of such skills set in cooperation with DRIVES. Also, to share required best practices and skills to be transferred to allow implementing the norms. This paper provides an overview of the current version of the skills set, and an example of what type of best practice is shared.

Keywords: Cybersecurity · DRIVES · Skills set · Best practices

1 Introduction

In this paper the current developments on the new skills for cybersecurity engineering and the background of the EU Blueprint project DRIVES are discussed. An introduction to the job role pool from DRIVES project will be presented. The scrutiny of a cybersecurity engineer job role will be further analysed and described.

In addition, the paper describes the typical steps for cybersecurity threat analysis and cybersecurity requirements analysis in automotive projects, applying norms like SAE J3061 and ISO 21434.

It also outlines how the working group SOQRATES consisting of leading automotive suppliers and experts is cooperating to develop best practices for fields such as cybersecurity related best practices.

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 499–510, 2020.
https://doi.org/10.1007/978-3-030-56441-4_37

2 DRIVES

2.1 Need for New Skills for Future Automotive Design

In order to deliver sector-specific skills solutions, the New Skills Agenda for Europe has launched the action Blueprint for Sectorial Cooperation on Skills [12, 13]. DRIVES project [11, 38] represents an automotive sector and was launched in the first wave of five sectors specific Blueprint projects. The aim is to develop concrete actions to satisfy short- and medium-term skills needs to support the overall sector: DRIVES project foreseen:

1. Implementation of a common European automotive skills reference framework where industry, academia and VET-providers can access and get to the pool of skills and job roles needed to help them to grow;
2. Enabling mutual recognition of automotive skills across Europe by DRIVES digital badge in order to enhance and support the automotive workers and their applicability and comparability of skills across Europe (respectful to the current certification of formal and non-formal training networks);
3. Assessing and inclusion of existing, proven and newly developed trainings for the future job roles and/or needed skills. A first push of emerging job roles in Europe by ca. 1200 attendees in trail courses according to the future DRIVES reference framework scheme;
4. Deployment of the Apprenticeship Marketplace by enhancing its effectiveness for automotive job seekers. Linking to the DRIVES framework of automotive skills and job roles to facilitate dissemination of common job requirements, which will be available for job seekers, training providers (namely universities), VET providers and other stakeholders.

One of the outputs foreseen and mentioned in the paragraph 3 above is the DRIVES Learning Platform. This platform offers courses developed by DRIVES partnership (http://learn.drives-compass.eu). The training materials result from identified reskilling and upskilling needs stemming from the sector after new mobility dynamics and automotive industrial transformations. The DRIVES learning platform is being constantly updated. The first set of courses for learning experience is being prepared. The teaching offer will be partially available online as MOOC courses and courses on demand as on-site courses in partner regions. DRIVES partnership is working on 30+ trainings to be fully available in 2021.

Figure 1 shows the learning compass service platform concept which will become available in 2020 and rolled out EU-wide in 2021.

2.2 Job Role Strategy

In the DRIVES project job role-based reference skills sets are developed and training providers are linked Europe wide with these skills sets [19–21, 23–29, 32, 37].

Figure 2 shows parts of the list of job roles being developed as a pilot skill sets and linked trainings. Each job role is described with skills units (area of skills), skills elements per skill unit (a specific skills topic) and performance criteria per skills element

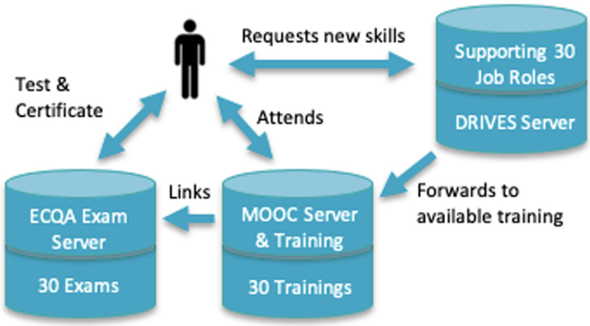


Fig. 1. Proposal of future cooperation training platform scheme - in cooperation with ECQA as a pilot certification body

(a specific ability to do a certain task). The certification for pilot implementation is supported by the ECQA (European Certification and Qualification Association).

3 Cybersecurity Related Job Roles

In cooperation with [10] SOQRATES (www.soqrates.de) starting from 2018 the cybersecurity [30, 31] related competences have been brainstormed, discussed, grouped and documented. In parallel to this, task clauses of SAE J3061, and ISO 21434 [22] have been mapped onto Automotive SPICE to get an understanding how a combined process audit would like.

The competences roadmap for …

- Cybersecurity manager
- Cybersecurity engineer
- Cybersecurity tester

… is shown in Fig. 3 below.

The structure is based on the European skills definition structure of the NVQ (National Vocational Qualification) standards and structures units and elements.

Each element is assigned to a specific cybersecurity job role with a level of competence based on the Bloom's taxonomy of skills levels.

According to Blooms (see the terms in Fig. 3):

- **Comprehension** means that you understand.
- **Application** means that it can be applied in usual projects.
- **Analysis** means that this knowledge can be used in complex situations.

For each element so called performance criteria are defined. Find below 2 examples to understand the term performance criteria.

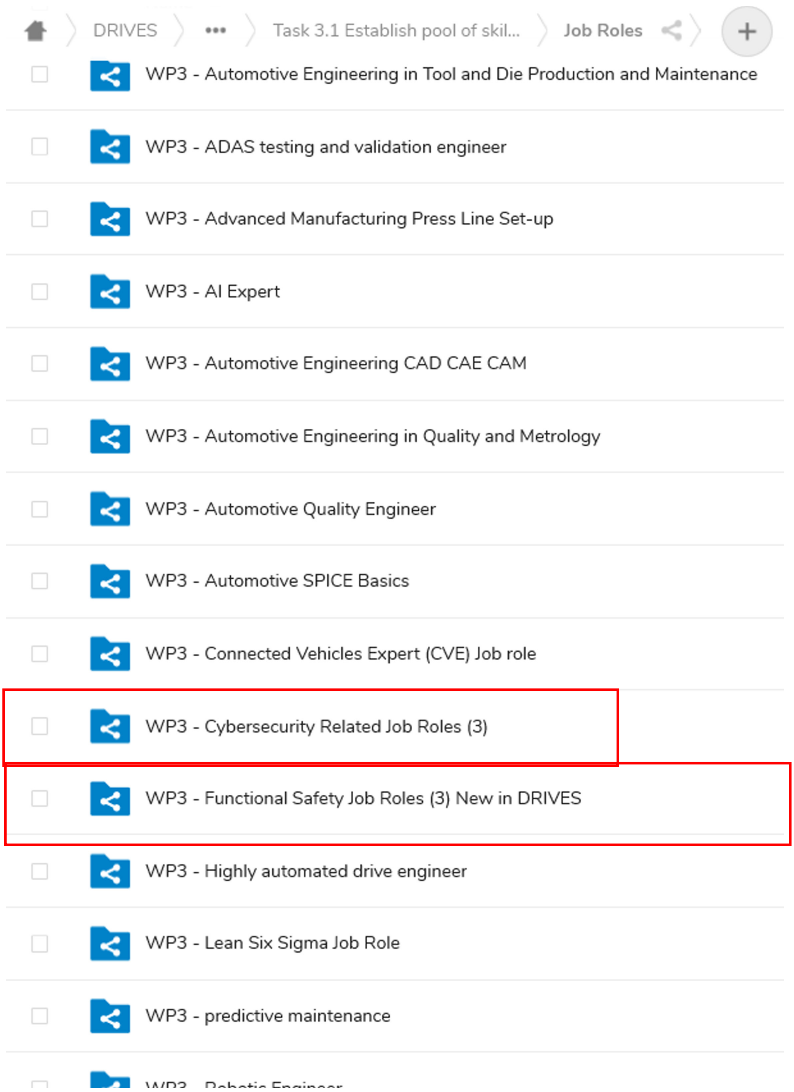


Fig. 2. Example of job roles skills' definitions/trainings being developed in DRIVES

Example 1:

Unit CYBER.U3 - Element 1: System Threat Analysis and Cybersecurity Goals
This element addresses

- Types of attacks
- Known threat lists
- Cybersecurity assets
- Drawing a system item picture including the assets that could be attacked
- Threat and Risk Analysis (TARA)
- Cybersecurity goals

Units (U) and Elements (E) of the skill card	Cybersecurity Engineer	Cybersecurity Manager	Cybersecurity Tester
Unit 1 Cybersecurity Management			
U1.E1 Legal Aspects and Privacy		application	
U1.E2 Organisational Structure		application	
U1.E3 Cybersecurity Planning		application	
Unit 2 Cybersecurity Operation and Maintenance			
U2.E1 Life Cycle Assessment		analysis	
U2.E2 Cybersecurity processes and audits		analysis	
U2.E3 Incident Response Management		analysis	
U2.E4 Supply Chain Security		analysis	
Unit 3 Engineering aspects of cybersecurity			
U3.E1 System Threat Analysis and Cybersecurity Goals	analysis	comprehension	
U3.E2 System Design and Vulnerability Analysis	analysis	comprehension	
U3.E3 Software Design and Vulnerability Analysis	analysis	comprehension	
U3.E4 Software Detailed Design and Cybersecurity	analysis	comprehension	
U3.E5 Cybersecure hardware and firmware design	analysis	comprehension	
Unit 4 Testing aspects of cybersecurity			
U4.E1 Cybersecurity verification at SW level		comprehension	analysis
U4.E2 Cybersecurity verification at HW level		comprehension	analysis
U4.E3 Cybersecurity verification at system level		comprehension	analysis

Fig. 3. Cybersecurity job role competences roadmap divided to units and elements

Performance Criteria:

The student must be able to show evidence of competencies for the following performance criteria (PC) (Table 1):

Table 1. Performance criteria for system threat analysis and cybersecurity goals

Performance criterion	Evidence check: the student can demonstrate
CYBER.U3.E1. PC1	The student knows the different types of attacks
CYBER.U3.E1. PC2	The student is able to access the known threat lists
CYBER.U3.E1. PC3	The student is able to identify and document cybersecurity assets (asset analysis)
CYBER.U3.E1. PC4	The student is able to draw a system item picture showing the system structure and the cybersecurity assets as potential attack targets
CYBER.U3.E1. PC5	The student is able to perform a TARA (Threat and Risk Analysis) and document it
CYBER.U3.E1. PC6	The student is able to derive cybersecurity goals from the TARA (Threat and Risk Analysis) and document them

Example 2:

Unit CYBER.U3 - Element 2: System Design and Vulnerability Analysis

This element looks at the system level related cybersecurity methods.

This includes

- Applying cybersecurity design patterns on system level

- Performing an attack tree analysis
- Performing vulnerability analysis and integrating a proper defence mechanism
- Integrating cybersecurity views into the system architectural design
- Writing cybersecurity requirements

Performance Criteria:
The student must be able to show evidence of competencies for the following performance criteria (PC) (Table 2):

Table 2. Performance criteria for system design and vulnerability analysis

Performance criterion	Evidence check: the student can demonstrate
CYBER.U3.E2. PC1	The student knows cybersecurity design patterns on system level and how to apply them
CYBER.U3.E2. PC2	The student is able to perform an attack tree analysis
CYBER.U3.E2. PC3	The student is able to perform a vulnerability analysis and integrating a proper defence mechanism
CYBER.U3.E2. PC4	The student is able to integrate cybersecurity views into the system architectural design
CYBER.U3.E2. PC5	The student is able to write and trace cybersecurity requirements

4 Cybersecurity Related State of the Art Development

In parallel to the skills set development the SOQRATES working party started to exchange best practices per element and the minimum state of the art is agreed and potential exercises to delivery such a qualification in a training are designed [3–7, 10, 18].

In SOQRATES there is the rule that if a group of partners agrees the same approach it will be declared as state of the art by the group.

The following describes an example outcome of such exchange concerning the above mentioned 2 elements U3.E1 and U3.E2.

Example conclusions from practice:

Threat Model. There is not one threat model, in practice there is a threat model on system level, and for the software a threat model per session (programming session, diagnose session, normal operating session) [15].

Trust Boundaries: Threat model scenarios have a scope (like system design) and such boundaries are also drawn into the threat models.

Assets: When drawing a threat model, the assets list must be integrated to show the potential attack targets and surfaces. When doing a TARA (Threat Analysis and Risk Assessment) also the attack on those assets is evaluated and delivers a risk level.

Cybersecurity Assurance Levels (CAL): Even if the norm defines a CAL to be a qualitative rating, some leading companies use a TL (Threat Level, [14–17]) 1 to 4, the working party has access (by partnership) to the former AVL research results with a method to assign a TL to threats. Thus, cybersecurity goals can be assigned to a TL 1 to 4, the higher the CAL the more methods are applied.

Cybersecure Related SW Functions. There is a defined structure to list and evaluate so called cybersecurity related SW functions, - the threat is described, the potential attack to the SW function is described, and cybersecurity requirements to defend against the attack are written to this function.

Classification as Cybersecurity Affected Data. There is a defined structure to list and evaluate so called cybersecurity related data, - the threat is described, the potential attack to the SW data is described, and cybersecurity requirements to defend against the attack are written to these data.

Traceability: Cybersecurity goals are linked to cybersecurity related functions and cybersecurity related data, and this results in cybersecurity requirements. The cybersecurity requirements are assigned to system, hardware, software and from there copied to an A-SPICE like implementation process with traceability (Fig. 4).

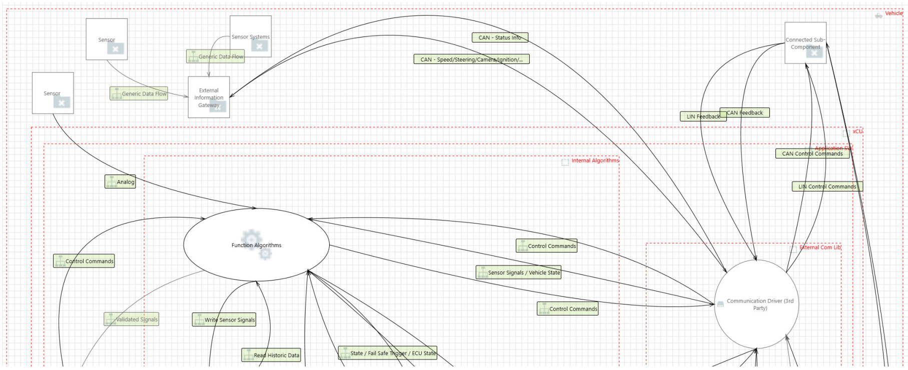


Fig. 4. Example threat model on system level (of a supplier ECU).

The threat model above uses the STRIDE tool and elements are circles and arrows show a potential attack flow from one to the next element. Squares are external interfaces.

Each of the attack flows can be classified by e.g. the type of the attack. See Fig. 5.

Each attack (each arrow on the threat model) and type of attack is considered in the TARA and is classified in the TARA. See the line called "Threat Level Assessment" in Fig. 6.

Each line in the threat level assessment results in a classification with a threat level, a description of the threat, and typically (if applicable) a safety rating, if it has Automotive Safety Integrity Level (ASIL) impact.

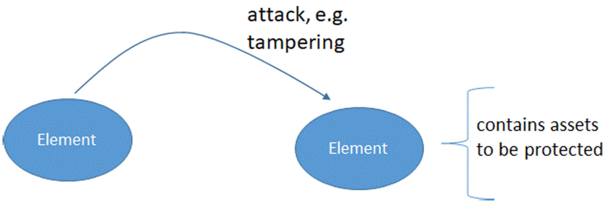


Fig. 5. Attack flow with a classification of attack type

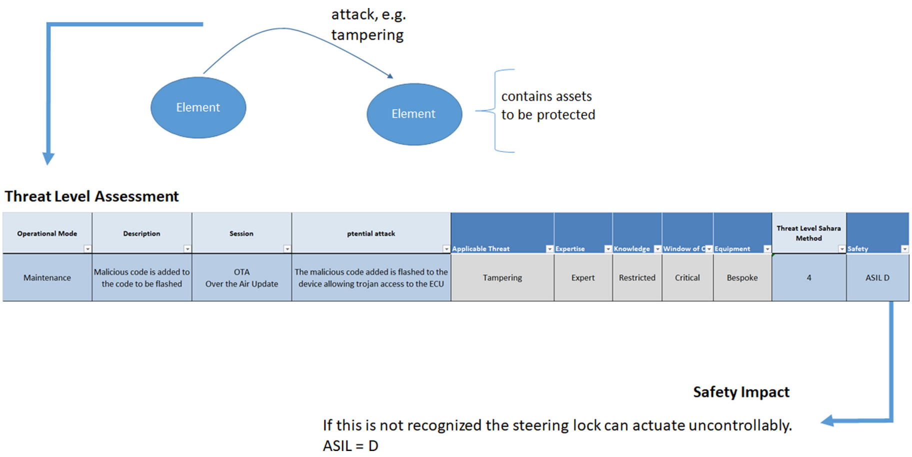


Fig. 6. Mapping of potential attack flows to lines in the Threat Level Assessment

The counteraction to overcome the security threat leads to so called either (1) a cybersecurity function (protection function) or (2) cybersecurity relevant data (measures to protect specific data assets). See Fig. 7.

And for each of the counter measures requirements are written on system, software and hardware level.

Figure 7 also shows how the analysis leads to requirements and how they are linked and integrated to the traceability concept of Automotive SPICE [1, 2, 8] typical specifications.

These examples are not to be seen as the one and only solution but in SOQRATES potential solutions that can cover the aspects of the skills set are agreed and documented for the partnership.

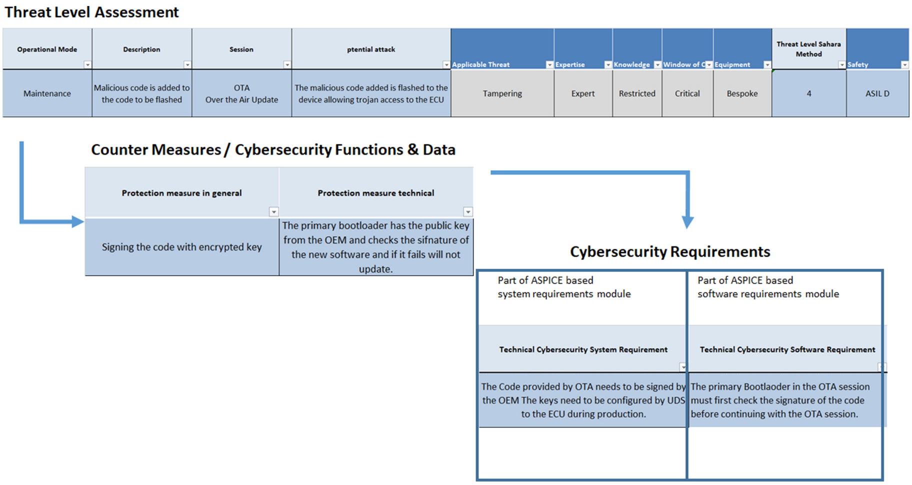


Fig. 7. Traceability to requirements and interface to Automotive SPICE

5 Relationship with the SPI Manifesto

A platform were such new cross-cutting approaches can be discussed is EuroAsiaSPI[2]. Its mission is to develop an experience and knowledge exchange platform for Europe where Software Process Improvement (SPI) practices can be discussed and exchanged and knowledge can be gathered and shared [33–36]. The connected SPI manifesto defines the required values and principles for a most efficient SPI work. One main goal is to support changes by innovation and include all affected stakeholders.

The principle "**Use dynamic and adaptable models as needed**" means that cybersecurity norms and views in future need, to be integrated into the existing processes.

Another important platform for such new cross-cutting approaches is the European DRIVES project. DRIVES is a BLUEPRINT [11, 12, 38] project for the automotive industry and creates a strategy that supports the development of new business visions for 2030. It will also emphasise the combined use of different norms.

Acknowledgements. We are grateful to the European Commission which has funded the Blueprint project Development and Research on Innovative Vocational Skills (DRIVES, 2018-2021, www.project-drives.eu). The project DRIVES – Project number 591988-EPP-1-2017-1-CZ-EPPKA2-SSA-B is co-funded by the Erasmus + Programme of the European Union.

The European Commission support for the production of this publication under the Grant Agreement Nº 2017-3295/001-001 does not constitute an endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Work is partially supported by Grant of SGS No. SP2020/62, VŠB - Technical University of Ostrava, Czech Republic.

We are grateful to a working party of Automotive suppliers SOQRATES (www.soqrates.de) who exchange knowledge about such assessment strategies. This includes: Böhner Martin (Elektrobit), Brasse Michael (HELLA), Bressau Ernst (BBraun), Dallinger Martin (ZF), Dorociak Rafal (HELLA), Dreves Rainer (Continental Automotive), Ekert Damjan (ISCN), Forster Martin (ZKW), Geipel Thomas (BOSCH), Grave Rudolf (Elektrobit), Griessnig Gerhard (AVL), Gruber Andreas (ZKW), Habel Stephan (Continental Automotive), Hällmayer Frank (Software Factory), Haunert Lutz (Giesecke & Devrient), Karner Christoph (KTM), Kinalzyk Dietmar (AVL), König Frank (ZF), Lichtenberger Christoph (MAGNA ECS), Lindermuth Peter (Magna Powertrain), Macher Georg (TU Graz & ISCN), Mandic Irenka (Magna Powertrain), Maric Dijas (Lorit Consultancy), Mayer Ralf (BOSCH Engineering), Mergen Silvana (TDK/EPCOS), Messnarz Richard (ISCN), Much Alexander (Elektrobit), Nikolov Borislav (msg Plaut), Oehler Couso Daniel (Magna Powertrain), Riel Andreas (Grenoble INP & ISCN), Rieß Armin (BBraun), Santer Christian (AVL), Schlager Christian (Magna ECS), Schmittner Christoph (Austrian Institute of Technology AIT), Schubert Marion (ZKW), Sechser Bernhard (Process Fellows), Sokic Ivan (Continental Automotive), Sporer Harald (Infineon), Stahl Florian (AVL), Wachter Stefan (msg Plaut), Walker Alastair (Lorit Consultancy), Wegner Thomas (ZF).

References

1. Automotive SPICE © 3.1. Process Assessment Model, VDA QMC Working Group
2. Automotive SPICE © Guidelines. 2nd Edition Nov 2017, VDA QMC Working Group 13, November 2017
3. Macher, G., Sporer, H., Brenner, E., Kreiner, C.: Supporting cyber-security based on hardware-software interface definition. In: Kreiner, C., O'Connor, R.V., Poth, A., Messnarz, R. (eds.) EuroSPI 2016. CCIS, vol. 633, pp. 148–159. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-44817-6_12
4. Macher, G., Messnarz, R., Kreiner, C., et. al.: Integrated Safety and Security Development in the Automotive Domain, Working Group 17AE-0252/2017-01-1661. SAE International, June 2017
5. Macher, G., Much, A., Riel, A., Messnarz, R., Kreiner, C.: Automotive SPICE, safety and cybersecurity integration. In: Tonetta, S., Schoitsch, E., Bitsch, F. (eds.) SAFECOMP 2017. LNCS, vol. 10489, pp. 273–285. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66284-8_23
6. Messnarz, R., Kreiner, C., Macher, G., Walker, A.: Extending automotive SPICE 3.0 for the use in ADAS and future self-driving service architectures. J. Softw. Evol. Process **30**(5), e1948 (2018)
7. Messnarz, R., Kreiner, C., Riel, A.: Integrating automotive SPICE, functional safety, and cybersecurity concepts: a cybersecurity layer model. Softw. Qual. Prof. **18**, 13 (2016)
8. Messnarz, R., Sehr, M., Wüstemann, I., Humpohl, J., Ekert, D.: Experiences with SQIL – SW quality improvement leadership approach from Volkswagen. In: Stolfa, J., Stolfa, S., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2017. CCIS, vol. 748, pp. 421–435. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64218-5_35
9. SAE J3061. Cybersecurity Guidebook for Cyber-Physical Vehicle Systems, SAE - Society of Automotive Engineers, USA, January 2016
10. SOQRATES. Task Forces Developing Integration of Automotive SPICE, ISO 26262 and SAE J3061. http://soqrates.eurospi.net/
11. EU Blueprint Project DRIVES. https://www.project-drives.eu/. Accessed 17 Apr 2020
12. GEAR 2030. High Level Group on the Competitiveness and Sustainable Growth of the Automotive Industry in the European Union (2017)
13. European Sector Skill Council: Report, Eu Skill Council Automotive Industry (2013)

14. Macher, G., Sporer, H., Berlach, R., Armengaud, E., Kreiner, C.: SAHARA: a security-aware hazard and risk analysis method. In: Design, Automation Test in Europe Conference Exhibition (DATE), pp. 621–624, March 2015
15. Microsoft Corporation. The STRIDE threat model (2005)
16. Macher, G., Armengaud, E., Brenner, E., Kreiner, C.: A review of threat analysis and risk assessment methods in the automotive context. In: Skavhaug, A., Guiochet, J., Bitsch, F. (eds.) SAFECOMP 2016. LNCS, vol. 9922, pp. 130–141. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-45477-1_11
17. Macher, G., Höller, A., Sporer, H., Armengaud, E., Kreiner, C.: A comprehensive safety, security, and serviceability assessment method. In: Koornneef, F., van Gulijk, C. (eds.) SAFECOMP 2015. LNCS, vol. 9337, pp. 410–424. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24255-2_30
18. Riel, A., Kreiner, C., Messnarz, R., Much, A.: An architectural approach to the integration of safety and security requirements in smart products and systems design. CIRP Ann. **67**(1), 173–176 (2018)
19. Riel, A., et al.: EU project SafEUr – competence requirements for functional safety managers. In: Winkler, D., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2012. CCIS, vol. 301, pp. 253–265. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-31199-4_22
20. Riel, A., Tichkiewitch, S., Messnarz, R.: The profession of integrated engineering: formation and certification on a European level. Acad. J. Manuf. **6**, 6–13 (2008)
21. Riel, A., Draghici, A., Draghici, G., Grajewski, D., Messnarz, R.: Process and product innovation needs integrated engineering collaboration skills. J. Softw. Evol. Process **24**(5), 551–560 (2012)
22. ISO/SAE 21434. Road vehicles – Cybersecurity engineering, ISO and SAE, Committee Draft (CD) (2018)
23. Christian, K., Messnarz, R., Riel, A., et al.: The AQUA automotive sector skills alliance: best practice in an integrated engineering approach. Softw. Qual. Prof. **17**(3), 35–45 (2015). 11 p.
24. Messnarz, R., et al.: Integrating functional safety, automotive SPICE and Six Sigma – the AQUA knowledge base and integration examples. In: Barafort, B., O'Connor, R.V., Poth, A., Messnarz, R. (eds.) EuroSPI 2014. CCIS, vol. 425, pp. 285–295. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-662-43896-1_26
25. Kreiner, C., et al.: Automotive knowledge alliance AQUA – integrating automotive SPICE, Six Sigma, and functional safety. In: McCaffery, F., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2013. CCIS, vol. 364, pp. 333–344. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-39179-8_30
26. Messnarz, R., Spork, G., Riel, A., Tichkiewitch, S.: Dynamic learning organisations supporting knowledge creation for competitive and integrated product design. In: Proceedings of the 19th CIRP Design Conference – Competitive Design, 30–31 March 2009, p. 104. Cranfield University (2009)
27. Messnarz, R., Kreiner, C., Riel, A., et.al.: Implementing functional safety standards has an impact on system and SW design - required knowledge and competencies (SafEUr). Software Quality Professional (2015)
28. Messnarz, R., et al.: Implementing functional safety standards – experiences from the trials about required knowledge and competencies (SafEUr). In: McCaffery, F., O'Connor, Rory V., Messnarz, R. (eds.) EuroSPI 2013. CCIS, vol. 364, pp. 323–332. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-39179-8_29

29. Messnarz, R., König, F., Bachmann, V.O.: Experiences with trial assessments combining automotive SPICE and functional safety standards. In: Winkler, D., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2012. CCIS, vol. 301, pp. 266–275. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-31199-4_23
30. Messnarz, R., Much, A., Kreiner, C., Biro, M., Gorner, J.: Need for the continuous evolution of systems engineering practices for modern vehicle engineering. In: Stolfa, J., Stolfa, S., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2017. CCIS, vol. 748, pp. 439–452. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64218-5_36
31. Much, A.: Automotive Security: Challenges, Standards and Solutions. Softw. Qual. Prof. **18**, 4–12 (2016)
32. Stolfa, J., et al.: Automotive quality universities - AQUA alliance extension to higher education. In: Kreiner, C., O'Connor, R.V., Poth, A., Messnarz, R. (eds.) EuroSPI 2016. CCIS, vol. 633, pp. 176–187. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-44817-6_14
33. Korsaa, M., et al.: The SPI manifesto and the ECQA SPI manager certification scheme. J. Softw. Evol. Process **24**(5), 525–540 (2012)
34. Korsaa, M., et al.: The people aspects in modern process improvement management approaches. J. Softw. Evol. Process **25**(4), 381–391 (2013)
35. Messnarz, R., et al.: Social responsibility aspects supporting the success of SPI. J. Softw. Evol. Process **26**(3), 284–294 (2014)
36. Messnarz, R., Ekert, D.: Assessment-based learning systems - learning from best projects. Wiley Inersci. Softw. Process Improv. Pract. **12**(6), 569–577 (2007). https://doi.org/10.1002/spip.347. Special Issue on Industrial Experiences in SPI
37. Messnarz, R., Ekert, D., Zehetner, T., Aschbacher, L.: Experiences with ASPICE 3.1 and the VDA automotive SPICE guidelines – using advanced assessment systems. In: Walker, A., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 549–562. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_42
38. Stolfa, J., et al.: DRIVES—EU blueprint project for the automotive sector—a literature review of drivers of change in automotive industry. J. Softw. Evol. Process **32**(3), e2222 (2020). Special Issue: Addressing Evolving Requirements Faced by the Software Industry

Achieving Data Privacy with a Dependability Mechanism for Cyber Physical Systems

Gilbert Regan[1(✉)], Fergal Mc Caffery[1], Pangkaj Chandra Paul[1], Jan Reich[2], Ioannis Sorokos[2], Eric Armangeud[3], Marc Zeller[4], and Simone Longo[5]

[1] Lero @DKIT, Dundalk, Ireland
{Gilbert.regan,Fergal.mccaffery, Pangkajchandra.paul}@dkit.ie
[2] Fraunhofer IESE, Kaiserslautern, Germany
{Jan.reich,Ioannis.sorokos}@iese.fraunhofer.de
[3] AVL List GmbH, Graz, Austria
Eric.armengaud@avl.com
[4] Siemens, Munich, Germany
Marc.zeller@siemens.com
[5] General Motors, Turin, Italy
Simone.longo@gm.com

Abstract. Cyber-Physical-Systems (CPS), such as smart cars and implanted medical devices, are systems of collaborating computational entities. The open and cooperative nature of CPS poses a significant new challenge in assuring dependability. The DEIS project addresses this important and unsolved challenge through its key innovation which is the concept of a **D**igital **D**ependability **I**dentity (DDI). A DDI contains all the information that uniquely describes the dependability characteristics of a CPS or CPS component. Data protection and privacy is a key component of dependability and is regulated by the General Data Protection Regulation (GDPR) for all European Union (EU) and European Economic Area (EEA) citizens.

In this paper we present an overview of the DDI. Additionally, we provide our concept of how the DDI can support assurance of technical system security requirements derived from GDPR. Additionally, we demonstrate how this concept is implemented in an automotive use case.

Keywords: Dependability · Cyber Physical System · Data privacy · Cyber security · GDPR

1 Introduction

Within Europe support for development and integration of Cyber-Physical Systems and the Internet of Things is seen as essential for the future. As the embedded world meets the Internet world there will be an increasing number of interacting systems with strong connectivity utilized in both society and in industry. Connectivity between embedded systems and computing devices is predicted to grow massively– there are 2.5 quintillion bytes of data created each day, projected 200 billion of smart devices by 2020 [1]. The

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 511–524, 2020.
https://doi.org/10.1007/978-3-030-56441-4_38

main trends are directed towards increase of connectivity and collaborative functions [2], the establishment of trust including safety and cyber-security [3], and the integration of artificial intelligence algorithms [4, 5].

The integration of personal data is growing fast. This can be due to more advanced human – machine collaborations (e.g., cognitive systems, collaboration with robots), or to monitoring and control of processes with human interactions (e.g., monitoring of an oven in a production line). The general data protection regulation (GDPR) [6] has been developed in this context to provide a legal framework enabling the processing of data with appropriate protection of natural persons.

Assuring dependability of CPS is the core challenge of the DEIS project [7]. Dependability is qualitatively defined as 'the ability to deliver service that can justifiably be trusted', and quantitatively defined as 'the ability to avoid service failures that are more frequent and more severe than is acceptable to its user(s)' [8]. The DDI targets (1) improving the efficiency of generating consistent dependability argumentation over the supply chain during design time, and (2) laying the foundation for runtime certification of ad-hoc networks of embedded-systems. This paper presents the concept of how the DDI can assure GDPR requirements for data privacy. The concept is demonstrated through an industrial use case from the automotive domain. This paper is organized as follows: Sect. 2 presents related works on GDPR and privacy assessment. An overview of the DDI is presented in Sect. 3, and the research methodology is introduced in Sect. 4. In Sect. 5, we present our concept of how the DDI can assure compliance with the GDPR technical security requirements, while in Sect. 6 we exemplify GDPR assurance in one of our DEIS industrial partner's use case. Finally, Sect. 7 concludes this work.

2 Related Work

The GDPR is a recent development in EU law regulating the protection of data privacy of all EU and EEA citizens, both within and outside the geographical boundaries of those areas [9]. GDPR encompass not only technical requirements for systems which collect and process citizen data but also the intent, processes and management applied by the organisations operating such systems. GDPR compliance requires data controllers to implement adequate data protection mechanisms as part of their data processing systems. Article 24 explicitly states the controller's responsibility to implement appropriate measures and policies to address risks involved in the data processing. Article 25, paragraph 1, further specifies that the design and implementation of these measures must take place as early as the determination of the means of processing, aka 'data protection by design'.

A privacy-focused threat modelling methodology can be used to identify the privacy-related risks of a CPS system. Cloud Privacy Threat Modeling (CPTM) [10], LINDDUN [11], and Privacy Risk Analysis Methodology (PRIAM) [12] are widely used privacy-centric threat modelling methodologies. CPTM is focused on eliciting and addressing privacy requirements for cloud computing environments; however the approach could be adapted to other application domains as well. The approach fits with the typical development lifecycle applied in CPS development as well as with the assurance process framework supported by the ODE and DDI. The PRIAM application encompasses two phases. In the first phase, the analyst identifies key qualitative and quantitative information relevant to the system's privacy, for example harmful events (harms). In the second phase, the risk level of each identified harm is evaluated as the combination of the likelihood of occurrence and severity of worst-case impact. PRIAM is relatively easy of adopt and apply by the use case owners, who are not security experts. LINDDUN is a privacy engineering framework applicable to generic systems analysis. The framework focuses on identifying threats categorised under one of its acronym's categories; linkability, identifiability, non-repudiation, detectability, disclosure of information, unawareness and non-compliance. LINDDUN's approach is more direct when compared to CPTM and PRIAM, directly targeting the above privacy properties.

With this context in mind, GDPR compliance is supported by DDIs via assurance of technical CPS requirements during development and assurance of protection of GDPR personal data rights during operation. The former aspect involves performing adequate risk analysis and mitigation of privacy risks associated with personal data of EU citizens as part of the CPS' operation. The latter aspect involves assuring that the chain of systems implementing the data processing service adequately protects the data rights of citizens.

3 Overview of DDI

A Digital Dependability Identity (DDI) [13] captures the various artifacts of the safety lifecycle in a model-based way and to establish a relationship between the argumentation and the supporting evidence models. By establishing traceability across the artifacts, DDIs represent an integrated set of dependability data models ("What is the evidence data?") that are generated by engineers and are reasoned upon in dependability arguments ("How is the evidence data supporting the claim?"). A DDI is an evolution of the classical modular dependability assurance model. In a DDI, several separately defined dependability aspect models are formally integrated which allows for comprehensive dependability reasoning. DDIs are produced during the design of a component or system, certified when the component or system is released, and then, continuously maintained over the lifetime of the component or system.

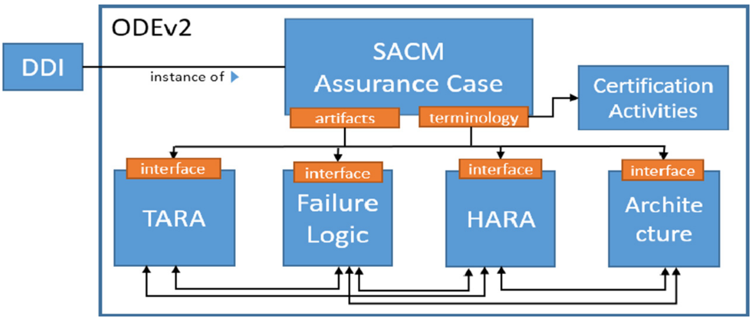


Fig. 1. Digital Dependability Identity (DDI)

Figure 1 illustrates the concept of traceability between a *Structured Assurance Case Metamodel (SACM)* [14] assurance argument and dependability evidence models, SACM provides the assurance case backbone for creating the required traceability. Moreover, it is formally relating evidence models to the assurance argument. The DDI meta-model formalizing the traceability and evidence semantics is the so-called *Open Dependability Exchange (ODE)* meta-model. For security assurance, it contains e.g. threat and risk analyses (TARA) and attack trees, while for safety assurance, hazard and risk analyses (HARA), architecture modeling and failure propagation modeling such as fault trees, FMEA or Markov chains are supported.

Based on the DDI concept, it is possible to automate the safety argumentation of a system. Since the DDI represents both the argumentation in a model-based (using SACM) as well as the respective artifacts created during the system development (e.g. requirements, functional architecture, etc.) and the safety engineering life-cycle (e.g. hazard lists, safety analysis results, etc.), relationships between argumentation and the corresponding artifacts can be automatically established [15]. Details on the DDI framework as well as an open-source version of the ODE meta-model can be found at Github [16].

4 Research Methodology

Over the course of DEIS, the consortium partners recognised that addressing privacy concerns for the use cases being examined would demand more tailoring in the threat analysis, requirements elicitation and countermeasure design compared to addressing generic confidentiality, integrity and availability concerns. Assuring compliance with GDPR in particular involves establishing claims regarding the provision of specific technical security capabilities e.g. security by default and by design and data protection rights. While such claims can be constructed and managed using the SACM facilities that are integrated in the ODE, supporting them with adequate evidence requires the employment of appropriate privacy threat modeling, requirement elicitation, privacy capability identification, design and implementation.

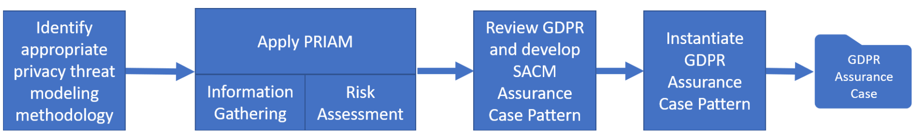


Fig. 2. Research methodology overview

In Fig. 2, an overview of the research methodology followed by the DEIS partners can be seen. The TARA package of the DDI's ODE metamodel is generic and can be adapted to application-specific security threat analysis methods. In the first stage of Fig. 1, we reviewed privacy-specific threat analysis methodologies, including CPTM, LINDDUN and PRIAM. Out of the 3 methodologies reviewed, PRIAM was selected due to its ease of adoption and application by the use case owners, which were not security experts. PRIAM provides detailed guidance and examples of tables for its information gathering phase, as well as comprehensive diagrams for its risk assessment phase. The separation of the process into the two phases was also beneficial for managing the complexity of the tasks involved for novices in security analysis.

In the second stage of Fig. 2's process, the use case owners applied the two PRIAM phases with support from the consortium security experts. For each use case, information gathering artefacts (data tables, data flow diagrams) and risk assessment artefacts (harm trees, see below) were produced and reviewed. For the risk assessment phase of PRIAM, a variant of the Fault Tree Analysis technique, typically applied in reliability and safety analysis, is adapted as Harm Tree Analysis. An example of harm tree analysis with DDI and tool support can be found in [17]. The results of the PRIAM were then captured in a DDI. Towards this end, a mapping from the PRIAM to the ODE's TARA was also defined in Table 1 under Sect. 5.

Having the necessary evidence for supporting privacy and GDPR claims, those claims were expressed as SACM assurance cases. In the interest of re-usability, assurance case patterns can be employed to abstract the details of a specific system from the argument, defined for GDPR in the fourth stage of Fig. 2. The pattern can then be re-applied to the DDI of a different system; see Sect. 6 for more details. The DDI capturing the system model and PRIAM analysis results was used to instantiate the GDPR pattern for each use case, as per the final stage of Fig. 2.

5 GDPR Compliance Through DDI

The DDI can encompass GDPR privacy requirements in a holistic assurance case, which contains the argumentative basis for the satisfaction of technical system security requirements derived from GDPR. Design time DDI assurance involves performing and documenting adequate risk analysis and mitigation of privacy risks associated with personal data. At runtime operation, DDIs are used to ensure that the chain of systems implementing the data processing service adequately protects the data rights.

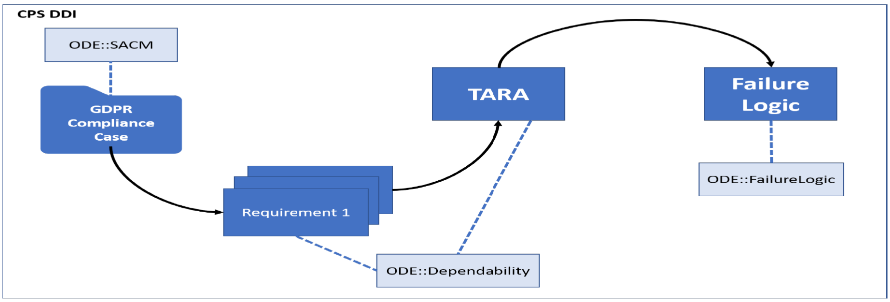


Fig. 3. GDPR compliance via DDI

Figure 3 demonstrates how the DDI packages form a continuously traceable assurance case to assure GDPR risk assessment and adequate allocation of GDPR security requirements. The backbone of the assurance case is the SACM package of the ODE, which contains an argument clearly tracing GDPR requirements to system measures realizing them. For the purposes of GDPR compliance assurance, GDPR argument patterns are combined with the information from the ODE ProductPackages of a CPS DDI (ODE::Dependability, ODE::FailureLogic). The supporting evidence for both SACM assurance arguments can be partially provided by a specialized risk analysis that targets concerns for security of privacy. PRIAM is one such approach to identify, diagnose and mitigate CPS privacy risks. PRIAM consists of two phases, information gathering and risk assessment. During the first phase, analysts probe the system to identify qualitative and quantitative information on key properties that are relevant to security of data privacy. Under PRIAM, the system and its operational context is viewed as a combination of the following seven components: *System*, *Stakeholders*, *Data*, *Risk Sources*, *Privacy Weaknesses*, *Feared Events*, and *Privacy Harm*.

A PRIAM information gathering phase typically begins by identifying the System and its Stakeholders. PRIAM components are further parameterized via categories and attributes. Categories identify specific types of components that are relevant or present in the subject system. For instance, potential Data categories include 'health data', 'location data', 'financial data' etc. Attributes identify specific aspects of components. For example, Data attributes can include precision, retention and delay.

Upon completion of the information gathering phase, the target of the risk assessment phase is to compute the risk level i.e. severity and likelihood of occurrence for each Privacy Harm. Towards this end, PRIAM adopts an approach equivalent to Fault Trees or Attack Trees, named Harm Trees. Harm Trees can be analyzed qualitatively to identify combinations of weaknesses and/or risk sources that can lead to harms and quantitatively to compute harm risk levels. After the Harm Tree analysis, security requirements are expressed in the DDI that mitigate the identified risk causes. These requirements are allocated to system design elements realizing them. Thus, DDIs offer the possibility to provide a continuous traceability from GDPR articles to their realization in system elements, which greatly simplifies change impact analyses as well as provides a transparent means to check GDPR compliance automatically.

A SACM pattern can be defined to capture the satisfaction of technical requirements necessary by GDPR. This pattern can use the PRIAM analysis and produced artifacts as evidence, to assure that the risk levels across all privacy harms have been mitigated or eliminated by appropriate protection measures. The SACM pattern can be incorporated into a DDI alongside the above information transformed from PRIAM. PRIAM is effectively a combination of a specialized TARA and specialized SATs, respectively performed via its information gathering and risk assessment phases. Thus, it conveniently maps directly to the security data structures in the ODE. In Table 1, a mapping of significant elements from PRIAM to the ODE's TARA package can be seen. Using this mapping, automated processes can be developed to transform input from tool support for PRIAM into DDIs.

Table 1. - PRIAM to ODE::TARA mapping

PRIAM element	TARA element
Privacy harm	Attacker Goal/Security Risk
Feared event	Attack
Privacy weakness	Vulnerability/Security Violation
Risk source	Threat Agent
System::Controls	Security Capability/Security Control
System	Asset/Vulnerable Item

6 Exemplification in Automotive

This section describes the Assurance case pattern and the output of PRIAM for one of the DEIS industrial partner's (General Motors (GM)) use case. The Dependable Physiological Monitor System (DPMS) use case describes the sensing environment inside the vehicle that monitors the health condition of drivers and passengers. In the case of the drivers' health deteriorating, a high level emergency manager feature makes mitigating decisions depending on the severity of the deterioration in the drivers health, for example taking control of the car and parking in a lay-by, or notifying emergency responders. As the system is designed to communicate over open channels and manage sensitive and personally identifiable information, security and privacy must also be assured.

6.1 DPMS Assurance Case Pattern

An example of the assurance case pattern instantiated for the DPMS use case can be seen in Fig. 4. The argument presented in the figure is a 'module', an isolated argument unit. The argument is linked to other modules to separate their scope, ease management and maintain clarity.

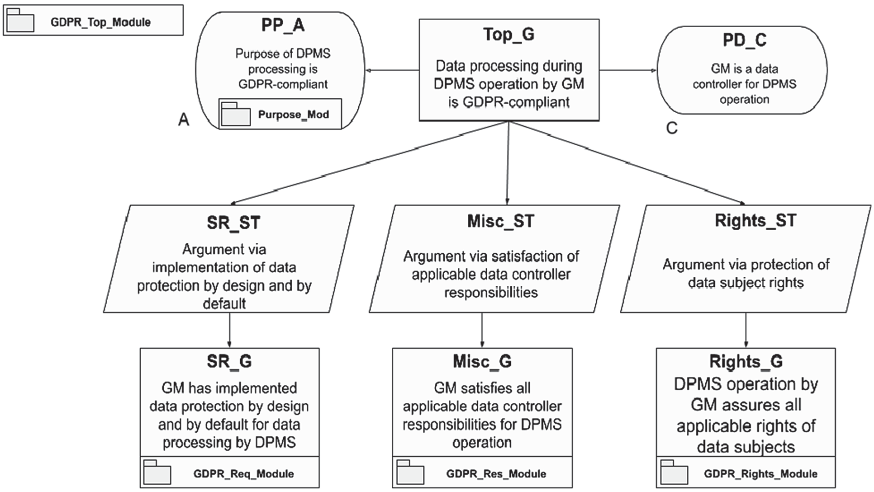


Fig. 4. Top-level argument for GDPR compliance of DPMS operation

At the top of the figure, a claim regarding the compliance of the DPMS with GDPR regulations is established. Within the argument presented in Fig. 4, matters concerning the lawfulness of the data processing purpose are addressed by the 'PP_A' assumption. PP_A claims the purpose of processing of the DPMS is compliant with the relevant GDPR articles (1–3 and 5–11). While PP_A is an assumption within the scope of the 'GDPR_Top_Module' argumentation, it clearly requires further explanation and justification. Thus, PP_A is represented as an external reference to an element in the 'Purpose_Mod' argument module. The overall claim of Fig. 4 is supported in 3 ways:

1. The first strategy claims that the processing performed as part of the DPMS operation has security implemented by design and by default, as per the relevant GDPR articles 24 and 25. An explicit claim that GM, as the developer of DPMS, has implemented appropriate measures further supports this argument. As this goes beyond the scope of the high-level claim in the current module, it is linked as an external claim 'GDPR_SR_G', found in module 'GDPR_Req_Module'.
2. The second argument strategy aims to address additional responsibilities on the part of the data controller (per GDPR articles 30, 31, 33, 34 and 37 to 39). As before, given the scope of the involved supporting claims, these are addressed in a separate argument module.
3. The final aspect assures that GDPR data subject rights are protected as part of the DPMS data processing, per GDPR articles 15 to 22.

Expanding on the line of argumentation under (1) above and from element 'SR_G' in Fig. 4, the approach shown in Fig. 5 can be used.

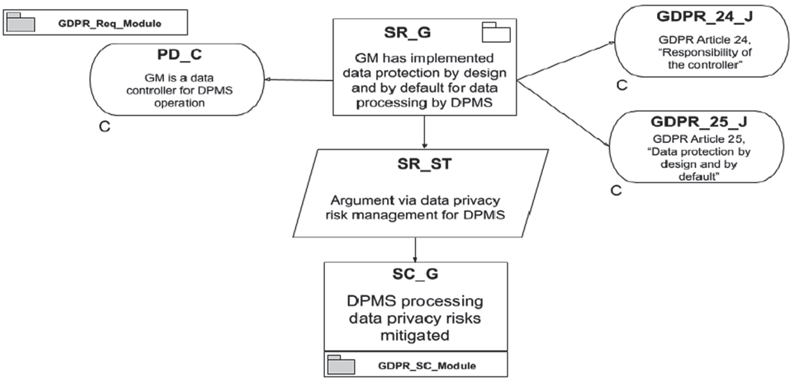


Fig. 5. Data protection argument for DPMS

Using this line of argument, the relevant GDPR articles are satisfied on the condition that all significant data privacy risks are addressed i.e. mitigated appropriately, per 'GDPR_SC_G'. In Fig. 6, the argument for depicting how privacy risks during processing have been addressed is shown. PRIAM is applied to identify significant privacy risks and address them appropriately. Two example (claims regarding) harms are included in the argument, 'GDPR_ID_G' and 'GDPR_Dec_G'. The former relates to the general risk of identifying the data subject during DPMS operation. The latter refers to the specific risk of inadvertently transmitting over an open channel unencrypted personal data during operation. Appropriate means for addressing each risk are explained within each argument module respectively. The results of PRIAM contribute to this argument by allowing identification of the relevant harms and linking them with systemic weaknesses that are addressed via appropriate measures. The contextual element 'PRIAM_C' links to the PRIAM guidelines [2.3], 'PRIAM_A' argues that PRIAM is equivalent to the assessment required by GDPR and 'PRIAM_A2' argues that all significant security harms have been identified. 'PRIAM_A' and 'PRIAM_A2' could be further argued in separate modules but not in this example, due to its simplicity.

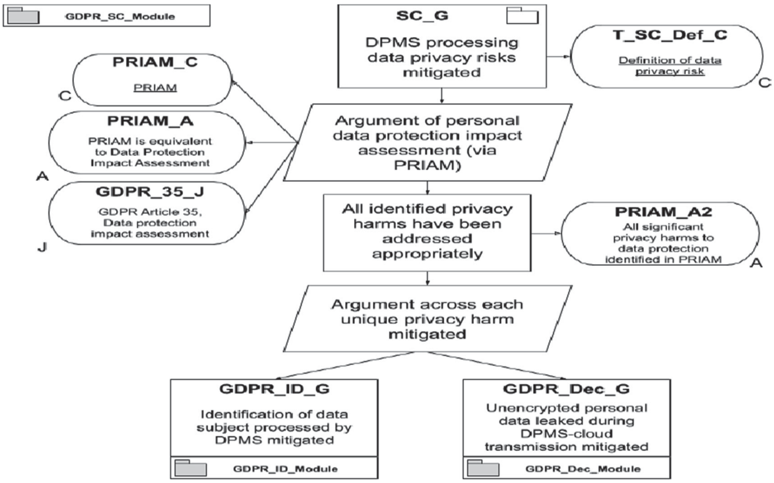


Fig. 6. PRIAM data privacy argument for DPMS

6.2 Application of PRIAM to DPMS Use Case

In Fig. 7, a data flow model depicting the data processing and system architecture involved in the DPMS system is shown. Using this model, analysts can more easily identify and correlate PRIAM elements such as privacy harms, risk sources etc.

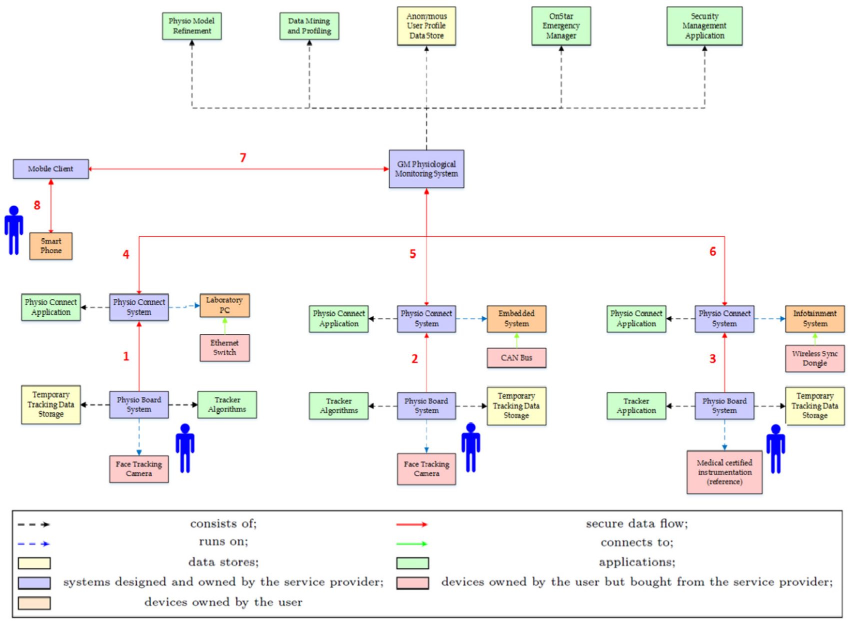


Fig. 7. Applying PRIAM for DPMS - Data flow model

These elements are gradually identified, refined and compiled, as indicated in Tables 2, 3, 4 and 5 below. Due to page limitations, only a subset of each table is shown.

Table 2 contains DPMS *Harms to Privacy*. Privacy Harm is the negative impact on data subjects or society due the effects of feared events being realized.

Table 2. Harms

Harm	Example of event	Categories	Victims	Intensity	Severity
H.1	Denial of a job due to inferred health problems	Psychological, financial	Low	High	High
H.2	Increased premium for health insurance	Financial	Medium	High	Maximum

Table 3 contains DPMS *Feared Events*. Feared Events are systemic events, resulting from exploitation of privacy weakness that may lead to privacy harm.

Table 3. Feared events and their attributes

Code	Feared events	Relevant scenarios	Scale	Irreversibility
FE.1	Excessive collection of physiological data	Collection of user's physiological related data at a higher volume or precision than the user consented for	High	Medium
FE.2	Use of physiological and identification data for unauthorized purpose	Targeted advertising, sending offers about car insurance policy	High	Medium
FE.3	Excessive data inference from physiological and profile data	Health risk profiling	High	High
FE.4	Disclosure of identifiable physiological related data to unauthorized actors or public	Publication of driving patterns online, selling user health profiles to employers, data brokers or insurance companies without consent	High	High

Table 4 contains DPMS *Sources of Risk*. Sources of risk are individuals or organizations processing data and causing direct or indirect harm to data subjects.

Table 4. Sources of risks

Code	Category	Individual-organisation	Insider-outsider	Value of exploitation	Motivation/fear	Background information	Access rights to personal data	Relationship with data subject
A.1	Service provider rogue	Organisation	Insider	High	High	Extensive system knowledge	All data	Semi-trusted
A.4	Third party	Organisation	Outsider	High	High	Publicly available data, some knowledge about the user	Derived data	None

Table 5. Privacy weakness and exploitability

Code	Privacy weakness	Exploitability
V.2	Unencrypted physiological data in live mode	Medium
V.5	Insufficient system audit	High
V.6	FTC sends biometric unencrypted data via serial bus	Medium
V.9	Weak anonymization of data	Medium
V.22	Not prohibiting re-identification through contracts with third parties	High

Table 5 contains *Privacy Weakness* and their level of exploitability. Privacy Weaknesses are functional or implementation weaknesses in data protection mechanisms or the absence of such mechanisms, leading to privacy harm. At a later stage in the PRIAM process harm trees are constructed; one such example can be seen in Fig. 8. The tree depicts the relationship between harms, feared events, sources of risk and weaknesses. By safeguarding against the combinations of sources of risk and weaknesses listed across harm trees, developers can mitigate risks for those harms appropriately.

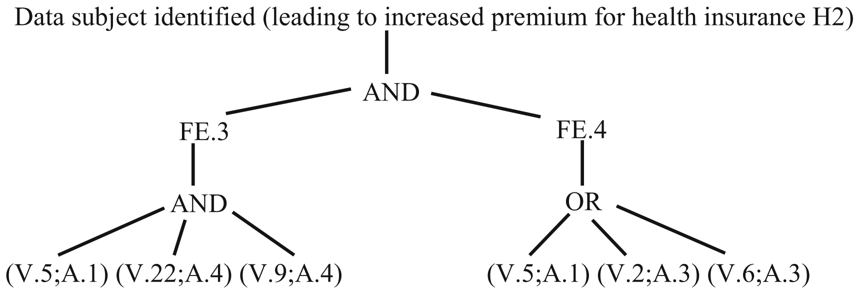


Fig. 8. Example harm tree for DPMS

7 Conclusion

In this paper we demonstrated our concept of how the DDI can support assurance of GDPR technical data security requirements. The DDI can encompass GDPR privacy requirements in a holistic assurance case, which contains the argumentative basis for the satisfaction of technical security requirements derived from GDPR. Assurance case patterns are utilized to abstract system details and these patterns can be re-applied to a different system's DDI. The concept involves conducting a privacy risk assessment to identify potential harms to a data subject, events that may lead to those harms, system privacy weaknesses, and individuals or organisations that could cause harm. To assess the risk, that is severity and likelihood of a harm occurring, Harm trees are utilized. The results of the risk assessment are then captured in the DDI through the TARA package. To assist this capture, a mapping of the PRIAM elements to the TARA elements has been provided.

Capturing this privacy risk assessment data in the DDI's TARA package should ensure CPS compliance with the GDPR technical security requirements at both design time and runtime. Design time DDI assurance involves performing and documenting adequate risk analysis and mitigation of privacy risks associated with personal data. At runtime, DDIs are used to ensure that the chain of systems implementing the data processing service adequately protects the data rights. Furthermore, traceability between the mitigating security requirements expressed in the DDI and the system design elements that realise them, greatly simplifies change impact analyses in addition to providing a transparent means to check GDPR compliance automatically.

Acknowledgement. This paper is supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No 732242. It is also supported in part by Science Foundation Ireland grant 13/RC/2094.

References

1. Marr, B.: How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read, Forbes, 21 May 2018. https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/#49ce131860ba
2. European Commission, 2016. Digitising European Industry Reaping the full benefits of a Digital Single Market, COM (180 final) (2016). http://ec.europa.eu/transparency/regdoc/rep/1/2016/EN/1-2016-180-EN-F1-1.PDF
3. European Commission, 2015. EU initiatives on cybersecurity and cybercrime. In its COM 1922 (2015). http://ec.europa.eu/information_society/newsroom/image/document/2017-3/factsheet_cybersecurity_update_january_2017_41543.pdf
4. European Commission, 2018. Artificial Intelligence for Europe, COM 237 (2018). https://ec.europa.eu/transparency/regdoc/rep/1/2018/EN/COM-2018-237-F1-EN-MAIN-PART-1.PDF
5. McKinsey: Artificial Intelligence - Next Digital Frontier, June 2017. http://www.odbms.org/2017/08/artificial-intelligence-the-next-digital-frontier-mckinsey-global-institute-study/
6. European Union. REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)
7. DEIS. http://deis-project.eu/. Accessed 19 Feb 2020
8. Avižienis, A., Laprie, J.-C., Randell, B.: Dependability and its threats: a taxonomy. In: Jacquart, R. (ed.) Building the Information Society. IIFIP, vol. 156, pp. 91–120. Springer, Boston, MA (2004). https://doi.org/10.1007/978-1-4020-8157-6_13
9. European Union: Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data. https://eur-lex.europa.eu/eli/reg/2016/679/oj
10. Gholami, A., Laure, E.: Advanced Cloud Privacy Threat Modelling (2016). https://arxiv.org/abs/1601.01500 Accessed 06 Apr 2020
11. Wuyts, K., Scandariato, R., Joosen, W.: Empirical evaluation of a privacy focused threat modelling methodology. J. Softw. Syst. **96**, 122–138 (2014). https://doi.org/10.1016/j.jss.2014.05.075
12. De, S.J., Le Métayer, D.: PRIAM: a privacy risk analysis methodology. In: Livraga, G., Torra, V., Aldini, A., Martinelli, F., Suri, N. (eds.) DPM/QASA 2016. LNCS, vol. 9963, pp. 221–229. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-47072-6_15
13. Schneider, D., Trapp, M., Papadopoulos, Y., Armengaud, E., Zeller, M., Hoefig, K.: Wap: digital dependability identities. In: Software Reliability Engineering (ISSRE), pp. 324–329. IEEE (2015). https://doi.org/10.1109/issre.2015.7381825
14. Object Management Group (OMG). Structured Assurance Case Meta-model 2.0 (SACM) (2018). http://www.omg.org/spec/SACM/2.0/PDF

15. Reich, J., Zeller, M., Schneider, D.: Automated evidence analysis of safety arguments using digital dependability identities. In: Romanovsky, A., Troubitsyna, E., Bitsch, F. (eds.) SAFECOMP 2019. LNCS, vol. 11698, pp. 254–268. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-26601-1_18
16. DEIS Consortium. DEIS Project (2019). https://github.com/DEIS-Project-EU
17. Whiting, D., Sorokos, I., Papadopoulos, Y., Regan, G., O'Carroll, E.: Automated model-based attack tree analysis using HiP-HOPS. In: Papadopoulos, Y., Aslansefat, K., Katsaros, P., Bozzano, M. (eds.) IMBSA 2019. LNCS, vol. 11842, pp. 255–269. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32872-6_17

Supporting Process Design in the Autonomous Era with New Standards and Guidelines

Masao Ito(✉)

NIL Inc., 2-17-7, Kinuta, Setagaya, Tokyo, Japan
nil@nil.co.jp

Abstract. It isn't easy to define a general and clear process in a new system such as an autonomous vehicle. The new technology is complex and lacks experience. In this paper, we reconsider the process used for system development. We assume that the process changes depending on the characteristics of a project [1]. Of course, this is a general agreement. In autonomous vehicles, new technological elements such as AI and ensuring safety are further required. So, it is meaningful to consider how to think about the development process. New standards, guidelines and documents such as UL4600 are emerging for autonomous vehicles. In addition to many standards, dealing with these presents difficulties. We believe that better process design is possible by using the Toulmin model.

Keywords: Process design · Toulmin model · Autonomous vehicle · AI · UL4600 · SCSC-153A

1 Introduction

Currently, many types of research and developments on autonomous vehicles (SEI Level 4 or 5) [2] are underway. Autonomous vehicles do not assume human operations. On the other hand, the standards and guidelines up to now are, of course, premised on human operations. Because of this difference, it is necessary to consider safety differently than before.

In an autonomous vehicle, the machine recognizes the environment and controls it, instead of humans. The main issues regarding safety are as follows.

(A) Is it possible to ensure safety when an autonomous vehicle recognizes the environment and steers itself?
(B) How should we evaluate new technologies such as AI that performs control on behalf of humans?

Several new guidelines and standards have been published or planned to address these challenges. Although not necessarily exhaustive, some are shown in Table 1.

UL4600 [3] is for autonomous vehicles. In UL4600, the regulations mainly use the safety case. It also includes guidelines for AI systems. BSI 1880 [4] is a standard for autonomous vehicle control systems and covers a wide variety of vehicles. SCSC-153A [5] is also targeted at autonomous vehicles. This document has cross-references to UL4600.

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 525–535, 2020.
https://doi.org/10.1007/978-3-030-56441-4_39

Table 1. Recent standard, gridlines, and documents for AS

ID	No.	Title	Type	Issue date
(1)	UL4600	Standard for Safety for the Evaluation of Autonomous Products	Standard	April/2020
(2)	BSI I880	Guidelines for developing and assessing control systems for automated vehicles	Standard (PAS)	April/2020
(3)	SCSC-I53A	Safety Assurance Objectives for Autonomous Systems	Document	Jan/2020
(4)	MISRA	Guidelines for Automotive Safety Arguments	Guidelines	September/2019
(5)	DIN SPEC 9200I-I	Artificial Intelligence – Life Cycle Processes and Quality Requirements – Part I: Quality Meta Model	Standard (PAS)	April/2019
(6)	ISO/IEC 29II9-II	Testing AI-Based Systems	TR	–

As a document related to the safety case, MISRA gives guidelines on the safety and logical structure of automobiles [6]. It mainly relates to Part 3 of ISO 26262 and gives detailed guidelines for applying the Safety Case. The structure of the safety case is shown here as a template.

The following are the standards/guidelines specialized for AI. The DIN SPEC 92001 series are life cycle processes and quality requirements related to AI [7]. At present, only quality metamodels are published. As for handling AI testing, ISO IEC 29119-11 (Software and systems engineering—Software testing—Part 11: Testing of AI-based systems) is being established, but it is under development [8].

Here, we would like to consider what kind of efforts are made concerning the claims of these documents from the viewpoint of the process, rather than comparing individual standards and guidelines. In Table 1, we use UL4600 in (1) and SCSC-153A in (3), which have descriptions directly related to autonomous vehicles and AI, as examples in this paper.

In the next chapter, we briefly examine UL4600 and SCSC-153A. In Sect. 3, we reconsider the process of ensuring safety through a survey of these standards. In Sect. 4, we propose a method to support process design, using various standards and guidelines. We use the Toulmin model.

2 UL4600 and SCSC-153A

2.1 Characteristics

UL4600 aims to "*support(s) state-of-the-art safety case approaches, which permit standardizing an approach to safety while at the same time enabling the use of rapidly evolving technology, tools, and methods. It is both technology neutral and development process agnostic)*" [9]. That is, the objective is to react quickly as the autonomous vehicle evolves. Also, this standard has many prompt lists, predetermined lists of risk categories. These lists assist us in creating a safety case.

The UL 4600 chaptering is very interesting (Fig. 1). The first half of the standard (up to Chapter 8) is probably the one that the standard setters consider important, namely the safety case (Chapter 5), Risk Assessment (Chapter 6), and autonomy (Chapter 8). It has become. The second half is a relatively traditional title, but it's still a bit different from the general one; Software and system engineering process (Chapter 9), dependability (Chapter 10), data and network (Chapter 11) and V & V and testing (Chapter 12).

UL4600	SCSC-153A
1 Preface (Informative)	1 Introduction
2 Scope	2 Computation-Level Framework: Description
3 Referenced Publications	3 Computation-Level Framework: Objectives
4 Terms, Definitions, and Document Usage	4 Autonomy Architecture-Level Framework: Description
5 Safety Case and Arguments	5 Autonomy Architecture-Level Framework: Objectives
6 Risk Assessment	6 Platform-Level Framework: Description
7 Interaction with Humans and Road Users	7 Platform-Level Framework: Objectives
8 Autonomy Functions and Support	8 Summary
9 Software and System Engineering Processes	Appendix A Computation-Level Framework: Justification
10 Dependability	Appendix B Computation-Level Objectives: Justification
11 Data and Networking	Appendix C Platform-Level Framework: Justification
12 Verification, Validation, and Test	Appendix D Comparison with AAIP Body of Knowledge
13 Tool Qualification, COTS, and Legacy Components	Appendix E Comparison with UL4600
14 Lifecycle Concerns	Appendix F Comparison with OECD Principles on AI
15 Maintenance	Appendix G Known Issues
16 Metrics and Safety Performance Indicators (SPIs)	Appendix H Abbreviations
17 Assessment	Appendix I References
Annex A (Informative) – Use with ISO 26262 and ISO/PAS 21448	Appendix J Contributors

Fig. 1. Structure of UL4600 and SCSC-153A chapters

The SCSC-153A aims to focus "*on aspects directly related to autonomy, and enabling technologies such as AI and ML, rather than more general safety engineering or system engineering, where it is assumed that relevant general standards, guidelines and best practice will be applied*".

The whole is divided into three levels: compute level, autonomous architecture level, and platform level. Each level has a projection. At the computational level, there are five projections: Experience, Task, Algorithm, Software, and Hardware. It has the objective for each projection.: "*Each objective is accompanied by a discussion that illustrates how the objective contributes to AS safety. This is followed by examples of approaches that could be taken to satisfy, or partially satisfy, the objective.*"

2.2 Processes

In the introduction to UL4600, there is the following statement: *"Traditional safety standards are prescriptive"*. And these traditional standards provide *"how to do safety (process, work products)"*. So, UL4600 says it provides a goal. By the way, the traditional standards are ISO 26262 [10], ISO/PAS 21448 [11], IEC 61508, MIL-STD

882 and so forth. If we use the term process in a sense that is more like a procedure, we think there is probably a misunderstanding here. Indeed, ISO 26262 defines the work products, but it doesn't have the procedure.

What is the process is a delicate matter, but it is described as follows in the standard ISO 12207 [12] for software life cycle processes. *"... this document does not prescribe any particular sequence of processes within the life cycle model. The sequence of the processes is determined by project objectives and by selection of the life cycle model"*. ISO 26262 conforms to ISO 12207 Just because you specify a process does not mean that you specify a chronological order. I will discuss this point in the next chapter.

SCSC-153A, like its title, is a description of objectives regarding safety assurance. Therefore, there is not much description of the process. From the few descriptions, there is the following description regarding security as an example.

We can't get the same assurance evidence as to the normal development process.: "*..., from a practical perspective, most ML pipelines make extensive use of open source frameworks and tools which, generally speaking, do not provide the same type of assurance evidence. as is delivered by development processes for critical software*".

In Table 10 of SCSC-153A, Mapping Projections to Typical Software Development, there is a correspondence between standard life cycle and Computation level projection. For example, Experience projection is related to design and implementation. Task projection is related to everything from Plan to Test.

3 Rethink the Process

UL4600 claims that existing safety standards are prescriptive (c.f. Sect. 2.2). To be sure, there is such an aspect, but I think that it may be said to be the influence of the times when the standard was established. At the initial examination stage of ISO 26262, I believe that passenger cars of Level 3 and above were not the centre of interest. Generally, in the area where the change is drastic, and the technical accumulation is small when we try to make some rule, it becomes the goal-based writing method. Also, the explanations are exemplary or an exhaustive list. Here, we will reconsider the process and consider a way that is effective in both formats.

3.1 Is the Process a Kind of Recipe?

Here, we will reconsider the process. First of all, think about whether a process is a time-based definition of activities like a recipe.

Refer to the definition of ISO 12207: the process is a *"set of interrelated or interacting activities that transforms inputs into outputs"*. Of course, I'm not saying here a sequence of activities. More aggressively:

This document does not prescribe any particular life cycle model. Instead it defines a set of processes, termed life cycle processes, which can be used in the definition of the system's life cycle. Also, this document does not prescribe any particular sequence of processes within the life cycle model.

However, the input and output may implicitly determine the order. Now, if activity A outputs a work product and activity B takes that the one as input, then activity A must precede activity B. it may be possible to say that the order of execution of activities is determined by specifying input and output without defining the order explicitly.

In ISO 26262-3, we need to define items and create work products called item definitions. To enact the Hazard Analysis and Risk Assessment (HARA) activity, we need the item definition as an input.

```
5.5 Work products
5.5.1 Item definition resulting from requirements in 5.4.
...
6 Hazard analysis and risk assessment
6.3 Inputs to this clause
6.3.1 Prerequisites
The following information shall be available: — item definition
in accordance with 5.5.1.
```

There is an order relation between the item definition and HARA via the work product called the item definition. But this doesn't seem to have any further meaning. It's just like you can't test that code without coding it. The actual working process is not that simple.

In Agile development, we write the code for testing as preparation for unit testing before implementation. Of course, we do test after coding, but we can design test cases and then write the test code. The test-first approach might have the potential to produce good quality code with validation in mind. You may also notice design mistakes in the process of creating test cases.

We will consider the above in detail. Let's say design is D and coding is C. The test is T. Abstractly, we can think the following order.

$$\mathrm{D} \prec \mathrm{C} \prec \mathrm{T} \tag{1}$$

The symbol '$\prec$ ' indicates the execution order here.

Now, decompose T1 as follows.

$$\mathrm{T} = \mathrm{T1} \prec \mathrm{T2} \prec \mathrm{T3} \tag{2}$$

Here, T1 is test case creation, T2 is test code creation, and T3 is test execution.

Simply, (1) becomes

$$\mathrm{D} \prec \mathrm{C} \prec (\mathrm{T1} \prec \mathrm{T2} \prec \mathrm{T3}) \tag{3}$$

However, if you think about test first, you can also:

$$\mathrm{D} \prec \mathrm{T1} \prec \mathrm{T2} \prec \mathrm{C} \prec \mathrm{T3} \tag{4}$$

Alternatively,

$$\mathrm{T1} \prec \mathrm{D} \prec \mathrm{T2} \prec \mathrm{C} \prec \mathrm{T3} \tag{5}$$

The input and output of each process element do not change in either case. Only the choice is different.

Next, consider the process instance. Now assume that {di1, di2, …, din} is an instance of process element D.

di1, …, d1k: DP1 is a critical element, and I want to confirm its feasibility early. On the other hand, d1l, …, d1n: DP2 is an easy element so that we can design it later.

At this time, if you take a strategy to tackle difficult issues first, we can get the sequence showing below:

$$\mathrm{DP1} \prec \mathrm{CP1} \prec \mathrm{TP1} \prec \mathrm{DP2} \prec \mathrm{CP2} \prec \mathrm{TP2} \tag{6}$$

This is a kind of incremental approach. It is easy to find for us by using parenthesis.

$$(\mathrm{DP1} \prec \mathrm{CP1} \prec \mathrm{TP1}) \prec (\mathrm{DP2} \prec \mathrm{CP2} \prec \mathrm{CP2}) \prec \ldots \tag{7}$$

The work products are the same in both cases. The only difference is the choice of the process designer.

That is, the work product definition of each activity does not uniquely determine the order in which the activities are enacted.

3.2 How We Define a Recipe for a Project?

As mentioned earlier, each project will determine the appropriate process based on the given conditions. At this time, how will the process be decided? Usually, you will make a trade-off of QCD based on your development experience. If you are trying a new process model for the first time, you are going to make a trial and use that experience. Also, we design a process. In this section, we consider ways to support process design.

The method we propose uses the Toulmin model [13]. The Toulmin model is the idea behind the GSN to express the safety case. Initially, in an attempt to clarify the structure of everyday discussions, the figures were written for the explanation, and there was no precise definition. Various people are expanding, but here we consider the original expression (Fig. 2).

- Data: The facts or evidence used to prove the argument.
- Claim: The statement being argued.
- Warrants: The general, hypothetical (and often implicit) logical statements that serve as bridges between the claim and the data.
- Qualifiers: Statements that limit the strength of the argument or statements that propose the conditions under which the argument is true.
- Rebuttals: Counter-arguments or statements indicating circumstances when the general argument does not hold true.
- Backing: Statements that serve to support the warrants.

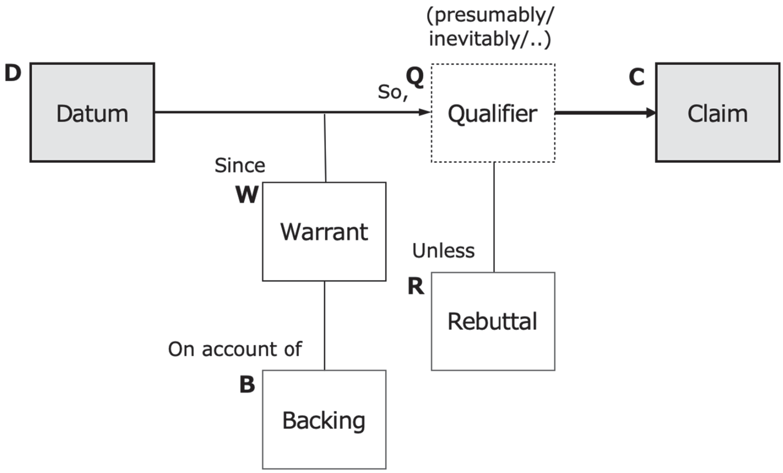


Fig. 2. Toulmin model

By reversing this structure in time, we can see the similarity with GSN. However, in general, there is no equivalent to Qualifier, which means the confidence factor of the claim. The standard GSN does not have an equivalent for this Qualifier.

Let me give you an example. Now, let's say that we have to carry out development using technology elements that we have never used (D). Since it is unclear whether the technical element can be used, it may be possible to choose to proceed with development (W) while checking. Generally, this method is known as incremental development (B). From now on, perhaps (Q), select the development process in incremental development (C). However, it may become difficult if there are restrictions on costs and delivery (R).

In this way, it is possible to evaluate if an argument is defined. Of course, since there is a modality there, it cannot be said to be entirely correct, and in some cases, Claim may not be established. It is possible to evaluate including that.

4 Toulmin Model and New Technology Elements

At present, it is not easy to cover all aspects of how to ensure safety in the new field of autonomous vehicles and the new technological field of AI. It is also possible that the goals for ensuring safety will change as the technology evolves. Therefore, developers need to follow the revisions of many existing standards/guidelines/documents. Furthermore, the number of target documents is expected to increase in the future. For example, there is an IEEE P7000 series on ethics that is not included this time.

In this paper, we propose the following method. We are organizing the claims for safety in each document using the Toulmin model. It gives users a centralized way to access what a document requires. Also, you can easily change or add.

On the other hand, there are difficulties. For example, terms are not always consistent across standards. For example, the SCSC-153A has 'Experience' projection at the computation level. This refers to the dataset used for training. However, similar things may not be called Experience in other documents. We allow replacement using terminology dictionaries, but not a complete solution. When you use Experience literally, it spoils its meaning by replacement. The final solution can only be obtained by referring to and understanding the original document. The main purpose of this scheme is to get to the relevant part of the required document without leakage.

Also, multiple W/B/R may be required for a particular Claim. For the graphical description, you can represent everything as nodes, but I don't think it is a proper method. This is because the structure that is too complicated impairs the intelligibility of the diagram. We express that there is another W/B/R in Qualifier as a link.

Methods of integrating knowledge are often difficult to use continuously in real problems (e.g. Unified Process). We would like to use Toulmin's model as a way to reach all the relevant parts of a proper document without leaking them. We are not trying to integrate everything.

For SCSC-153A, take an example from COM1-1.

Belonging to the computational experience are the following four Objectives.

COM1-1: Data is acquired and controlled appropriately.
COM1-2: Pre-processing methods do not introduce errors.
COM1-3: Data captures the required algorithm behaviour.
COM1-4: Adverse effects arising from distribution shift are protected against.

These are Objectives for the data used in ML. The following sentence can be found in the Example.

If a complete data set is acquired from an external party then care should be taken to ensure that it has not been subject to "Data Poisoning"; for example, the addition of a small number of maliciously crafted samples can create a backdoor" The same techniques used to confirm the authenticity of information downloaded from the Internet (e.g., checksums) may be helpful here.

An example of the representation using the Toulmin model is shown in Fig. 3.

Here, the data happens to be the same as the node name in the Toulmin model. The qualifier is "with other warrants" because there can be other Ws. For example, prevention of semantic errors in data definition is equivalent (whether the vehicle speed is expressed in MKH or MPH in vehicle speed data). This can also be added as W or B.

The left side at the top of the figure shows the target system, the center shows the technology category, and the right side shows the related phases in the life cycle; R: requirement (analysis), D: design, I: implementation, T: test. Figure 3 shows that ML for autonomous vehicles is a factor to be considered in design (D) and implementation (I).

You can find a similar example in UML4600. There are the following rules regarding data for ML.

8.5.3 Machine learning training and V & V shall use acceptable data.
c) Data provenance: historical record of data and its origins
NOTE: This can support better understanding of data, track back sources of errors, and provide auditability and quality trails

Similarly, Fig. 4 shows the thing written by the Toulmin model.

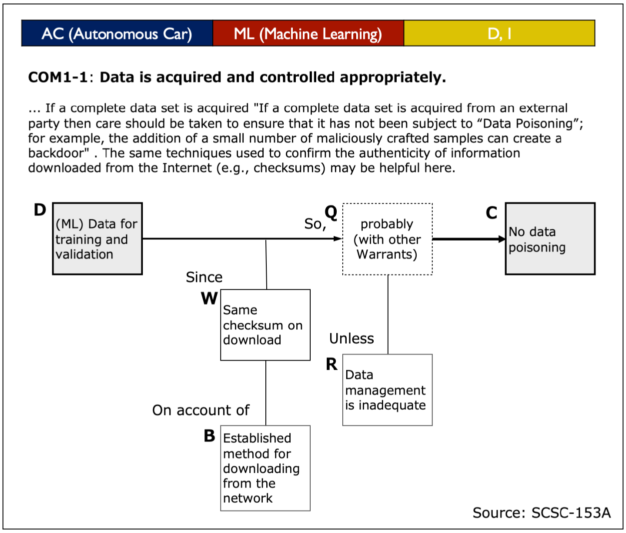


Fig. 3. A Toulmin model from SCSC-153A COM1-1

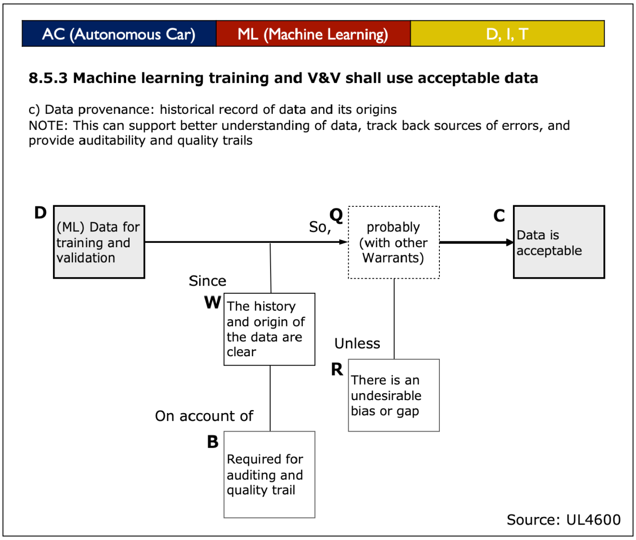


Fig. 4. A Toulmin model from UL4600 8.5.3

5 Summary

As we see, several standards and guidelines have been established or are being developed to ensure safety for the practical application of autonomous vehicles. These documents also reflect the fact that many technologies in the new AI category are used in the new platform of autonomous vehicles. As technology evolves rapidly, so does the way it ensures safety. This document provides an overview of these new documents and highlights their characteristics.

Next, we consider the relationship with the process. The discussion here is traditional. That is, a process does not necessarily describe a temporal sequence. The process designer understands the process elements, considers the process instance, and defines the process for each project based on the trade-off.

Also, these standards/guidelines are less likely to be related to their life cycle than existing standards (e.g. ISO 26262). Both UL4600 and SCSC-153A, which were taken up this time, are so-called goal-based descriptions. Therefore, process designers encounter difficulties when considering what to achieve at what timing when designing a process that considers safety.

In this paper, we have proposed a method for organizing these new documents using Toulmin's model. I don't think of using a graph with a large argument structure. As shown in Sect. 4, for one goal (requirement), one argument structure is made into one card using the Toulmin model. Depending on the stage of development, we will collect the necessary cards and judge whether they are sufficient. Proper maintenance is easy. You can also add new cards as needed.

We believe this method will be useful as we adapt to new standards and guidelines that will continue for the next few years.

References

1. Korsaa, M., et al.: The SPI manifesto and the ECQA SPI manager certification scheme. J. Softw. Evol. Process **24**(5), 525–540 (2012)
2. SAE J3016. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles, J3016_201806: SAE International (2018)
3. UL4600. Standard for Safety for the Evaluation of Autonomous Products, UL (2020)
4. BSI PAS 1880:2020. Guidelines for developing and accessing control systems for automated vehicles (2020)
5. SCSC SASWG, SASC-153A. Safety Assurance Objectives for Autonomous Systems (2020)
6. MISRA. Guidelines for Automotive Safety Arguments (2019)
7. DIN SPEC 92001. Artificial Intelligence – Life Cycle Processes and Quality Requirements – Part 1: Quality Meta Model (2019)
8. ISO/IEC 29119-11. Software and systems engineering—Software testing—Part 11: Testing of AI-based systems (under development)

9. https://medium.com/@pr_97195/an-overview-of-draft-ul-4600-standard-for-safety-for-the-evaluation-of-autonomous-products-a50083762591. Accessed 4 2020
10. ISO 26262:2018. Road vehicles – Functional safety (2018)
11. ISO/PAS 21448:2019. Road vehicles – Safety of the intended functionality (2019)
12. ISO/IEC/IEEE 12207:2017. Systems and software engineering—Software life cycle processes (2017)
13. Toulmin, S.E.: The uses of argument updated ed. Cambridge University Press, p. 97 (2003)

Requirement Definition Taking a Flyer on DO-254

Alastair Walker(✉)

Lorit Consultancy GmbH, 5020 Salzburg, Austria
alastair.walker@lorit-consultancy.com

Abstract. Defining requirements that both capture the essential properties at a given hierarchical level and meet the appropriate needs of that hierarchical level is never any easy task. ISO 26262 has a fairly minimal coverage on how to achieve the hierarchical relations of hardware requirements and also the abstract way the standard has evolved over two versions adds some confusion to this subject. The DO-254 Design Assurance Guidance for Airborne Electronic Hardware conversely gives very good guidance on this topic, albeit that DO-254 is now 20 years old, does that really matter? There are techniques from more modern guidance that can augment the definition in DO-254 and support ISO 26262 hardware definition and implementation. In this paper we look at the strategy defined in DO-254 and combining this with more recent thinking on the subject to propose a method for assisting the ISO 26262 requirement definition activities. We also consider at what level ISO 26262 hardware parts should or should not be traceable to hardware safety requirements.

Keywords: Requirements · Hardware component · Hierarchy

1 Introduction

ISO 26262 [1] encourages the concept of tailoring and this in short should encourage engineers to tailor the numbers and levels of requirements. Unfortunately, this is often not the case in ISO 26262 projects and teams often prefer to stick to defined text in a standard irrespective whether normative or informative. For this reason, in this paper we look at the different approaches in different industry standards on this topic and specifically to the levels of abstraction in hardware requirements. For doing this, we take the airborne DO-254 [2] standard as a reference.

2 DO-254 Approach to Hardware Requirements and Design

The DO-254 is arguably the best-known hardware standard used in safety related industries, the most recent version dating back to April 2000. Figure 1 illustrates the design lifecycle and the two levels of hardware design i.e. Conceptual and Detailed.

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 536–542, 2020.
https://doi.org/10.1007/978-3-030-56441-4_40

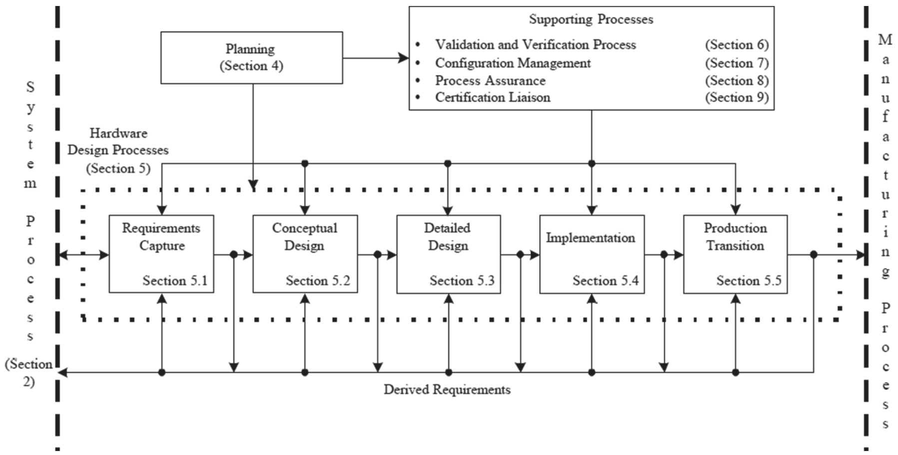


Fig. 1. DO-254 hardware design lifecycle [2]

2.1 DO-254 Conceptual Design Activities

At the conceptual design level, a high-level description is generated including architectural constraints and including any on software. All major components are identified, derived requirements particularly for interfaces are defined.

2.2 DO-254 Detailed Design Process

At the detailed design level, the design data for the hardware is generated based on the requirements and conceptual design data. This may include component data, VHDL, test methods and hardware-software interface data. Also, at this stage architectural design techniques are implemented e.g. safety monitors and diversity.

2.3 DO-254 Derived Requirements

Derived requirements in DO-254 may address for example, the interfaces between different design assurance levels, the range of data inputs, reset states or performance of timing related functions. Derived requirements at all stages feedback into the requirements processes.

Figure 2 illustrates one key difference between DO-254 and ISO 26262, in that DO-254 considers all hardware and not just hardware covered by functional safety requirements. Albeit the topic is mentioned briefly in ISO 26262 part 5. Note software aspects are not covered in this paper.

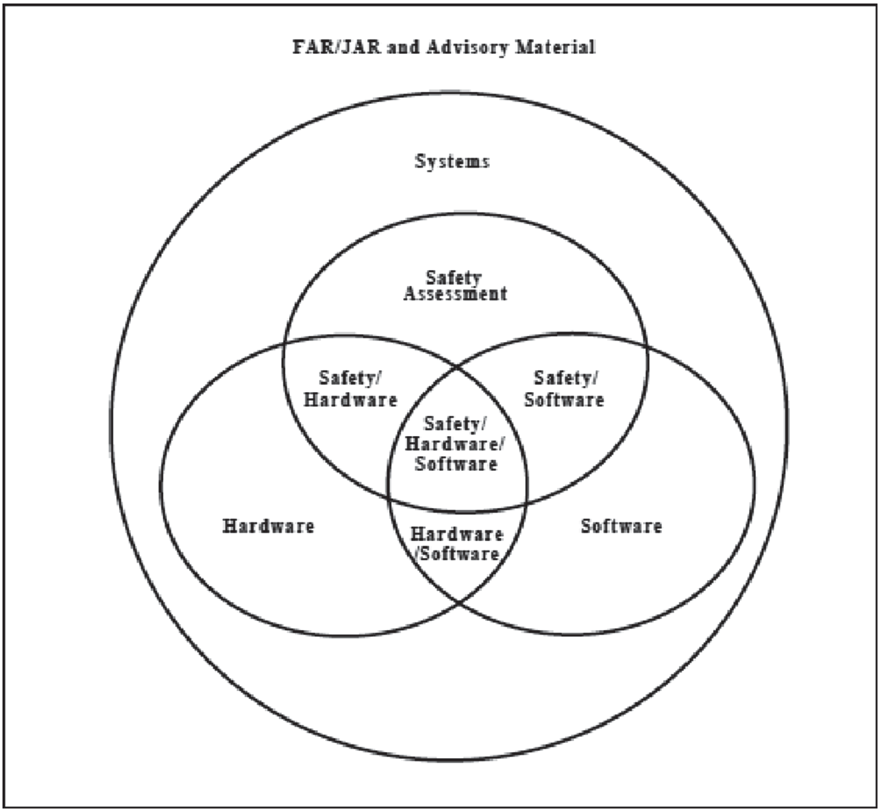


Fig. 2. DO-254 relationships among airborne systems safety assessments hardware and software processes [2]

3 ISO 26262 Approach to HW Hierarchy

ISO 26262 has no real guidance on whether one hierarchical level or two are necessary. Section 7 of part 5 does require hardware to be split into architectural and detailed hardware. As the focus of ISO 26262 is functional safety orientated, there is no definition of the nature defined in Fig. 2.

Hardware architectural design - represents all hardware components and their interactions with one another.

Hardware detailed design - is at the level of electrical/electronic schematics representing the interconnections between hardware parts composing the hardware components.

Where ISO 26262 shows a few weaknesses is in the way the hardware aspects are split between part 5, 10 and 11. Ultimately, as part 11 was written for the second addition of the standard in 2018, then it is understandable that the information has been separated, however this does not help with consistency in the standard. Figure 3 illustrates the relationship between requirements and design phases for hardware as defined in part 10 of ISO 26262. This does not show the detail of hardware architectural against hardware detailed design.

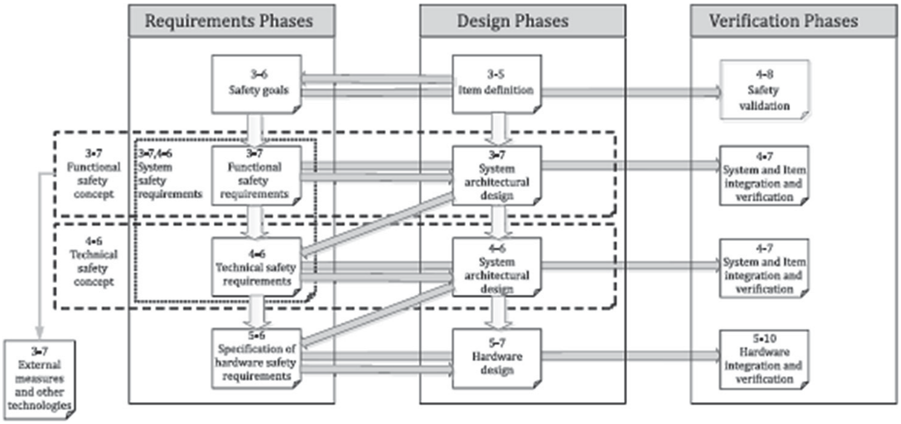


Fig. 3. Safety requirements, design and test flow from concept to hardware [1]

4 ISO 26262 What Is Architectural Design in Comparison to Detailed Design?

ISO 26262 has an interesting take on this point in part 5, that contradicts the concepts of part 11 and the considerations for semiconductors.

Hardware architectural design represents all hardware components and their interactions with one another. Hardware detailed design is at the level of electrical/electronic schematics representing the interconnections between hardware parts composing the hardware components.

The above statement from part 5 does not represent the reality of the semiconductor industry, as modern devices such as safety compliant microcontrollers may have many abstract levels of architecture before the level of hardware part is reached. Also, the relationship of hardware architectural level to software is not well represented in ISO 26262. In part 11 section 4.7.6 has the following statement, that does link the two areas [1]:

NOTE 1: Firmware and any micro-code running on programmable HW elements, irrespective of whether they are classified as CPUs or not, can be considered to be SW elements.

This statement is then defining a software architectural design level and a software detailed design level below the hardware architectural level and hence detailed design is traceable to software safety requirements.

In ISO 26262, a component is a non-system level element which is comprised of one or more hardware parts (or software units). Hardware parts have three levels of abstraction, a hardware part, a hardware elementary subpart and a hardware subpart. A hardware part can comprise of a diverse range of devices a CPU, a resistor or a flash array. Elementary subparts being a flip-flop of an ALU. Finally, a subpart can be represented by an ALU of a CPU or a register bank of a CPU.

5 Process Considerations

If considering definition of processes for hardware development, the split between the abstract levels need clear definition, as does the relationship between requirements and design implementation. In part 5 of ISO 26262, section 7.4.1.5, the following note is both confusing and misleading. This note is often held up as a reason for not defining requirements down to hardware part level, despite it being informative [1]:

NOTE: The traceability of hardware safety requirements is not required down to the hardware detailed design. No hardware safety requirements are allocated to hardware parts that cannot be divided into subparts. For example, it is neither meaningful nor beneficial to try to establish hardware traceability down to each capacitor and resistor, etc.

Safety requirements can absolutely be assigned to both hardware parts and hardware subparts. When considering for example a capacitor used across an isolation barrier there may be very specific requirements for the capacitor to meet Y Class ratings (Y1 or Y2) complying with IEC 60384-14 [3] hence there is a clear requirement for that capacitor to meet dielectric strength requirements. This is a major topic in IEC 60601-1 [4] where Y class capacitors are deemed to provide a means of protection against electrocution. The levels of abstraction for hardware requirements and hardware design can then be represented as shown in Fig. 4, which expands on Fig. 3 and takes into consideration the points discussed above.

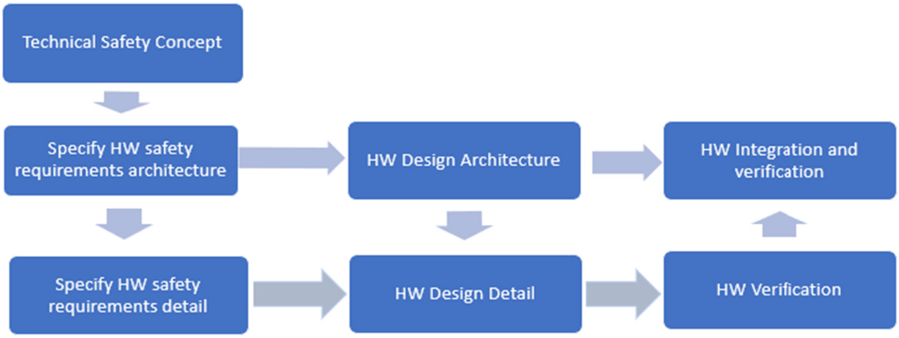


Fig. 4. HW safety requirements, design and verification

Figure 4 can then expand the processes to the software aspects, if programmable devices are to be incorporated into the item.

Returning to the topic of one or two levels of hardware requirements i.e. architectural and detail, it can be argued that only one level would be necessary for a relatively simple circuit, the example of an operational amplifier is often used. Here, gain, bandwidth, common-mode rejection and phase reversal, can all be verified at ISO 26262 hardware component level, without defining specific requirements for each hardware part. However, this needs to be assessed on a case by case basis in each project. Following the lead given in DO-254, and particularly for more complex items, semiconductors or programmable devices two levels of abstraction are really required, particularly for safety relevant hardware.

6 Hardware Engineering SPICE PAM

Also, noteworthy to mention in this paper is the hierarchal structure defined in the PAM for Hardware Engineering Processes [5]. The first version of the PAM released in November 2019 refers to one level of hardware requirement abstraction, while the software and mechanical plug-in modules for PAM Automotive SPICE® [6] refer to two levels of abstraction. Also, to note the PAM for Automotive SPICE® still defines two levels of abstraction.

7 Conclusion

Guidance and definition on hardware requirements abstraction is not as clear in ISO 26262 as parts it could be, much of which is down to the iterative way the standard has evolved. The considerations defined in DO-254 are more appropriate and kept generally at a simpler level which could certainly leave more room for interpretation; however, these can be applied to products from large items down to small circuits in a consistent fashion.

ISO 26262 at present (edition 2) leaves topics such as semiconductors, the relationship to software in programmable devices in a slightly confusing position. The note suggesting the requirements should not be traceable to hardware parts is also not helpful as in reality certain hardware parts may very well be defined by safety requirements.

One level of hardware requirement is difficult to argue as an acceptable solution unless the item being developed is of a fairly simplistic nature.

DO-254 taking the approach in a normative manner that requirements need to be defined at conceptual, detailed design and derived levels, provides a more consistent definition.

References

1. ISO 26262:2018. Road vehicles – Functional safety
2. DO-254. Design Assurance Guidance for Airborne Electronic Hardware
3. IEC 60384:2010. Fixed capacitors for use in electronic equipment - Part 14: Sectional specification - Fixed capacitors for electromagnetic interference suppression and connection to the supply mains

4. IEC 60601-1 Medical electrical equipment – Part 1: General requirements for basic safety and essential performance
5. PRM/PAM Hardware Engineering (HWE) 1.0 – intacs™ Working Group 'HW Engineering Processes'
6. Automotive SPICE® Process Assessment/Reference Model Version 3.1 VDA QMC Working Group 13/Automotive SIG

An ICS Based Scenario Generator for Cyber Ranges

Xabier Larrucea[1,2](✉) and Alberto Molinuevo[1]

[1] TECNALIA, Basque Research and Technology Alliance (BRTA), Derio, Bizkaia, Spain
{xabier.larrucea,alberto.molinuevo}@tecnalia.com
[2] University of the Basque Country, Vitoria-Gasteiz, Spain

Abstract. Cybersecurity is playing a key role in industrial control systems' scenario because they are one of the targets of cyber-attacks. In order to mitigate the impact in industrial scenarios, it is necessary to train people not only in IT breaches, but also in OT breaches in industrial settings. Training is one of the weaknesses identified within the industry especially by practitioners, and the use of cyber ranges is motivated. In this sense, this paper deals with the deployment of Industrial Control Systems scenarios based on honeypots for training purposes. An example illustrates the deployment of a scenario within a cyber range.

Keywords: Cyber range · Cyber scenario deployment · Honeypot

1 Introduction

Cybersecurity is a hot topic in industry [1], and its relevance is increasing [2] because Industrial Control Systems (ICS) are being considered one of the targets of cyber attacks [3]. In this context, cyberthreats are becoming one of the most relevant issues for the industry including manufacturers [4]. In fact, this study [4] reveals that 60% of manufacturing companies have been subject to a cybersecurity incident, and "almost a third of whom suffered some financial loss or disruption to business as a result" [4]. Based on the same report, 33% of companies do not provide awareness briefs or formal training to their employees. So the industry is facing training as a challenge [5, 41].

Training is one of the weaknesses identified for the industry [6]. In fact, the industry, academia and governments are required to provide means for enhancing cyber skills to key personnel [7]. It is being evidenced that "college graduates are forced to enroll in numerous certification programs in order to develop critical industry useful hands-on skills that are not taught in college" [8]. So traditional training such as attending courses at universities [9], is not enough for training personnel in industry, and they are required to develop their skills in real scenarios [10]. Industrial scenarios such as Industrial Control Systems (ICS) [11] are the most common industrial settings [12]. In addition, "Awareness and Training" is a category of the NIST framework for improving cybersecurity [13]. Therefore, training must be considered from different angles.

There are several technologies for teaching cybersecurity such as the cloud [14]. But industrial scenarios basically require hands on experience, and they propose the use

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 543–554, 2020.
https://doi.org/10.1007/978-3-030-56441-4_41

of cyber ranges such as Emulab [15] or DETERlab [16]. These environments emulate not only Information Technology (IT) networks but also Operational Technology (OT) infrastructures such as supervisory control and data acquisition tools (SCADA). Hands-on labs are being used as testbeds [16], and some research works are focused on IT/OT infrastructure for dealing with the training [17]. There is not a silver bullet for training, but all practitioners must evaluate the needs [18]. Once the needs are clear enough, practitioners evaluate the cost of the training [19] which is not always evident especially in hands on scenarios. The effectiveness of trainings does not produce real tangible Return On Investments (ROI) [20], and this includes the definition of cyber range scenarios.

Nowadays, cybersecurity is playing an important role in industrial control systems [21]. These scenarios are complex, and they require specific settings for testing their specific infrastructure. Honeypots have been used for intrusion detection in industrial scenarios [22], and they have been used for defining and simulating ICS [23]. There is a quite number of honeypots that can be used. Conpot [24] is one of the honeypots used for simulating ICS, and it is the honeypot used in our paper for simulating ICSs.

Therefore, based on the aforementioned situation, the definition and the implementation of ICS scenarios based on honeypots for training purposes are envisaged as challenges for the industry. In fact, this paper deals with this situation, and it presents a novel approach for generating ICS based scenarios within a cyber range for training purposes.

This paper is structured as follows. Firstly, it provides a background on cyber ranges and honeypots. Secondly, the architecture and tool are described in detail. Thirdly, an example illustrates the use of the tool. Finally, the main conclusions are provided at the end of this paper.

2 Background

2.1 Cyber Ranges

A cyber range is a virtual environment used mainly for training purposes. It is also used for research and development activities in order to simulate complex situations. In fact, there are several cyber ranges uses such as attack & defence, evaluate a novel technology, or to evaluate risks. In this sense, there are testbeds [25] such as SECUSIM, RINSE, ARENA, and LARIAT for simulating and testing networks scenarios [26]. In this sense, Emulab [15] o DETERlab [16] have been proposed as open source solutions. Emulab is a Network Lab Bed developed by the University of Utah Computer School. This network experiment bed consists of a series of computers, servers and routers, and a dedicated software system to manage operations. Some recent research in this field suggests platform independent mechanisms for partitioning experimental networks and breaking it down into smaller experiments that can be run sequentially [27].

Developing cybersecurity competencies through the use of Cyber Ranges, and their use for research topics is of increasing interest. In addition, Cyber Ranges include tools that help improve the stability, security and performance strength of IT/OT systems.

From a physical point of view, these Cyber Ranges are composed of a host room (Security Operations Center (SOC)) which controls the Cyber Range and includes monitoring tools (IPS, IDS, SIEMS). Then there is a second room which is where the training and research exercises are performed. The scenarios must be defined and modeled in the Cyber Range. During training, the Cyber Range allows Capture The Flag (CTF) or Red-VS-Blue exercises. CTF is a challenge in which different people compete to attack a series of servers and if they succeed, they earn a series of points. In the case of Red-VS-Blue 2 teams are created, one to attack and the other to defend.

Cyber Ranges are having a major impact at the country or region level for national security reasons, and states are positioning themselves with the aim of improving their capabilities [28]. The industry is a key element in any country and begins to be threatened by the effects of attacks from cybercriminals. Investment in cybersecurity is increasing due to the growing real threat of cybercrime, and it is a must for every company and Country. In order to mitigate the risks from cybersecurity, the need to investigate possible attack and defence scenarios has been identified, and for this, Cyber Ranges play a fundamental role in training and investigating these possible complex scenarios. In order to implement and adapt a Cyber Range environment, it is necessary to use appropriate software for the identified needs and to develop the appropriate scenarios to be evaluated. Therefore, this paper deals with the scenario's generation. Finally, it should be noted that computer security education and training activities are critical [29]. In fact, recent studies such as [29] suggest the use of Moodle e-learning systems for interactive training.

According to the study carried out by [25] an updated taxonomy of a cyber range provides a meaningful description of each element involved in a cyber range. Based on this definition A scenario defines the execution environment as well as the storyline that indicates the execution steps of a test or a training exercise. It accurately represents the operational environment and training requirements, and drives training execution to ensure the achievement of training objectives [25].

2.2 Honeypot/Honeynet

As stated before, Honeypots have been used for intrusion detection in industrial scenarios [22]. Therefore, the generation of scenarios based on honeypots is a need and a challenge for the industry. This is due to the fact industrial companies are not willing to devote resources on replicating their networks by using honeypots, and there is a scarce of skilled staff. Traditionally, as defined in Fig. 1, honeypots can be used as mechanisms to disorient potential attackers. Industrial settings such as ICS [23] are quite different from a common used of honeypots (Fig. 1). Conpot [24] is one of the honeypots used for simulating ICS, and it is the honeypot used in our paper for simulating ICSs.

There are two types of honeypots [30]: honeynet and honeytokens. But basically, they rely on honeypots. Honeynets are complex entire networks of honeypots, and honeytokens *are not a computer; instead they are a digital entity, such as an Excel spreadsheet. Even though they are not a computer, they share the same definition and concept of honeypots, no one should be interacting with them. Any interaction with a honeytoken implies unauthorized or malicious activity* [30].

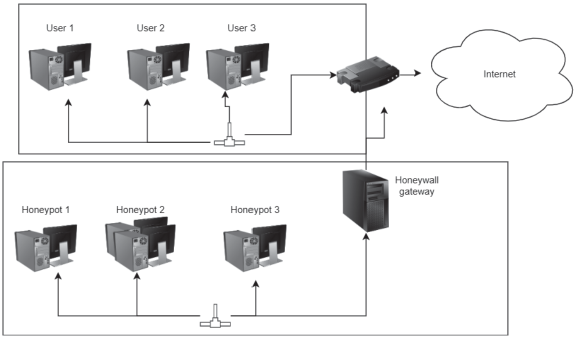


Fig. 1. Traditional network topology using honeypots

There are several projects developing honeypots such as Dionaea [31] mainly developed in python, and SNARE [32] which is a web application honeypot. Other projects such as HonSSH [33] are used for creating two separate SSH connections and it is jointly used with other honeypots. Another honeypot simulating systems operating functions is Sebek [34].

From a broad perspective, there are two kinds of honeypots: those with high interaction and those with low interaction. In the context of SCADAs, we use low interaction honeypots because they are aligned with the behaviour of this kind of environments. Conpot [24] is one of the honeypots used in this context. We use Conpot because we can simulate several parameters, and the behaviour of a PLC such as PLC Siemens S7-200. In the context of SCADAs, literature reflects some related works such as [35] where a framework called SCyPH is proposed in the context of SCADAs and the electric sector.

In the honeynet context, there is an open project called Honeynet and it is hosted in https://www.honeynet.org/ where there is a community researching on this kind of networks [36]. Honeynets are composed of four core elements [37]:

- Data control system which monitors all the activities of an attacker within the honeynet
- Data capture system which contains the activity carried out by an attacker
- Data collection system stores all captured data in one central location.
- Data analysis system provides the functionality to analyse the data being gathered.

In the context of this paper, our scope is to be limited to honeypots context, but allowing its extension to honeynets in a near future.

3 Proposed Architecture

3.1 Deployer Architecture

Our proposed architecture is designed to generate scenarios based on honeypots to be deployed and executed in a cyber range. The architecture (Fig. 2) consists of 4 different layers:

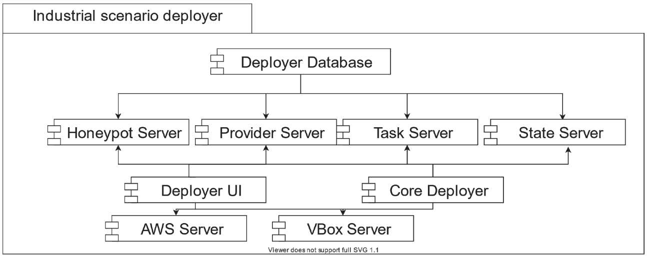


Fig. 2. Proposed architecture

- Data Layer contains a MySQL database subcomponent, which stores the catalogues of providers and honeypots, the different tasks to be processed and the states that these tasks can have.
- Operations Layer manages all the allowed operations on the database through 4 subcomponents.
- Deployer Layer contains all the business logic of Honeypot Manager and the User Interface.
- Instances Layer contains a subcomponent to manage all the operations allowed on each of the catalog providers (create, destroy, start and stop VM instances).
  - Scheduling: this feature allows to decide when a VM instance has to be created and/or destroyed, and how much time the VM instance has to be running or stopped, waiting to be restarted.
  - Log capture: Honeypot Manager collects the logs of the honeypots, and stores them in its file system.
  - User Interface: Honeypot Manager offers a web Swagger User Interface to facilitate the performance of the operations.

This architecture is composed by the following elements:

- Deployer Database: it contains the main MySQL database of the component.
- Provider Server: it manages all allowed operations (CRUD) on the information of virtual machines providers (AWS and Virtualbox).
- Honeypot Server: it manages all allowed operations (CRUD) on the available honeypots catalog.
- Task Server: any request of deployment, undeployment, starting or stopping VMs and honeypots is represented as a Task. This subcomponent manages all allowed operations on the existing tasks.

- State Server: it manages all allowed operations (CRUD) on the states (pending, running, stopped and expired) that the tasks can have.
- Core Deployer: it implements the business logic of Honeypot Manager.
- Deployer UI: the user interface (SwaggerUI).
- AWS Server: it allows to interact with EC2 instances in Amazon Web Services.
- VBox Server: allows to interact with VM instances in VirtualBox.

There is a tool developed following this architecture. This tool is in charge of deploying the configuration of different honeypots in a cyber range.

3.2 Deployer Entrypoint

It is worth to mention, and to describe the deployer entrypoint description. Each parameter is defined as follows:

- name: the name assigned to the honeypot in the catalogue.
- provider: the provider where the honeypot must be deployed ("aws" or "vbox").
- is_master: indicates if the VM instance will be master or slave.
- target_ip: the target IP of the machine where the honeypot will be deployed.
- schedule: indicates when to start and stop the instance, and how much time the instance remains running and stopped.
- start: timestamp when the instance must be launched ('0' to launch it immediately).
- stop: the timestamp when the instance must be stopped ('0' to never stop).
- time_unit: Minutes, Days, Hours or Minutes.
- time_up/time_down: how much time (in time units) the instance must be running/stopped.
- rules: indicates what rules must be applied to the incoming and outcoming traffic.
- protocol: "tcp" or "udp".
- inbound_ports/outbound_ports: the ports to be opened.
- logs: the different logs to be retrieved.

3.3 Main Process

As an entry point, a REST function is provided, which receives as an input parameter a configuration of honeypots to be deployed. For each honeypot, a task is added to the database, with the status Pending (Fig. 3).

The Core Deployer subcomponent is continuously checking if there are tasks to process (pending, running or stopped tasks). Depending on the state of each task, the Core Deployer takes the decision on how to proceed:

- Pending: If the task is expired in time, it is set to EXPIRED. If it is in time and it is not a future task, a new VM instance is created, the honeypot is deployed and the task is set as RUNNING.
- Running: If the task is expired in time, it is set as EXPIRED and the VM is stopped. If not, the task remains as RUNNING, unless the task is scheduled, and it has to be stopped. If so, the task is set as STOPPED and the VM instance becomes stopped too.

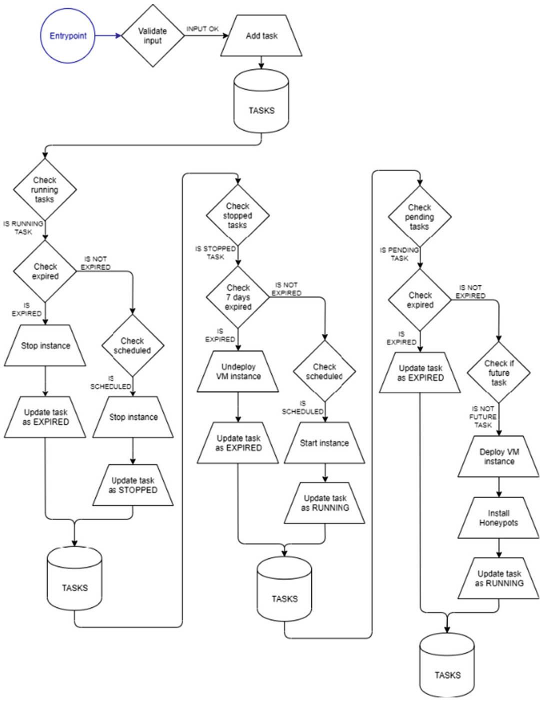


Fig. 3. Deployer main process

- Stopped: If the task is expired more than 7 days, it is set as EXPIRED and the VM instance is destroyed. If not, the task remains as STOPPED or becomes finally EXPIRED, depending on the scheduling.

4 Case Study

4.1 Introduction

The proposed architecture and tool are used within the frame of a research strategic project (SENDAI) in the Basque Country which is an industry driven geographic area. There is a Basque Cybersecurity centre and there is a cyber range managed by Tecnalia among other research actors. This region is aware about the need to train people not only in IT cybersecurity aspects but also in OT environments such as SCADAs.

The proposed architecture and tool are used within this cyber range in order to generate an industrial laboratory depending on different requirements such as networks, the number of honeypots, and so on. Concerning honeypots, we have used Conpot and the PLC Siemens S7-200 template.

It is worth to remind that the objective of the proposed architecture and tool is just focused on the deployment phase of an exercise in a cyber range. Therefore, we need to describe the running deployment tool and the scenario.

4.2 Running Deployment Tool

For running the tool, we launch it through the command line and the process automatically launches the Honeypot manager Task Server (Fig. 4), VBox server and the other components. Once we launch the process, Task Server receives a JSON as input parameter, indicating a provider and the honeypots to be deployed, including their start and expiration timestamps. This information is stored in a database as a new pending task for each honeypot.

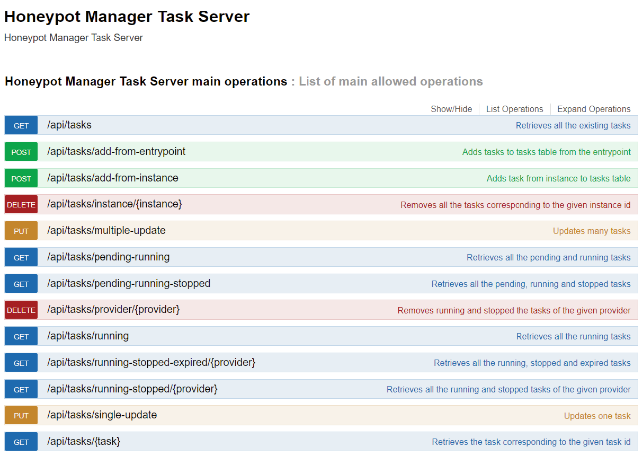


Fig. 4. Honeypot manager task server

Figure 4 shows the different GET, POST, DELETE and PUT messages from the Honeypot Manager.

As stated before, there is a checking process, and Deployer Server is continuously checking, pending and running tasks:

Pending tasks: if a pending task is found, and the current date matches its planning date, the Deployer Server creates a new virtual machine instance for the honeypot associated with the task. Once the instance is up, the honeypot is installed via SSH. Finally, Deployer Server calls Task Server to set the task state as “running”.

Running tasks: if an expired running task is detected, Deployer Server destroys its associated virtual machine instance and calls Task Server to set the task state as “expired”. The non-expired running tasks continue as “running”.

Other considerations: when a running task is expired, and another running tasks are not, Deployer Server destroys all the virtual machine instances, and Task Server sets the expired running task state as “expired” and the non-expired running task as “pending”. This is because Deployer Server is thought to manage virtual machine infrastructures, and not for manage operations like “start” and “stop” instances. So Deployer Server cannot destroy just one instance and leave the others running.

4.3 Scenario Deployed

The scenario deployed by itself, is not critical for this paper because its purpose is just to illustrate that a scenario is deployed. This exercise is based on [38], and on the exercises provided by ENISA [39]. Therefore, the scenario is a tested exercise, and its validity has been already proved. According to [25], our scenario is defined by:

- **Purpose:** students must capture and analyse a S7comm packet.
- **Environment:** the topology of the exercise is a SCADA where the SIEMENS honeypot is used.
- **Storyline:** Students must analyse S7comm packets.
- **Type:** it is a static scenario.
- **Domain:** Network analysis.
- **Tools:** Students will use Wireshark.

As result, students must analyse the S7comm packets as it was provided during the BlackHat conference [40]. Students must identify Type, subtype, sequence, protocol ID.

5 Conclusions

This paper presents an architecture and tool for generating ICS scenarios based on honeypots. The main motivation of this research is to facilitate the deployment of different networks and configurations based on users’ needs. This deployment is used in a cyber range environment which purpose is to train skilled people in industrial settings.

The main contribution of this paper is the deployer architecture and the deployer main process where different virtual machines are managed. Each virtual machine has a honeypot and simulates a SCADA.

The exercise used in this paper is just used for illustrating the process. Our purpose is not to describe in detail the exercise carried out within the cyber range. Obviously, we are still gathering information from the exercises carried out by practitioners. In addition, we are considering different scenarios for different practitioners, and therefore we need to generate different configurations.

One benefits of this approach is that we are able to schedule the behaviour of a VM with the “start”, “stop”, “time_unit”, time_up” and “time_down” parameters. These parameters can simulate a potential the appearance of an intruder, or the time required for finishing an exercise.

This research has a set of limitations. Firstly, the honeypot used is just the Conpot and always the same Siemens configuration because the purpose is to demonstrate our ability to deploy several honeypot within the cyber range. In this sense, several research works are improving the security of S7comm and S7commm plus protocols, and we are starting some research works aligned to these protocols. Secondly, these honeypots are deployed in virtual boxes. We are in the process of refining and improving the approach by adding AWS.

Acknowledgements. This work has been partially funded by the Sendai - Segurtasun Integrala Industria Adimentsurako (KK-2019/00072) Elkartek framework programme from the Basque Government, and partially funded by the SPEAR project (787011) H2020.

References

1. Thames, L., Schaefer, D. (eds.): Cybersecurity for Industry 4.0. Springer Series in Advanced Manufacturing. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-50660-9
2. Lezzi, M., Lazoi, M., Corallo, A.: Cybersecurity for Industry 4.0 in the current literature: a reference framework. Comput. Ind. **103**, 97–110 (2018). https://doi.org/10.1016/j.compind.2018.09.004
3. CCN-CERT. Cyber threats and Trends 2019 (2019). https://www.ccn-cert.cni.es/informes/informes-ccn-cert-publicos/4041-ccn-cert-ia-13-19-threats-and-trends-report-executive-summary/file.html. Accedido 02 July 2020
4. Make UK. Cyber security and manufacturing: a briefing manufactureres. https://www.makeuk.org/-/media/cyber-security-and-manufacturing-a-briefing-for-manufacturers.pdf
5. Thames, L., Schaefer, D.: Industry 4.0: an overview of key benefits, technologies, and challenges. In: Thames, L., Schaefer, D. (eds.) Cybersecurity for Industry 4.0. SSAM, pp. 1–33. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-50660-9_1
6. Baines, T., Lightfoot, H.W.: Servitization of the manufacturing firm. Int. J. Oper. Prod. Manag. **34**, 2–35 (2014)
7. Tropina, T.: Public–private collaboration: cybercrime, cybersecurity and national security. Self- and Co-regulation in Cybercrime, Cybersecurity and National Security. SC, pp. 1–41. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-16447-2_1
8. Fenton, D., Traylor, T., Hokanson, G., Straub, J.: Integrating cyber range technologies and certification programs to improve cybersecurity training programs. In: Auer, M.E., Tsiatsos, T. (eds.) ICL 2018. AISC, vol. 917, pp. 632–643. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-11935-5_60
9. Larrucea, X.: Modelling and certifying safety for cyber-physical systems: an educational experiment. In: 2016 42th Euromicro Conference on Software Engineering and Advanced Applications (SEAA), Limassol, Cyprus, pp. 198–205 (2016). https://doi.org/10.1109/seaa.2016.28
10. Korwin, A.R., Jones, R.E., et al.: Do hands-on, technology-based activities enhance learning by reinforcing cognitive knowledge and retention? **1**(2) (1990). (Spring 1990)
11. Cuppens-Boulahia, N., Lambrinoudakis, C., Cuppens, F., Katsikas, S. (eds.): CyberICPS 2016. LNCS, vol. 10166. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-61437-3
12. Lee, S., Lee, S., Yoo, H., Kwon, S., Shon, T.: Design and implementation of cybersecurity testbed for industrial IoT systems. J. Supercomput. **74**(9), 4506–4520 (2017). https://doi.org/10.1007/s11227-017-2219-z
13. National Institute of Standards and Technology. Framework for Improving Critical Infrastructure Cybersecurity, Version 1.1, 05 dic 2017. https://csrc.nist.gov/publications/detail/white-paper/2017/12/05/cybersecurity-framework-v11/draft
14. Salah, K., Hammoud, M., Zeadally, S.: Teaching cybersecurity using the cloud. IEEE Trans. Learn. Technol. **8**(4), 383–392 (2015). https://doi.org/10.1109/tlt.2015.2424692
15. Network Emulation Testbed https://www.emulab.net/

16. Cyber-Defense Technology Experimental Research Laboratory Testbed. http://deter-project.org/
17. Morelli, U., Nicolodi, L., Ranise, S.: An open and flexible cybersecurity training laboratory in IT/OT Infrastructures. In: Fournaris, A.P., et al. (eds.) IOSEC/MSTEC/FINSEC 2019. LNCS, vol. 11981, pp. 140–155. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-42051-2_10
18. Scaffidi, C.: What training is needed by practicing engineers who create cyberphysical systems? 298–305 (2015). https://doi.org/10.1109/SEAA.2015.19
19. Corallo, A., Lazoi, M., Lezzi, M.: Cybersecurity in the context of industry 4.0: a structured classification of critical assets and business impacts. Comput. Ind. **114**, 103165 (2020). https://doi.org/10.1016/j.compind.2019.103165
20. Kweon, E., Lee, H., Chai, S., Yoo, K.: The utility of information security training and education on cybersecurity incidents: an empirical evidence. Inf. Syst. Front. (2019). https://doi.org/10.1007/s10796-019-09977-z
21. Asghar, M.R., Hu, Q., Zeadally, S.: Cybersecurity in industrial control systems: issues, technologies, and challenges. Comput. Netw. **165**, 106946 (2019). https://doi.org/10.1016/j.comnet.2019.106946
22. Baykara, M., Das, R.: A novel honeypot based security approach for real-time intrusion detection and prevention systems. J. Inf. Secur. Appl. **41**, 103–116 (2018). https://doi.org/10.1016/j.jisa.2018.06.004
23. Naruoka, H., et al.: ICS honeypot system (CamouflageNet) based on attacker's human factors. Procedia Manuf. **3**, 1074–1081 (2015). https://doi.org/10.1016/j.promfg.2015.07.175
24. Conpot. http://conpot.org/. Accessed 17 Mar 2020
25. Yamin, M.M., Katt, B., Gkioulos, V.: Cyber ranges and security testbeds: scenarios, functions, tools and architecture. Comput. Secur. **88**, 101636 (2020). https://doi.org/10.1016/j.cose.2019.101636
26. Ferrag, M.A., Ahmim, A.: Security Solutions and Applied Cryptography in Smart Grid Communications. IGI Global, Hershey (2016)
27. Yao, W.-M., Fahmy, S.: Flow-based partitioning of network testbed experiments. Comput. Netw. **58**, 141–157 (2014). https://doi.org/10.1016/j.comnet.2013.08.029
28. Fang, B.: Positions of states toward cyberspace and cyber-relating regulations. Cyberspace Sovereignty, pp. 243–320. Springer, Singapore (2018). https://doi.org/10.1007/978-981-13-0320-3_8
29. Beuran, R., Tang, D., Pham, C., Chinen, K., Tan, Y., Shinoda, Y.: Integrated framework for hands-on cybersecurity training: CyTrONE. Comput. Secur. **78**, 43–59 (2018). https://doi.org/10.1016/j.cose.2018.06.001
30. Spitzner, L.: Honeypots: catching the insider threat. In: 19th Annual Computer Security Applications Conference, 2003. Proceedings, pp. 170–179 (2003)
31. Dionaea Project. https://github.com/rep/dionaea. Accessed 17 Mar 2020
32. SNARE. https://github.com/mushorg/snare. Accessed 17 Mar 2020
33. HonSSH Project. https://github.com/tnich/honssh/wiki. Accessed 17 Mar 2020
34. Sebek Project. https://projects.honeynet.org/sebek/. Accessed 17 Mar 2020
35. Redwood, O., Lawrence, J., Burmester, M.: A symbolic honeynet framework for SCADA system threat intelligence. In: Rice, M., Shenoi, S. (eds.) ICCIP 2015. IAICT, vol. 466, pp. 103–118. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-26567-4_7
36. Lopez, J., Setola, R., Wolthusen, S.D. (eds.): Critical Infrastructure Protection 2011. LNCS, vol. 7130. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-28920-0
37. Sokol, P., Míšek, J., Husák, M.: Honeypots and honeynets: issues of privacy. EURASIP J. Inf. Secur. **2017**(1), 1–9 (2017). https://doi.org/10.1186/s13635-017-0057-4

38. Hui, H., McLaughlin, K.: Investigating current PLC Security Issues Regarding Siemens S7 Communications and TIA Portal. In: presentado en 5th International Symposium for ICS & SCADA Cyber Security Research 2018 (2018). https://doi.org/10.14236/ewic/ics2018.8
39. ENISA. Introduction to Network Forensics ICS/SCADA Environment Toolset, Document for students (2019). https://www.enisa.europa.eu/topics/trainings-for-cybersecurity-specialists/online-training-material/documents/introduction-to-network-forensics-ex1-toolset.pdf
40. Lei, C., Donghong, L., Liang, M.: The spear to break the security wall of S7CommPlus. https://www.blackhat.com/docs/eu-17/materials/eu-17-Lei-The-Spear-To-Break%20-The-Security-Wall-Of-S7CommPlus-wp.pdf
41. Schuster, S., van den Berg, M., Larrucea, X., Slewe, T., Ide-Kostic, P.: Mass surveillance and technological policy options: improving security of private communications. Comput. Stand. Interfaces **50**, 76–82 (2017). https://doi.org/10.1016/j.csi.2016.09.011

A Systematical Approach for "System Item Integration and Testing" in Context of ISO 26262

Martin Ringdorfer, Gerhard Griessnig(✉), Patrick Draxler, and Adam Schnellbach

AVL List GmbH, Hans List Platz 1, 8020 Graz, Austria
gerhard.griessnig@avl.com

Abstract. The development of automotive systems and components are often in context of a development of a safety critical system with respect to electric/electronic faults. To ensure that these systems are developed adequately safe, standards like the ISO 26262:2018 have evolved which should guide the engineers through the whole development. In this paper a possible testing approach for automotive safety critical systems on system level and on vehicle level is shown. Based on real-world examples, the approach for functional testing, fault-injection testing and robustness testing is investigated in detail. Several important steps which are needed to end up with a fully tested and calibrated safety critical product in context of ISO 26262:2018 are explored. Potential challenges that need to be solved during the testing/calibration activities are highlighted and discussed in detail.

Keywords: ISO 26262:2018 · Functional testing · Fault injection testing · Robustness testing · System and Item Integration and Testing · Safety validation

1 Introduction

The functionality of automotive control systems has increased in the past years. This has led to a continuous increase of complexity within electric/electronic (E/E) control systems. Since E/E systems rely on information sourcing, information processing and actuation behavior (e.g. sensors, control algorithms, electric actuation, …), single failures within the whole chain can cause inappropriate system reactions. In worst case, such failures might lead to unreasonable risks for the human life. From the engineering point of view, special attention must be put on such systems/functions in order to reduce the risk to an acceptable level. To identify such risks and to gain control over possible threads, more and more effort must be put into the development (implementation) and into the testing activities of such safety functions. This is where safety standards e.g. the IEC 61508 [1] or the ISO 26262:2018 [2] enter the game. For the automotive industry the functional safety standard ISO 26262:2018 is of interest, as it deals with the development of series production road vehicles. It is the task of such standards to define guidelines and recommendations on how to identify potential risks and on how to develop a safe product based on the best practices and state of the art.

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 555–567, 2020.
https://doi.org/10.1007/978-3-030-56441-4_42

Among all parts of the ISO 26262:2018, the part 4 is of special interest for this paper as this part defines how to test and validate safety critical control systems on system level and on vehicle level.

Referring to the SPI Manifesto [4], this paper addresses the principle 3 “Base improvement on experience and measurements” and the principle 6 “Use dynamic and adaptable models as needed” by updating of test model based on measured efforts for testing and gained experiences in executed industrial projects.

To give the interested reader an insight on testing of safety related systems, the paper is structured as follows: The first part gives an insight to the approach and to requirements as defined in the ISO 26262:2018 [2] in part 4 towards “System and Item Integration and Testing” (SIIT) as well as “Safety Validation” (SV). The second part explains two different approaches for function testing. Fault-injection testing (Does the device under test (DUT) show a proper safety reaction in case of E/E failures?) and robustness testing (Does the DUT not trigger a safety reaction when not needed?) are explained in detail. The experienced reader may have already be noticed that both methods might be contradicting with respect to calibration. This paper shows a potential way on how to establish a safe and robust system. Finally, a recommendation on the workflow for fault-injection and robustness testing is given and further improvements for future applications conclude the paper.

2 Requirements for Verification

The automotive functional safety standard defines many requirements towards organization, skills, processes, methods and expected evidences for the different phases of the product life cycle. The focus of this paper is on the verification and validation on system and vehicle level which are described in part 4 “Product development at the system level” of ISO 26262:2018. As the standard is written in natural language it is difficult for the reader to understand the systematic of the standard and the connections between the different safety activities. Therefore, a reasonable representation of the content is useful. For this purpose, the Fig. 1 from [3] shows the necessary safety activities which are requested for the “System and Item Integration and Testing” and “Safety Validation”.

The objectives for the SIIT are defined by “(1) define the integration steps and to integrate the system elements until the system is fully integrated (2) to verify that the defined safety measures, resulting from safety analyses at the system architectural level, are properly implemented; and (3) to provide evidence that the integrated system elements fulfil their safety requirements according to the system architectural design” [2].

To achieve these objects the SIIT starts with the “Item Integration and Test Strategy”. Within this plan the necessary test environments and equipment, the methods to derive test cases and the methods to perform tests shall be defined depending on the requested ASIL (Automotive Safety Integrity Level). This plan shall include the necessary integration steps and activities for HW/SW, System and Vehicle Level.

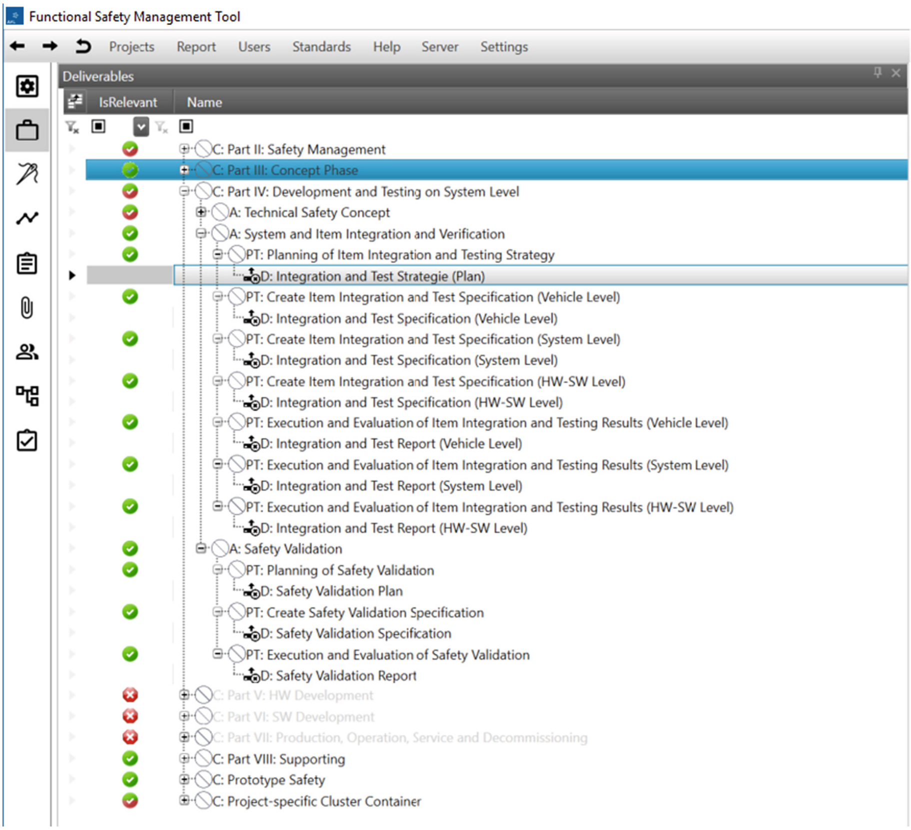


Fig. 1. Functional safety management tool [3]

After the planning activity the test case specification and finally the test case execution and evaluation will be performed (see Fig. 1). For the test case specifications and for the test case executions corresponding methods listed in tables (see Fig. 2) are defined within this standard.

Methods		ASIL			
		A	B	C	D
1a	Requirement-based test	++	++	++	++
1b	Fault injection test	+	+	++	++
1c	Back-to-back test	o	+	+	++

Fig. 2. Example table from [2] for methods for test case specification

The activities to achieve objectives for the SV which are “(1) to provide evidence that the safety goals are achieved by the item when being integrated into the respective vehicle(s); and (2) to provide evidence that the functional safety concept and the technical safety concept are appropriate for achieving functional safety for the item" [2] are similar to the SIIT (see Fig. 2).

It is not the intention of this paper to show that the derived testcases are complete or correct for the SITT and the SV, the focus is on the introduced systematic methodology to verify and validate that the product is safe and robust.

3 System and Item Integration and Testing

Common vehicle control function architectures (control system architectures) are structured in a multi-layer manner like shown in Fig. 3. In fault-free cases it is intended that sensor values or other input information are used by "nominal functions"[1] to determine actuator setpoints. Level 1 functions and calibration define the vehicle characteristic. Its focus is therefore e.g. the creation of a good driving behavior or the achievement of a low fuel consumption. In order to reach the target values, a high effort is spent to find a proper setting of the influencing Level 1 calibration.

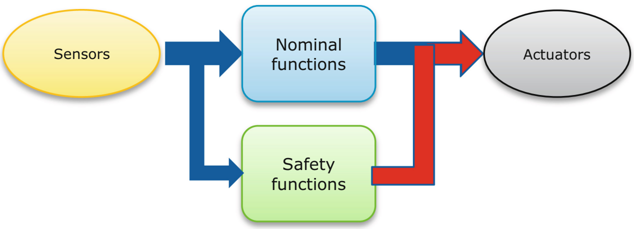


Fig. 3. Signal flow and functions

So-called "safety functions"[2] monitor the behavior of the control system. In case of E/E-fault occurrence, it is the duty of the safety functions to undertake all necessary actions (e.g. overwrite the setpoints from Level 1) to prevent the users[3] from unreasonable harm. The Level 2 functions are therefore instances that are working in the background and are typically not visible for the users. As representative examples for this paper, three different control systems have been tested and calibrated to prove the applicability of the proposed activities: a hybrid control system (HCU) and a transmission control system (TCU) consisting two variants (for conventional powertrains and for hybrid powertrains). The functional requirements for Level 2 were clustered in different safety mechanism and safety related functions (see Table 1). Although Level

[1] According to the E-Gas concept [5] these functions are often referred to as Level 1 functions.

[2] According to the E-Gas concept these functions are often referred to as Level 2 functions.

[3] Users in that sense covers all humans that can be affected by the vehicle or its functions (e.g. driver, passengers, maintenance staff or pedestrians).

2 is strongly dependent on Level 1, details on Level 1 are not explicitly mentioned in here. Testing these Level 2 functions is the challenge that should be addressed in the subsequent sections.

Table 1. Safety requirements and safety related functions

	HCU	TCU conventional	TCU hybrid
# of safety mechanisms/safety related functions	38	44	44
# of safety requirements (functional requirements)	361	412	420

3.1 Fault-Injection Testing vs. Robustness Testing

Before the testing of the safety function can start, it is required to have an aligned function calibration between Level 1 and Level 2. Whenever the Level 1 calibration changes, its relation to the safety function needs to be highlighted as the calibration changes might have an impact on the safety test results. Unintended interferences can appear e.g.:

- Level 2 reactions are too harsh compared to Level 1 reactions (e.g. due to parameter threshold violations)
- Level 2 reactions are not sufficiently debounced such that already triggered Level 1/diagnosis reactions cannot become effective

An aligned calibration data set can be derived by means of mathematical rules. These rules shall describe the relationship between the Level 1 function calibration data and the Level 2 function calibration data. The rules need to be applied to every dataset before the calibration/testing procedure starts (see Fig. 4) in order to make the safety responsible aware about potential interfering control actions.

To verify a software-intensive safety-related system, two testing approaches are necessary for functional testing which complement each other:

- Robustness testing
- Fault-injection testing

Robustness testing is a testing approach which can be used to verify if the system does not intervene when it is exposed to different operating scenarios which do not explicitly include any safety-critical situation (e.g. Is no safety reaction triggered in case of regular temperature conditions?). With respect to the Level 1 calibration dataset, this means that the Level 2 calibration parameter values must be chosen "sufficiently high" to not trigger unintended safety reactions. Considering only this aspect will lead to a robust, but not necessarily safe control system.

Fault-injection testing is a testing approach which can be used to verify if the system intervenes as specified in presence of a potentially safety-critical situation (e.g. Is a safety reaction triggered in case of overtemperature?). Together with the safety validation, it is the task of the fault-injection testing to ensure that the safety reaction takes place when needed (i.e. the Level 2 calibration parameter are chosen “sufficiently low”). Only if a system successfully passes

- a robustness testing campaign and
- a fault-injection testing campaign,

then it is adequately safe and robust with respect to safety-critical situations.

The workflow on how to combine the two approaches from the calibration perspective can be seen in Fig. 4. A detailed description of the efforts to be spent for each test approach and what challenges the engineers face with are explained in the subsequent sections.

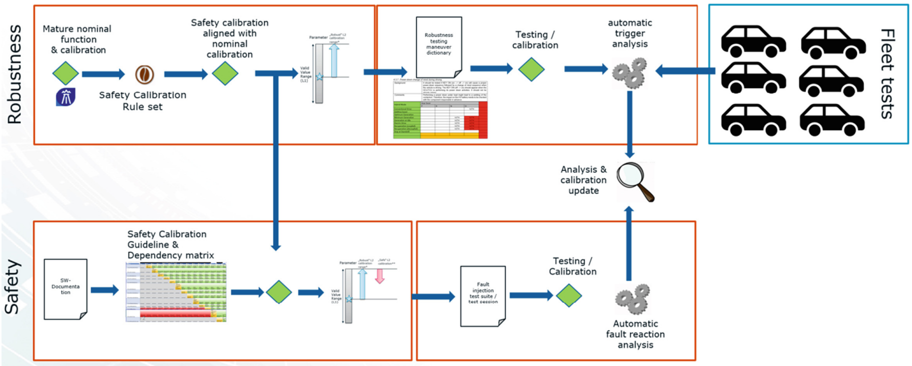


Fig. 4. Calibration workflow for fault-injection testing vs. robustness testing

3.2 Fault-Injection Testing

The expectation with respect to fault-injection testing is a dedicated safety reaction. For this purpose, the system is intentionally provoked by an injected fault. Typically, a fault is injected into the control system functions or into its related communication. By means of fault-injection, the complete fault propagation chain, shown in Fig. 5, shall be checked[4].

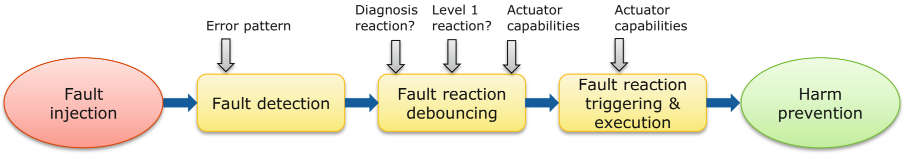


Fig. 5. Level 2 harm prevention strategy

[4] This implies that the fault reaction execution must be always switched on.

Starting points for the test activities are the input signals from the Level 1, the input signals from Level 2 and the functions defined for both levels[5]. All necessary information shall be available within the SW documentation and calibration guidelines (see Fig. 4). Together with the information about diagnosis functions, a formal test sequence (test session) can be derived. This sequence can either be used for manual testing or for automated testing. It is recommended to use an automated test approach to be able to perform the test loops in a time-efficient and highly reproducible way. To do so, the test specification is converted into a test automation sequence (test script) which can be executed e.g. on a Hardware in the loop system (HIL) (see Fig. 6) or on other test systems. It is recommended to define a test sequence is such a way (e.g. generic test description language) that it can easily be transferred between different test environments. The test sequences typically also contain automated check algorithms (pass-criteria[6]), that evaluate the effectiveness of the safety functions with respect to

- quantitative criteria (e.g. Is the reaction triggered when a threshold is exceeded?) and
- timing criteria (e.g. Is the reaction triggered sufficiently fast?).

When performing the test, a very important side-effect must be considered: As shown in Fig. 5, there might be an overlap between Level 1, Level 1 diagnosis and Level 2 with respect to the reaction that each system might trigger. When testing Level 2 functions, this overlap needs to be considered. Example: When Level 1 and Level 2 should trigger the same reactions, it is required to switch-off Level 1 reactions in order to observe just Level 2 reactions.

Due to the need of sharing test resources over different projects, the setup of modular test environments is recommended. To take this into account, a multi-HIL environment is of special interest. Multi-HIL in that sense means that different control systems e.g. HCU, TCU hybrid and TCU conventional can be tested on one HIL test bench. A schematic overview about such a HIL system can be derived from Fig. 6.

For fault-injection testing it is typically required to perform communication tests, where the communication between different control systems is corrupted (e.g. data corruption, message timeout [2]). Furthermore, fault-injection on hardwired interfaces (i.e. sensor lines) are performed. Such error patterns can contain signal drifts, short-to-ground, short-to-supply, … [2] and are typically injected by break-out-boxes. An addition to data corruption/manipulation outside of the control unit, also a fault-injection inside of the control unit (e.g. by manipulating calibration parameters) are applied.

[5] The overlap caused by calibration can be included in MOCA.

[6] In case of a non-successful evaluation of the safety functions, a root-cause analysis followed by rework-steps need to be performed. The same might hold true when calibration changes or functional changes have been applied to the control system.

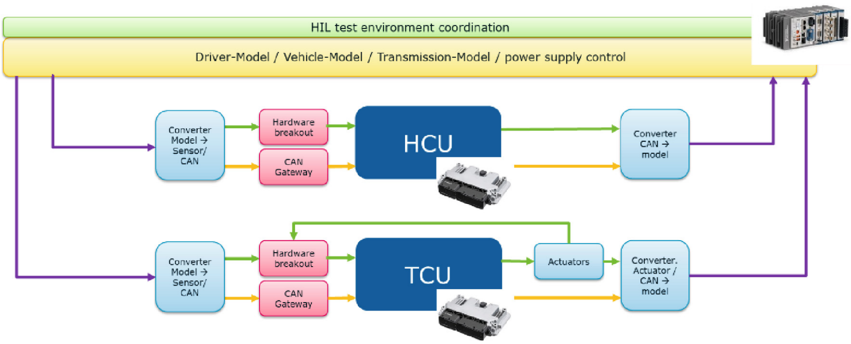


Fig. 6. HCU-TCU-Multi-HIL architecture (simplified)

When the fault-injection tests have successfully passed the HIL environment tests, then selected fault-injection tests are typically performed in the vehicle[7]. This is required to counterprove the reactions under real world conditions. Furthermore, the in-vehicle tests are required to check the functions and their calibration with respect to the validation targets.

3.3 Robustness Testing

The expectation with respect to robustness testing is that the system under test does not trigger any safety reactions when it is exposed to uncritical scenarios or maneuvers. It is a passive method were, in contrast to fault-injection testing, no "active" fault reaction is provoked.

The safety functions must observe the entire system consists of e.g. interfaces, Level 1 functions, sensors and actuators under all possible driving conditions and trigger safety reactions if required. A safety reaction without presence of E/E-faults would lead to customer dissatisfaction. Therefore, robustness testing tries to identify weak points in the safety functions and calibration by covering as much as possible combinations of vehicle driving scenarios and environmental and disturbance factors (e.g. temperature, humidity, EM radiation). Examples for aspects that characterize vehicle driving scenarios are mentioned in Fig. 7.

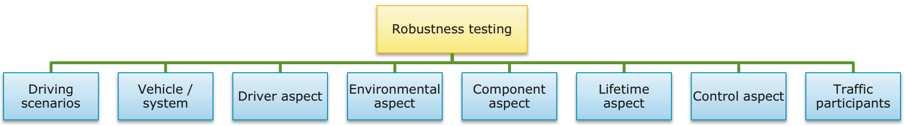


Fig. 7. Impact factors on driving scenarios over a vehicle lifetime (examples)

[7] Having a clever testing strategy in mind and having the test cases specified in a generic (ideally machine readable) manner would support to transfer of the test cases between the test environments.

Combining all possible influencing factors and variations to a set of driving scenarios would lead to an incredible number of test scenarios. As testing all scenarios is unrealistic with respect to development time and development costs, a more structured approach is required. In many projects, robustness testing is performed by a hands-on approach, where the interaction between driver and vehicle was selected to be the impact factor number one.

A robustness testing maneuver dictionary by considering the

- developed vehicle features (electric driving, creeping, snow-rock, …),
- potential driving scenarios (forward/rearward driving, speed range) and
- potential driver inputs (gear lever change, accelerator pedal input, …)

was created (see Fig. 4).

This dictionary contains several hundred scenarios that can occur during a vehicle lifetime. They are the starting point for the robustness test execution. The test execution is typically started in the vehicle (i.e. on a test track)[8]. When the functions have a certain maturity i.e. no unintended safety reactions are triggered, then the defined scenarios are backed-up by real world drives (e.g. test trips driven in real traffic situations). A robustness testing maneuver library is generated and can be tested by means of a complete vehicle fleet. The fleet test data is collected via a data logger and transferred via a cloud solution into a back-end where a further data processing (e.g. big data) is performed.

Figure 8 shows the steps that are required to achieve robust safety functions. Due to the fact, that all the functions are observed at the same time, and the scenarios cannot be exactly repeated, the amount of data that must be logged and analyzed increases dramatically compared to fault-injection testing. Efficient checking of each function manually would not be possible. To overcome this situation, an automated pre-analysis mechanism is applied. As a result of the pre-analysis procedure, a test report is generated. This report provides the safety experts a summary about all potential safety triggers including the information of trigger time and source[9]. Instead of analyzing all the data manually, the analyzing team just needs to focus on the trigger flag in the time window in which the safety intervention occurred.

[8] To perform an efficient robustness testing, it is proposed to "guide" the test driver through the test maneuvers. This is done by a human-machine-interface that gives the driver an explanation of the intended test before the test shall be performed. After the test has finished, a check routine is automatically triggered. The result of this check routine will immediately inform the driver if the test maneuver was successful or the test must be repeated.

[9] Please note, that depending on the test progress the "fault reactions triggering & execution" (see Fig. 5) might be switched off. The tester has no immediate feedback about potential safety triggers and its related fault reaction. As some safety triggers are only healed by an ignition cycle, a visual and audible HMI notification is provided to the tester.

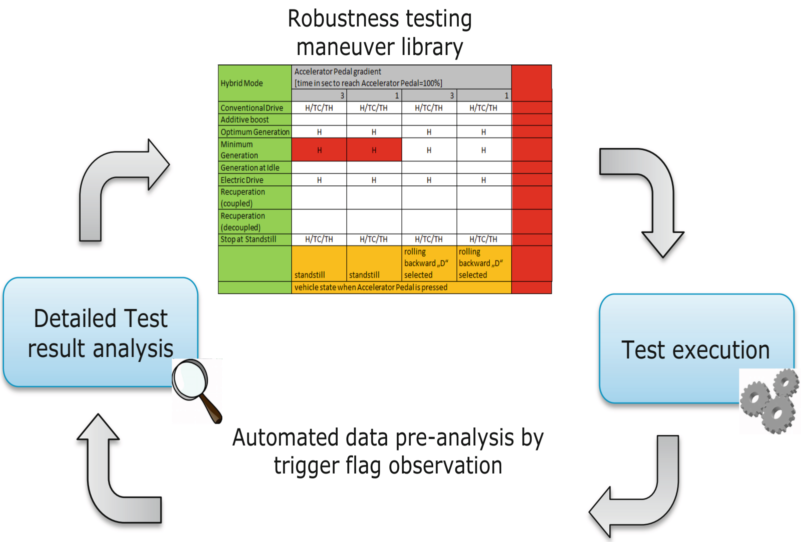


Fig. 8. Robustness-test cycle

When having robust safety functions under nominal conditions, it is recommended to increase the overall function and calibration maturity. This can be done by introducing further impacting factors (see Fig. 7). One possible approach to increase the maturity is the repetition of the tests under real-world conditions (e.g. by performing tests in different vehicles or with different drivers). As this possibility is time-consuming and expensive, another approach is of preferable interest (see Fig. 9). The preferred approach reduces the previously mentioned disadvantages, as a "semi-virtual" test environment is utilized. Semi-virtual in that sense means, that vehicle measurement data is introduced on a HIL as target values for the virtual environment. E.g. the measured vehicle speed is trusted to be the target speed for a HIL driver model. System parameters (e.g. lifetime aspects, environmental aspects, …) can be varied and the advantages provided by a HIL (e.g. 24/7-utilization) can get a chance.

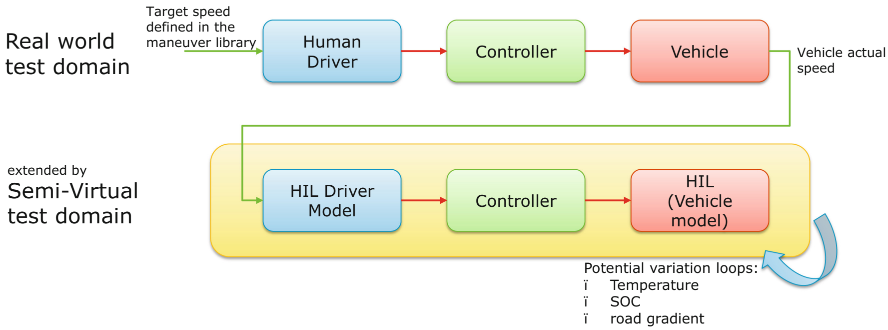


Fig. 9. Increase of function maturity by a semi-virtual test approach

Robustness testing is typically performed in parallel to the fault-injection testing or in a sequence like shown in Fig. 10. Special attention must be put on the fact that robustness testing and fault-injection testing have contradicting goals with respect to calibration. While fault-injection testing typically leads to a reduction of the safety margin, robustness testing typically tends to enlarge safety margins.

3.4 System Item Integration and Testing

In the previous sections, the focus was purely put on the test methods used within the safety function development. Remembering Fig. 3, it is obvious that the nominal functions and its related diagnosis functions must be considered in the safety development as well. Experience has shown that the functional growth and the reconcilement of all functions are done in a sequence like shown in Fig. 10.

Starting point for all required activities are released Level 1 functions[10]. After performing initial integration steps, the initial robustness is tested (second part in Fig. 10). As the safety functions are introduced the first time, it is required to minimize the interaction between Level 1 and Level 2. This is ensured by deactivating the diagnosis reaction in the Level 1 software. At a later stage, when the initial robustness tests are passed, the diagnosis reactions are switched on. Safety functions and diagnosis functions can now be harmonized.

As shown in Fig. 10 (third part), the safety test development is continued by performing fault-injection testing. Following the fault propagation chain mentioned in Fig. 5, the fault detection needs to be checked first. When the detection mechanisms work as expected, then the safety reactions are switched on. From that time onwards, the whole propagation chain can be tested and calibrated.

In a final step (fourth part in Fig. 10), it is required to activate the all currently switched-off functions (e.g. diagnosis) to check the overall interaction of Level 1 and Level 2. It is recommended to perform a final loop of robustness tests and fault-injection tests to conclude the item integration and testing[11].

[10] In minor cases it is also possible to start from a sufficiently mature Level 1 function. Level 1 function calibration and Level 2 function testing may go in parallel but bear the risk of function mismatch leading to additional development efforts.

[11] Between the single test steps shown in Fig. 10 it might be required to change the implemented functionality (e.g. bug fix or calibration update). Whenever this takes place, it might be that parts of the procedure must be repeated.

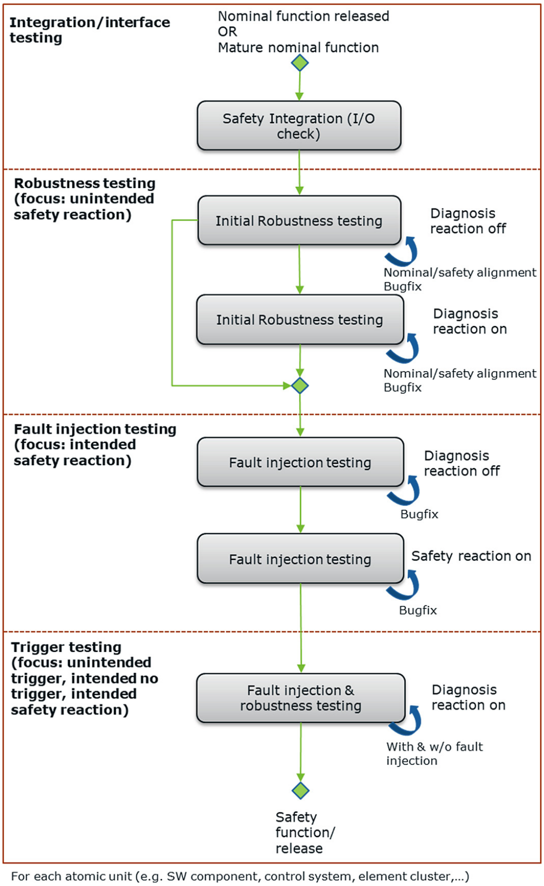


Fig. 10. Safety function testing and calibration workflow (process view)

4 Future Work

Although some proposals for future improvement have been stated in the paper, there are still unexploited potentials. It is planned to further increase the test depth and the test efficiency over the whole lifecycle.

Based on the recommendation of the ISO26262:2018 it is intended to enforce semi-formal and formal methods/approaches. One example is the application of additional formal aspects related to requirements engineering and test specification. Requirements are typically written in unconstrained natural language, which makes them prone to problems like for example ambiguity, complexity and vagueness [6]. Especially the specification of requirements/test sequences in a controlled natural language as e.g. proposed in [6] sounds promising for further automation.

First tests in introducing patterns already in an early development phase have shown very promising results (i.e. consistency checks, …). The gained results and the available lessons learnt will be used to extend and improve the test specification, test automation, test execution and test analysis activities.

5 Conclusions

This paper shows a systematical and "proven in use" approach for "System Item Integration and Testing" and "Safety Validation". It highlights and describes the necessary steps to perform a proper item integration in context of functional testing. A special focus is on the interaction between the two test approaches of fault-insertion and robustness testing. Details of both test methods and its relationship to different test environments are shown. Furthermore, it illustrates that the development of a safe product is not a straight-forward task. Especially the verification and validation activities are of big importance to bring a safe product on the market. Finally, to further improve the verification and validation activities an outlook on promising next steps is given.

References

1. International Electrotechnical Commission: IEC 61508 Edition 2.0: Functional safety of electrical/electronic/programmable electronic safety-related systems (2010)
2. International Standardization Organization: ISO 26262: Part 1-12: Road vehicles – functional safety (2018)
3. Wambera, T., Macher, G., Frohner, B.: Prozesssteuerung und domänenspezifische Dokumentation nach ISO26262 und ISO25119 während der Entwicklung und Integration von sicherheitsrelevanten Systemen, Diagnose in mechatronischen Fahrzeugsystemen XIII: Neue Verfahren für Test, Prüfung und Diagnose von E/E-Systemen im Kfz. TUDpress, Dresden (2019)
4. Pries-Heje, J., Johansen, J.: The SPI Manifesto (2009). https://2020.eurospi.net/images/eurospi/DownloadCenter/spi manifesto.pdf
5. Arbeitskreis EGAS: Standardisiertes E-Gas Überwachungskonzept für Benzin und Diesel Motorensteuerungen. Arbeitskreis EGAS, Version 6.0
6. Mavin, A., Wilkinson, P., Harwood, A., Novak, M.: Easy approach to requirements syntax (EARS). In: 2009 17th IEEE International Requirements Engineering Conference (RE 2009), pp. 317–322. IEEE, Atlanta (2009)
7. Holtmann, J., Meyer, J., von Detten, M.: Automatic validation and correction of formalized, textual requirements. In: 2011 IEEE Fourth International Conference on Software Testing, Verification and Validation Workshops, pp. 486–495. IEEE, Berlin (2011)

Experiences with Agile and Lean

Clean A-SPICE Processes and Agile Methods Are the Key to Modern Automotive Software Engineering

Improvement Case Study Paper for EuroSPI 2020 Keynote of Marelli Automotive Lighting

Thomas Winz[1], Sebastian Streubel[2](✉), Claudia Tancau[3](✉), and Shreyas Dhone[4](✉)

[1] Software Project Management, Marelli Automotive Lighting Reutlingen (Germany) GmbH/EES2, Reutlingen, Germany
thomas.winz@marelli.com
[2] Software Development, Marelli Automotive Lighting Reutlingen (Germany) GmbH/EES4, Reutlingen, Germany
sebastian.streubel@marelli.com
[3] Methodologies and Tools, Marelli Automotive Cluj S.R.L, Cluj-Napoca, Romania
claudia.tancau@marelli.com
[4] Process Engineer, Marelli Automotive Lighting Reutlingen (Germany) GmbH/EES2, Reutlingen, Germany

Abstract. With the development of our new generation of "high definition digital headlamp" we've succeeded in making the light distribution more flexible than ever before. "In the historical approach, which locks the requirements and delivers the product all in one go, the results are all or nothing. We either succeed completely or fail absolutely. The stakes are high because everything hinges on work that happens at the end" [1]. Our case study will highlight the challenges and process mitigation measures inside the development of a "light distribution as a software system" series production project.

In order to realize series production, the creation of our first Agile project within an Automotive SPICE (A-SPICE) framework is needed. Our case study is performing two A-SPICE Assessments to mitigate the unreasonable technical risks of the development of "light distribution as a software system". To achieve A-SPICE level 2 rating the planning and performance of software process improvement measures are essential.

Keywords: Assessment and process models ISO/IEC 15504 · Improvement case studies · Implementing A-SPICE · Implementing agile · Requirements

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 571–586, 2020.
https://doi.org/10.1007/978-3-030-56441-4_43

1 Introduction

The alignment of such project necessities as product development and process improvement activities is a very challenging issue for today's companies. Our deep conviction is, that any project developing groundbreaking technologies has the need for robust and sophisticated processes. In general, standards should give guidance to projects and their teams. At the same time the improvement of customer satisfaction and technology sophistication have a direct correlation to the performed process improvement activities.

Our Paper shall focus on the SPI Manifesto value: "We truly believe that SPI must involve people actively and affect their daily activities" and the corresponding principle: "Know the culture and focus on needs" [2]. Our case study shall provide an example and further evidence for a beneficial relationship between the values and principles of the Software Process Improvement Manifesto.

Our Paper is organized as follows: First we will relate this work to others and perform a general literature review. Then we will proceed with an in-depth presentation of the Improvement Case Study divided in four parts: Initial Project Situation Evaluation (first A-SPICE Assessment Result), Improvement Planning, Improvement Implementation, Improvement Evaluation (second A-SPICE Assessment Result). The paper will be closed with an overall conclusion.

2 Related Work

Process management is evolving into a crucial instrument for the management of projects and organization of companies. Timo Füermann and Carsten Dammasch have written in their book Process Management: "Today the process management can be seen as a decisive success factor, especially at a time when companies are characterized by increasingly complex structures, processes and products and in which the demands of customers continue to increase. Process management is a procedure that creates an overview and counteracts the growing complexity. In this way, added value can be increased and customer satisfaction rises. To do this, the requirements [of the process] must first be determined and then the structures changed so that fulfillment is possible at any time. When this is done, the processes are continuously improved and stabilized" [6].

There are several Software Process Improvement models (CMM, ISO and others) but the EuroSPI conferences present and discuss overall results from systems, software and services process improvement and innovation (SPI) projects in industry and research [2]. This year's event is the 27th of a series of conferences that focusses on the gained benefits and the criteria for success 10 years after the inception of the SPI manifesto [2] (Table 1).

Table 1. Value and principle of the SPI Manifesto [2]

	Principles				Value
	We trust that the following principles support the values				We truly believe that SPI
People	Know the culture and focus on needs	Motivate all people involved	Base improvement on experience and measurements	Create a learning organisation	**Must involve people actively and affect their daily activities** NOT to show-off or be focused on management alone
Business	Support the organisation's vision and business objectives	Use dynamic and adaptable models as needed	Apply risk management		**Is what you do to make business successful** NOT to live to deploy a standard, reach a maturity level, or obtain a certificate
Change	Manage the organisational change in your improvement effort	Ensure all parties understand and agree on process	Do not lose focus		**Is inherently linked with change** NOT continuing as we do today

EuroSPI's mission is to "develop an experience and knowledge exchange platform for Europe where SPI practices can be discussed and exchanged, and knowledge can be gathered and shared" [2]. It represents the necessary conditions to perform software process improvement activities with a high degree of success. The SPI Manifesto provides the context and the focus for the respective conference workshops and research papers of the congress.

2.1 Relation with Respect to EuroSPI 2019 Manifesto

SPI Manifesto was not known to team members throughout the improvement phase. Our case study does not propose new values and principles to be added to the SPI Manifesto. It highlights that our improvement activities show a beneficial relationship to the values and principles. Furthermore it confirms the emphasis on SPI Manifesto value: "We truly believe that SPI must involve people actively and affect" and the corresponding principle to "Know the culture and focus on needs" is a necessity for achieving a successful process improvement [2] (Fig. 1).

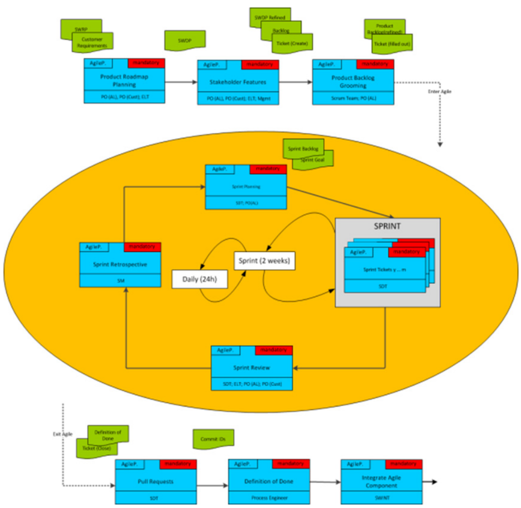


Fig. 1. Adapting standard Scrum process to company culture

It is confirmed that "Cultural values are deeply ingrained and rather resistant to change resulting in conflict if ignored. The way SPI initiative is introduced is imperative for success" [2]. The process improvement team in our case study needed to culturally translate the general A-SPICE practices, development concepts and Agile Scrum methods into a common language that the project team members are familiar with it [7]. That means to re-use familiar inhouse development methods and document naming rules. Those steps supported the team members to use their existing knowledge and have helped them to adapt to the expectations of new process improvement.

2.2 Relation with Respect to EuroSPI 2019 Improvement of Agile Software Development Process Based on Automotive A-SPICE: A Case Study

Komiyama work has "the goal to introducing Agile development conforming to A-SPICE with the goal to define and hybrid processes embodying both Agile and waterfall aspects and faces the issue of Harmonization of Scrum and A-SPICE by clarifying and resolving contradictory factors" [8].

The first deviation to our case study is the hominization efforts of Scrum and A-SPICE processes. In our case study only a project planning alignment and a process compliance proof are necessary to embed an Agile method into an A-SPICE context. The second deviation is the developer training effort of Komiyama work and our case study. In our case study Agile developers are enhanced regular project developers following the rigid A-SPICE process. This fact is successfully demonstrated within the second assessment. The third deviation to Komiyama refers to his commitment: "We will deploy the established processes to other area". In our case study the project process will not be deployed to other areas.

The last deviation that shall be identified is that the Komiyama's work implements "hybrid processes embodying both Agile and waterfall aspects". In our case study there is one standard project process that has fully A-SPICE capability. Within our case study the Agile sub team uses the standard of Agile Scrum methods [7]. The Agile Scrum method is aligned via compliance process matrix. See Table below (Table 2).

Table 2 Agile process compliance matrix against project process

Process	Applicable	Deviation	Comment
System requirements analysis and architectural design (SYRD)	Compliant	Handling of cust. req.	
Software requirements analysis and architectural design (SWRD)	Partially	Handling of cust. req.	
Module requirements analysis and design (MDRD)	Compliant		Agile methods are supplement to development methods
Software construction (SWCON)	Compliant		Agile methods are supplement to development methods

(*continued*)

Table 2 (*continued*)

Process	Applicable	Deviation	Comment
Module & SW INT test (MSIT)	Compliant		
Software testing (SWT)	Partially	LightSystem integration	
System test (SYT)	Compliant		
Product release (SPL.2)	Compliant		Part of current series release
Quality assurance (QA)	Partially	DOD SPICE enhancement	Agile methods are supplement to development methods
Configuration management (CM)	Compliant		
Problem resolution and change request management (PCM)	Compliant		
Project management (PM)	Compliant		

3 Improvement Case Study

The common Software Process Improvement period can be segregated into four distinct parts and those parts can further be separated with respect to the methods and techniques used [4]. This paper represents the following four periods:

1. Initial Project Situation Evaluation (First A-SPICE Assessment Result)
2. Improvement Planning
3. Improvement Implementation
4. Improvement Evaluation (Second A-SPICE Assessment Result)

Our case study is exceptional because of the development of “light distribution as a software system” in an Agile manner embedded in series production project in the context of A-SPICE Level 2. To be useful for our case study, the A-SPICE improvement of the overall project processes is too cumbersome. Our paper mainly focusses on the two A-SPICE processes “MAN.3 Project Management” and “SWE.1 Software Requirement Analysis” highlighting the importance of requirements handling. The critical linkage between Agile and A-SPICE is directly inherited from the above mentioned process and will be described in our case study in a general manner.

3.1 Context of the Case Study

The process improvement is performed within the context of a series AUTOSAR project. The duration of our case study from kickoff to start of production is three years. The first assessment is planned at the midst of the project runtime. The second assessment is planned just before the start of production.

The aim of our case study is to develop the headlight system consisting of two headlights assemblies which will be installed in the front of the vehicle. Earlier headlamp versions had been developed with a light distribution behavior using mechanical movements in the headlight mechanical systems. A highly innovative technology is created that uses new concepts for computing light distribution for high-resolution headlamps for the first time: “light distribution as a software system”. This new technology is established with our customer as partner. Our customer provided system requirement specification at car level including safety and security aspects that needed to be fulfilled.

The Feature Rollout Plan (FRP) of the case study contains 42 High-Level Features (e.g. fan driver control or measurement of temperature), which are split into 137 Low-Level Features. Each Low-Level Feature is represented by a high number of requirements, having in total a couple of hundreds of functional requirements. Each High-Level Feature is assigned to a maturity grade according to the customer release cycle. An example of a Low-Level Feature Structure could look like this:

- Technical Specification 1
- Technical Specification 2
- Technical Specification n + 1
- Fault Handling
- Diagnostic

The “light distribution as a software system” development represents four of those HLFs and contains a couple of hundred functional requirements. In addition to the new technology that is being used, the software requirements are not in a frozen state at the start but are evolving during the runtime as per our case study. Furthermore, the goal is to create completely new light functionalities together with our customer. It is agreed between our customer and us that “light distribution as a software system” is therefore evolving in an Agile Development approach.

An Agile Process needed to be created and embedded within the case study process context. The Agile method of Scrum is used and it follows the standard Scrum guide [7]. To respect and support the special situation of the “light distribution as a software system” the subproject needed to enhance the case study processes in the following proceedings:

- Elicitation of “light distribution software system” requirements
- Definition of “light distribution software system” requirements
- Define and refine the interfaces
- Implementation of the requirements for “light distribution software system”
- Define and improve the Agile Development Process

A-SPICE - Agile - Project Process Motivator

For purpose of our case study the “light distribution as a software system” should support animations that run when the car is locked. The software requirements for these animations were not approved at the beginning of the case study because it was not clear how animations could be realized with the new technology. We started with a simple implementation that supported hard coded animations and fulfilled the bare minimum expectations. Based on this simple implementation, requirements for a more sophisticated approach were generated.

During development, the idea of implementing animations with a play pack of an actual video evolved. The Agile team started with a feasibility study, verified limitations and developed a proof of concept. Based on this, requirements could be created, and the series software solution was developed in a series of sprints along with necessary tools.

For a fast and detailed refinement, an One Pager method was chosen to fulfill the quality aspects of a formal requirement. To have the feature verified within a sprint, unit tests where generated and a static code analysis was performed. Establishing pair programming and peer reviews activities for parts of our code was needed to distribute knowledge and know-how. The customer verified the accomplishments of each sprint. Together with our Product Owner the Agile team adjusted the targets of the next sprint based on the results of the previous sprint.

The possibility to play customized videos fully fulfills the original requirements and excels our customer expectations. Instead of having a fixed animation, the videos are fully flexible opening a vast range of possibilities without any technical draw backs. The customer can include their full corporate design in the video showing the benefits of a high-resolution light display systems. Any car variant can have its own animation without adding extra logistic effort.

This example illustrates the potential of continuous improvement of technology during our case study.

3.2 Initial Project Situation Evaluation

The first Assessment results did not meet the expectations of our customer and our management. Inconsistency between different planning levels, missing details on effort estimation as well as effort monitoring and detailed planning set for short durations had the consequence that the sprint and the release status reporting provided an incomplete information overview of the case study (e.g. implementation of customer requirements). Overall the quality of review was not proportionate to effort spent. Review documentation was inconsistent with the definition of review types.

Assignment of customer requirements to a process and metrics were based on an inconsistent and incomplete bilateral traceability and requirements coverage concept. Therefore, the requirements traceability had gaps and wrong links that could not be detected during extensive reviews. But project management process strengths are robust roles description and performed roles trainings.

Considering state of the project, emphasis shall be taken to concentrate on product and overall system, as looking at single processes or base practices is not enough. The main recommendations of the first assessment were:

- Consider the object of the project as a system and not as single elements, requirements or test cases, etc.

- Review and rework the traceability, starting from customer requirements, to ensure the traceability and consistency between customer requirements and design, implementation and verification (including rework of test levels defined in requirements).

3.3 Improvement Planning

After the first planned assessment in our case study the gap analysis is performed by Quality. It concludes that individual improvement tasks encompassing all processes must be fulfilled to meet the A-SPICE Level 2 rating. Our management decided to trace the process improvement by bi-weekly meetings alongside with our customer. The Agile development process for the development of "light distribution software system" has 11 improvement tickets. A short compilation of improvement tickets is:

- Requirement Handling: The Definition of Done (DOD) of the Agile process needs to confirm if the requirements have changed or not and add the references to the changed requirements.
- Quality assurance needs to check on regular base if the Agile process was not misused by changing/writing code without updating requirements or architecture. Otherwise we risk the coverage of hardware in the loop (HIL) and integration test.
- Training the new Agile manner approach expects that developers are empowered to do software requirements, links, SW architecture, detailed design, software development, unit test design, unit test execution, have safety know-how, etc., all together. Give us a proof that these developers are really empowered in this multi-competent way.

3.4 Improvement Implementation

Implementation involves executing the process improvements according to the gap analysis. It involves the structuring of task in task force teams and creating the necessary space in the release plan for different task forces to implement the measures. Our case study wants to highlight some of these task efforts in regard to the corresponding principle "Know the culture and focus on needs" and the Agile Scrum process [2, 7].

Planning Effort

Our case study provides – with its feature rollout plan and its release plan – a tight time frame. There is no necessity to plan smaller increments. The Agile sprint plan has been aligned to the release plan. Each sprint delivery increment is placed within a release integration time frame. All other planning efforts are focused on the adaptation of different stakeholder expectations onto one consistent time frame for each release. An overview is given in the diagram below (Fig. 2).

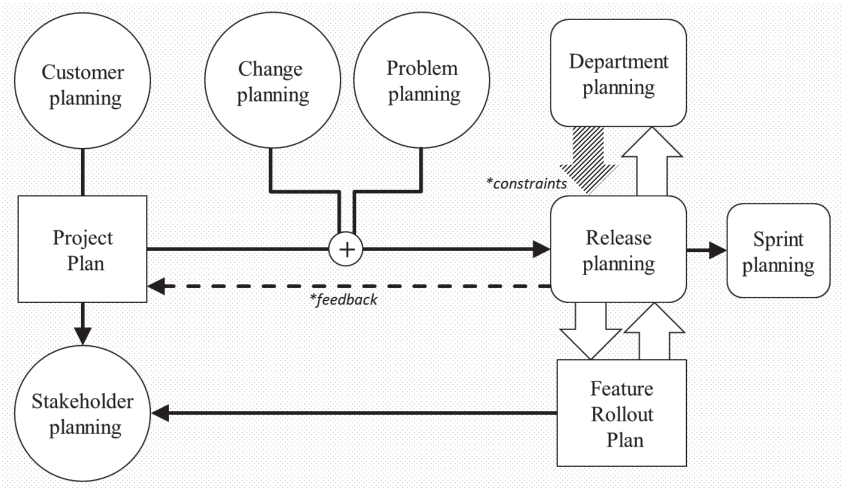


Fig. 2. Relationships of individual plans within our case study

The project management has many opportunities in our case study to eliminate uncertainties and unsureness, providing clarity: clear goals, understandable information, effective communication and clear division of tasks. The "light distribution software system" had many stakeholders that needed to be considered in software release planning for the first time. That is the reason for creation of structure plan in the case study, which shows how the individual stakeholder plans are related.

Missing the deadlines is a problem to the project team only if the members do not have a clear idea about their responsibilities. An example for individual responsibilities is the introduction of checklist for the AUTOSAR integration to give each sub-team (diagnosis, basic software, Agile development) the ability to coordinate deadlines with the customer in the case study.

Standardization Efforts of Management Reporting

Standardized reporting management systems help to deliver higher level of efficiency in projects. It ensures that project team members share the same best practice knowledge to create their reports in a similar way. Especially in software projects, one of the central questions which arise is how effectively the project progress is presented to management. It is a great challenge to prepare the essential data in the right scope.

Our Projects and therewith development cycles tend to get shorter with each product version. At the same time projects must be finished on certain milestones and within a given budget. Due to this, the expectations and the pressure on our new software applications rise constantly. To be able to foresee and intervene subsequently, there is a monthly upper management meeting where project situation and progress are reported to the management by the project team. The so called "System & Software Metric

Dashboard" is a metric for planning, controlling and measuring financial project performance. It should contribute to compare the plan and the actual values of one or more measurement figures, typically plotted against time. Software development metrics are quantitative measurements and can be understood easier by management. They measure the software performance of software teams. The measures include Series Volume, Development Costs, Costs and Invest.

The table below describes the metric, marks it with an ID and assigns the responsibilities. It should help to plan and monitor or track the system and software project deliverables (Table 3).

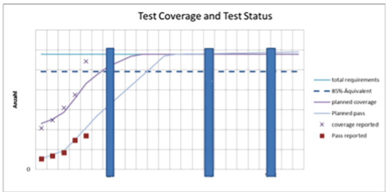
Table 3 Example of "System & Software Metric Dashboard"

Description	ID	Content	Update	Responsible role
Software maturity grade plan	MAN3_4	Diagram	2 times per release	SWR (SW project lead)
Software maturity status	SWE1_1	Planning, tracking	Weekly	SWR
Bug status	SUP10_1	Planning, tracking, diagram	Weekly	Problem resolution manager
Development requirements status	SWE_2_3	Planning, tracking	Weekly	SWR
SW test status	SWE_6	Planning, tracking	weekly	Test manager
SW unit test status	SWE4_1	Planning, tracking	Weekly	SWR
System requirements status	SYS2_1	Planning, tracking	Weekly	Project lead
System test status	SYS5_1	Planning, tracking	Weekly	Test manager

Except for the metric "software maturity grade plan" which is to be updated one to two times per release, the metrics were updated weekly. The status is reviewed in a weekly round of projects team members and responsible management stakeholders. During these meetings necessary actions can be defined and addressed.

Each metric needs to be described in a systematic manner. This includes the name, the metric owner, the main purpose and involves additionally a graph. This graph needs to be described in the same sheet. Finally, each metric must be readable by the user within a short time. It has a significant positive impact on the success of system- and software requirements and below is a case study example of "Metric Characteristic SW Test Status" (Table 4).

Table 4 Example of "Metric Characteristic SW Test Status"

	Description		
Metric Name	SW Test Status		
Metric Owner (role)	Test Manager		
Purpose	Graphical figure, showing the planned test coverage and test status and comparing both with actual test progress		
Abscissa (x-axis)	time in -		
Ordinates (y-axis)	no. of requirements		
Metric KPI		*Type*	*Source*
(1) Total Requirements	Total no. of requirements to be tested in test	goal line	requirement management tool
(2) Coverage (planned)	Planned no. of requirements to be tested	planning curve	SWR
(3) Test Passes (planned)		planning curve	SWR
(4) Coverage (actual)	Actual no. of requirements which are covered by test case	weekly tracking points	requirement management tool
(5) Passed (actual)	Actual no. of requirements which are tested with " pass in testing	weekly tracking points	requirement management tool

Requirements

By structuring the requirements and completing the requirement modules for system and software requirements, a rough analysis is already done by the Requirement Engineer based on their knowledge and experience. The assignment of the requirements to high-level features and low-level features implies the identification of dependencies between function and requirement. A detailed analysis by experts in order to identify and understand cause-effect relationships and providing a basis for problem solving and decision making must be followed: "According to past studies, approximately 60% of all error in system development project originate the phase of requirements engineering. These errors, however, are often discovered only in later project phases".

The Agile requirements are written together with customer; they are software detailed design (module) requirements but a part of our customer requirements. Planning of the Agile process overlaps the planning of the standard software development, so Agile requirements are treated in the same way as the rest of the requirements to consider the VDA guideline and to ensure that quality assurance is performed objectively without conflict of interest [5].

The same reviews, documentation criteria and coding guidelines apply to Agile methods and work products as to normal standard development process work products re-enforcing the VDA guidelines related to Automotive A-SPICE compliancy [5].

The purpose of the analysis and review process of requirements and their dependencies is to ensure correctness, completeness, technical feasibility, verifiability, traceability and to support the identification of risks. The impacts on costs, timing, technical concept, interfaces between the specified systems and other elements of the operating environment must be analyzed. Analysis criteria and involved roles are given in the requirement analysis (RA) template which must be used in order to document the analysis.

KPIs and metrics are defined to track the completeness of requirements and traceability according to the workflow of the requirements. For example, the defined KPI for Stakeholder Requirements is the traceability to our requirements. This means that the defined customer requirements should be linked 100% according to the Workflow of the requirements – the Link Model.

Besides the standard software project process as defined by us, there are two aspects particular to our case study: One is represented by an One Pager presented to the customer to establish technical features before they are officially written as requirements. The other is represented by the Link Model - Workflow of the requirements.

Elicitation of Requirements in an Agile Manner

To fulfill the quality aspects of a formal requirement, a requirement specifies a property of a software system that is technically necessary, verifiable, and attainable. It is expressed as a single sentence, using simple, precise and unambiguous vocabulary, making usage of the imperative “shall” to express a mandatory and contractual requirement. However, as Bergesmann describes there is also a project management aspect with respect to requirements: “Before the requirements can be meaningfully estimated and planned by the team, it must be ensured that they are as comprehensible, complete and in sufficient detail to be represented. In addition to the comprehensibility, completeness and level of detail, there are several other quality criteria in requirements engineering that must be met, as well as techniques for checking them. These can help prevent misunderstandings … from the outset and increase efficiency” [3].

Due to continuous improvements of technology during our case study, the requirements elicitation cannot end by a given final date of completion. Hence the software requirements are not in a frozen state at start of and during the case study. For a fast and detailed refinement, a One Pager method is chosen. For each Agile feature as described in Sect. 2.2 a separate One Pager document is created. They serve the purpose of defining the “How a feature shall be implemented” and “Why a feature shall be needed”. This enables us to improve the requirements in a cooperative way with our customer to make the best use of new technology and prevent misunderstandings. The One Pager method also solves the coordination issue between different stakeholders (e.g. design team on customer side, project leads, light departments, validation teams on both sides, customer and us), which do not necessarily use the same requirements management tool.

The One Pager documents contain informal requirements with enough details and they can be translated to formal requirements. The documents are structured in the following way:

- Purpose of the feature
- Context frame in the project
- Customer expectation
- Rough overview of necessary tasks
- Prove of Evidence
- Relation to existing formal requirements

The Agile process approach is supporting the continuous improvement of the system features specification using the One Pager Method as a basis of our development efforts. One Pager documents are exchanged on regular basis with the customer. Before each sprint, they include the current state of the development and the tasks to be performed.

The software implementation and function description are generated during the development process of the feature and included to the One Pager as well. This allows the PO the possibility to distribute increments of the features to future sprints according to stakeholder and our customer feedback and in alignment to the Feature Rollout Plan. This transition shifts the focus of the requirement elicitation process from overall project to focus on the increment to the next increment development.

The One Pager documents are used for generating detailed, understandable, complete and formal requirements for each sprint. Therefore, they include the necessary information for the Agile team to estimate the development efforts for the consecutive sprints. After accepting an increment of a feature by the Product Owner for the next sprints, the transition to formal requirements is done by our customer in conjunction with the One Pager. By the start of the current sprint the formal requirements should be in a frozen and released state in our requirement management tool.

Workflow of the Requirements – the Link Model

The Link Model represents the workflow of the requirements, from receiving them from the customer and refining them into our system down to the software detailed design level and software units, all this is done in the requirement management tool, to the test environments where the requirements are tested with the test results. The Link Model is a representation of the V-model implementation of the case study. It shows the relation and the traceability between requirements and requirement-related work products.

There are several ways to ensure the bi-directional traceability between requirements and work products. For requirements which are described in requirements management tool, the traceability is realized by tool functionality. Other methods to ensure traceability are naming conventions (e.g. from software detailed design to software units) and test reports which contain the unique requirement ID attached to the test case ID.

A way in which the Link Model is used represents the possibility of direct linking to customer requirements the system architecture requirements, software requirements, software architecture requirements, hardware architecture requirements and requirements related to the software and hardware detailed design (Fig. 3).

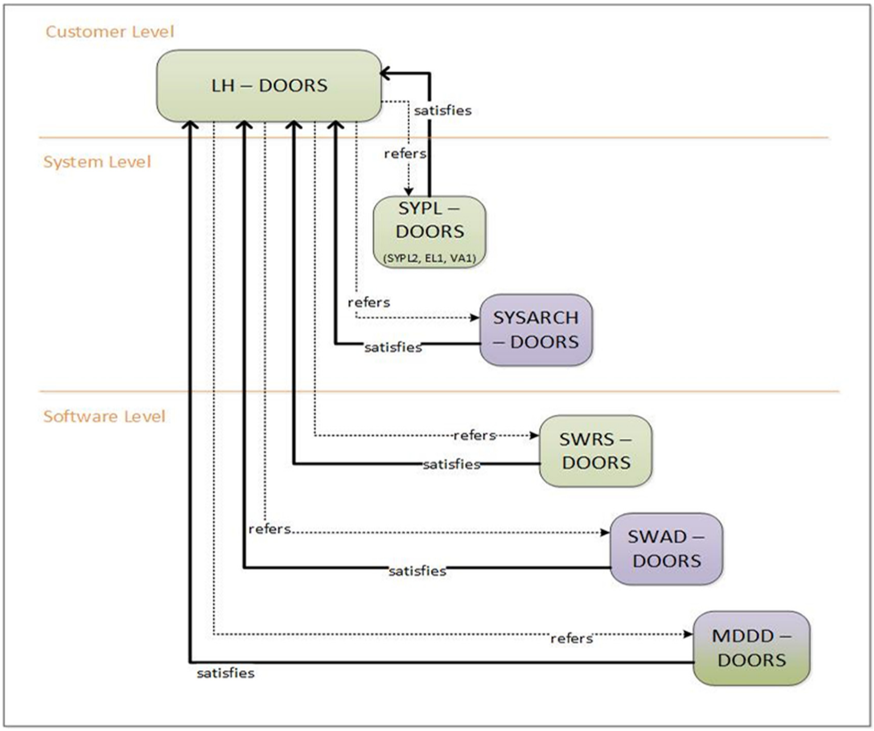


Fig. 3. Implementation of direct Link from lower level to customer level transfer of customer requirements during the elicitation process directly from the stakeholder requirements to other levels in the development process bypassing the level of system requirements ("SYPL2"). However, this bypass is not allowed for customer requirements to which an ASIL is assigned. (LH – Lasten Heft; SYPL ~ system requirement; SWRS ~ software requirement; MDDD ~ Software Unit)

Within the elicitation process, the requirements management tool attribute "Engineering Process Group" is used to allocate a requirement directly to a certain level (e.g. software requirements level). Traceability is kept by linking the stakeholder requirement to the respective (already existing) requirement at the corresponding level. If no requirement can be found at the corresponding level, a new object is created and linked. Afterwards the object is filled by the requirement engineer with requirement content (object text, attributes) and released after a review with the module owner.

To keep the traceability to system level in the case of requirements directly linked to customer level, the requirement engineer has to create an additional link or set of links from the level where the requirements were defined to the next superior levels in the Link Model, until it reaches the system level.

A part of the traceability is the linking requirements to a certain software release. Having this kind of traceability is a highly important and valuable topic in order to plan and to manage a software project.

Traceability to software release is done using a set of attributes defined in requirement management tool which have a direct correspondence in the Feature Rollout Plan (FRP):

- attributes High-Level Features (HLF) and (Low-Level Features) in requirement management tool corresponding to the same HLF and LLF used in the FRP and used in the release notes.
- attribute “Feature Cluster” in requirement management tool corresponding to the maturity grade (MG) model used in the FRP and used in the release notes.

With this system each single requirement is consistently linked to the planning of a certain software release and to the release note of that software release.

3.5 Improvement Evaluation

The following information is taken from second assessment report. The project has successfully implemented an improvement program and closed in 1.25 years 126 improvement tickets. The closure of each ticket is tracked and requires an evidence/proof that the problem is solved. An improvement rating is done by Quality team and linked all 126 tickets to BPs and GPs of A-SPICE and the closure of tickets led step by step to a better rating of BPs and GPs of A-SPICE over time. The major MAN.3 improvements are that the project scope is now defined and considered in all plans and the Agile method supports the release planning.

Agile process now includes efforts and tasks in DOD for requirements, SW architecture and testing. The process SUP.8 improvements are the review baselines, weekly team baselines, release baselines. Baseline plan or table is shared via dashboard that has been created for weekly tracking and coverage reporting of requirements. The Link model to have consistency criteria and baseline audits have been performed 11 times with 286 individually items checked every time.

4 Conclusions

The purpose of our research paper is the documentation of our case study to provide a clear, comprehensive and defensible argument for software process improvement activities. If the requirements elicitation cannot be given a final date due to continuous improvement of system features for a series production project, then an unreasonable project risk evolves. An Agile process approach can support the continuous improvement of system features with a systematic requirements elicitation activity. Supported by our evidence A-SPICE process creation, maintenance or change (SPI) helps to mitigate these unreasonable risks of series project development if applied in the culture context of the company and focused on the needs of the project management.

The main takeaway of our research paper is for our future projects to focus onto the elicitation of requirements in an Agile manner and provide a workflow concept of the requirement linkage. An improvement in planning should account also to increase, advance, and enhance the understanding of deadlines for all project stakeholders.

Thanks to excellence in engineering and customer satisfaction as encompass MARELLI Automotive Lighting has proved that it's possible to unite A-SPICE standards with Agile methods.

References

1. Layton, M.C., Ostermiller, S.J.: Quote was readjusted. Agile Project Management for Dummies, p. 45 (2017). (ISBN 978-1-119-40569-6)
2. Pries-Heje, J., Johnson, J.: Quote was readjusted. Manifesto Software Process Improvement (2020). https://2019.eurospi.net/images/eurospi/spi_manifesto.pdf
3. Pohl, K., Rupp, C.: Basiswissen Requirements Engineering, p. 1 (2015). ISBN 978-3-86490-283-3
4. Sami, M.: Quote was readjusted. The Software Process Improvement (SPI) – Reward or Risk (2018). https://melsatar.blog/2018/06/26/the-software-process-improvement-spi-reward-or-risk/
5. VDA: Quote was readjusted. Joint Quality Management in the Supply Chain Automotive A-SPICE Guidelines, pp. 58, 119 (2017)
6. Füermann, T., Dammasch, C.: Quote was readjusted. Prozess Management, p. 5 (2008). (ISBN 978-3-446-41571-3)
7. Scrum.org: Quote was readjusted (2017). https://www.scrum.org/resources/scrum-guide
8. Komiyama, T., NEC Corporation: Improvement of Agile Software Development Process Based on Automotive A-SPICE: A Case Study (2019). (WS6_Komiyama_18.09.2019_V1.1)

Building an Environment for Agility: A Releasable Inventory Tool

Marcos Pacheco[1], Antònia Mas[2], and Antoni-Lluís Mesquida[2](✉)

[1] Agile Coach & Trainer, Mallorca, Spain
marcos@marcos-pacheco.com
[2] Department of Mathematics and Computer Science, University of the Balearic Islands, Cra. de Valldemossa, km 7.5, 07122 Palma (Mallorca), Spain
{antonia.mas,antoni.mesquida}@uib.es

Abstract. In order to know the real status of all the applications and systems in a company, an inventory of all the elements that could be released can be very useful. This paper presents a *Releasable Inventory Tool*, which was developed to support Agility. It facilitates the creation of an environment aligned to Lean and Agile principles. The releasable inventory enables strategic, tactical, and operational decisions by providing details on current releasable elements, which are constantly monitored and evaluated to support continual improvement, alignment, and value creation.

Keywords: Continuous delivery · Software tool · Releasable elements · Lean principles · Agile manifesto

1 Introduction

There has been a big evolution in the industry of software development during the last 10 years. We are in a moment where companies cannot afford being behind their competitors in regards to the technical practices and approaches to deliver working software into productions hundreds of times per day if needed. Many organizations are taking "Continuous Improvement" as their mantra, and mainly in IT organizations this is extremely needed because their level of efficiency will determine how competitive they are within their market.

The 2019 Accelerate State of DevOps Report [1] represents six years of work and data from over 31,000 professionals worldwide. It is the largest and longest-running research of its kind, providing an independent view into the practices and capabilities that drive high performance. The results let us understand the practices that lead to excellence in technology delivery and powerful business outcomes. In [2], the authors indicate that high performing means delivering more value more quickly. Or in other words, the ability to ship higher-quality changes more frequently.

Continuous Integration (CI), *Continuous DElivery* (CDE) and *Continuous Deployment* (CD), called continuous practices [3], are some of the practices aimed at helping organizations to accelerate the development and delivery of software features without compromising quality [4], by reducing the amount of time needed to release software to customers.

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 587–597, 2020.
https://doi.org/10.1007/978-3-030-56441-4_44

CI is a software development practice where members of a team integrate their work frequently, usually with each person integrating at least daily [5]. CI includes automated building and unit and integration testing of each version of the software system.

CDE is a software engineering approach in which teams keep producing valuable software in short cycles and ensure that the software can be reliably released at any time [6]. CI is usually a part of the CDE practice.

CD is a software engineering process that focuses on the rapid delivery of software changes to end-users. In this process, incremental software changes are automatically tested and frequently deployed to production environments [7]. Software companies who are using CD have reported several benefits of this process, such as improved customer satisfaction, improved software quality, and savings in development effort [8].

For a delivery team (a delivery team is a group of professionals with all the needed skills and expertise that is capable to own a product end to end. Normally the different roles that compose a delivery team are: Software Engineers, Quality Assurance Engineers, and a Product Manager) it is important to be aware of all the systems, platforms, applications or programs (which from now onwards we will call *releasable elements*), that the team is supposed to maintain and improve. There are different actions that a company could perform to facilitate the work of a delivery team to bring value to the customer as soon as possible. One of these possibilities is to increase the knowledge about all the releasable elements by centralizing the information about those elements in an automated system. In many organizations, this data is not available or does not exist for old products and systems, the known monoliths.

This paper presents an information system called *Releasable Inventory Tool* that gathers a wide range of aspects regarding the releasable elements of a software development company in the tourism sector. The name of the organization is kept private. Many of its teams were very used to work on monoliths and old technologies. Teams were focused a lot on maintenance and very few on new developments. Those teams started to work on greenfield projects, but the need for maintenance on those monoliths was having a big impact on the greenfield projects' progress.

The Releasable Inventory Tool and the procedures related presented in this paper are part of an action carried out to promote the culture of Continuous Improvement of the company. They are aligned with the application of the principles of the Agile Manifesto [9, 10] and the Lean principles [11, 12]. Moreover, the development of this tool was aligned with the principle "Base improvement on experience and measurements" of the SPI Manifesto [13].

The *Releasable Inventory Tool* represents the first step towards excellence. To become an "elite performer" a company first should make visible all the systems, platforms, applications or programs it uses. Then, it should analyse their current state of health and, finally, perform the needed changes to improve these releasable elements.

The knowledge provided by this information system allows software developers to perform releases in the most productive possible way. Moreover, this *Releasable Inventory Tool* is not only relevant for the people that are currently working on the company. By analysing the power of the tool from the perspective of incorporating new employees, we can state that the procedures related to the integration, delivery, and deployment, represent a standardization that facilitates the integration of new members much faster.

The rest of the paper is organized as follows. Section 2 describes the *Releasable Inventory Tool* in detail. Section 3 discusses how this tool can help a company to fulfil the principles of the Agile Manifesto and the Lean Principles. Finally, Sect. 4 concludes the paper.

2 The *Releasable Inventory Tool*

The *Releasable Inventory Tool* is composed of a set of attributes that helps to know in-depth each releasable element of the company. The selection of attributes in this inventory is based on the experience of the first author, working in organizations in the middle of a technology transformation. Moreover, some of the ideas from Skelton and Pais [14] on team topologies were also considered for the attribute selection.

The *Releasable Inventory Tool* is the first step towards raising awareness of the state of the software that is already built. From this knowledge, it is possible to act quickly and in the right direction when some request comes from a customer and this request forces to generate a new delivery. Having such an inventory will be helpful in order to use the experience to not repeat the same mistakes in future projects.

The releasable elements that the *Releasable Inventory Tool* can contain in each record could be of different types:

1. An API that retrieves and updates information about users.
2. A third party system integrated into the company, such as a system for managing human resources.
3. An application, such as the actual web application of the organization.

The next sub-sections describe the fifteen attributes of each record for a releasable element of the *Releasable Inventory Tool*.

2.1 Business Area

Nowadays, in agile environments, delivery teams are often much better aligned to commercial teams, because this enables the delivery team to have clear conversations with the known “business”. This attribute represents the name of the business area the delivery team is working with.

2.2 Application Name

It can seem obvious, but sometimes different members use acronyms or they refer to the same releasable element by different names. This attribute represents the API name, service name or program name. For instance, if using GitHub, GitLab or similar, this attribute contains the repository name.

2.3 Criticality

Not all applications have the same criticality. Having a sales service down is not the same as having an organization's internal tool down. The impact on the business could be very different. In this case, the criticality will be relatively compared between all the applications that belong to the same business unit. For this attribute, one of the following three options should be selected:

- *Level 1*: Critical to keeping the business running. For instance, a part of the booking service being down.
- *Level 2*: Unknown, side effects not known, uncertainty on how it could affect the business.
- *Level 3*: Not important, the business can continue operating. For instance, some internal tooling.

2.4 Is Part of a Bigger Solution?

This attribute shows if the releasable element belongs to a bigger codebase. This would mean that when a change is released, it is released with other parts of the codebase that have not changed. For instance, if an organization has a big booking API for hotels and flights, and both belong to the same codebase, when there is a change needed on the hotels API, the flights API will be also released with no need. From the release point of view this could be problematic.

This attribute contains valuable information for making decisions on decoupling services later on.

2.5 Release and Rollback Process

Releases are critical because they mean putting value in the customer's hands. However, as the more manual a release is the highest chance to human mistake there is. This attribute shows how a release or rollback is performed. One of the following four options should be selected:

- *CI/CD*: No action will be necessary. Everything is triggered by the source control after each commit.
- *Automated*: It can be easily triggered by a platform (such as Jenkins) with a minimum of actions required by a user. There is a very low risk of human mistakes.
- *Semi-automated*: A set of scripts or jobs need to be triggered by a person to perform the release or rollback. There is a mid-level risk of human mistakes.
- *Manual*: It can be performed by manual actions of a person. No scripts at all or very few scripts need to be triggered by a person to perform the release or rollback. There is a very high risk of human mistakes.

This attribute provides valuable information by offering a view of how safe and easy are the releases within the organization.

2.6 Release and Rollback Documentation

This attribute will contain a link to the wiki or any other resource where the release process is documented. In case it is not documented, it will be left blank.

2.7 Infrastructure

This attribute will contain the infrastructure information relative to where the releasable element is running. More concretely, it contains information about:

- *Test environment*: If the element is running in a test environment, this field indicates where.
- *Production environment*: Refers to the different production environments where the element is running.

Finally, this attribute shows all the teams that can administrate these environments.

2.8 Git Repositories

This attribute will contain all links to the source version repositories.

2.9 Local Execution

This attribute will contain a link to the wiki or any other resource where the process to get the application running in a local machine is documented. In case there is no such documentation, it will be left blank.

2.10 Problem Identification Place

In case of a problem, incident within or outside the company working hours, this attribute contains the link of the relevant dashboards or places to look at for logs and an explanation of how to read them. This attribute also indicates if, in order to see the logs, it is needed to access directly the server, by FTP or SSH.

This information is valuable because it shows how easy it is to access that information, especially when every second matters.

2.11 Technology Stack

This attribute shows the current technology stack that is built and the different versions that they are in use.

This information is valuable because it shows if teams are supporting different versions within the same technology or if there are applications running on no longer supported versions.

2.12 Teams Working on It

When teams work with different parts of the same application dependencies might happen, due to being the same codebase. In this case, this attribute specifies which teams are likely to perform changes on the releasable element.

This information will be valuable to identify those releasable elements shared among more than one team and to analyse if different parts could be decoupled. In case no other team will need to change it, it will be left blank.

2.13 Specialists

Sometimes there are elements where the people who started them are no longer part of the company, or where most of the information is known just for a couple of persons. This fact can create a lot of dependencies and silos.

This attribute contains who are the people that know the releasable element best, with a level of knowledge that could teach others how to use it, including domain knowledge.

2.14 Maturity of the Element Coverage Test

The amount of test coverage can give hints of how safe can be the changes and releases of the element. This attribute indicates the level of coverage. One of the following three options should be selected:

- Over 80% coverage.
- Below 50% coverage.
- No test coverage at all.

This information can be valuable because it gives a view of how safe it would be to perform refactoring or changes in the codebase.

2.15 Database Logic

Old applications many times contain a lot of logic within the databases. In this case, it is important to know how much the application relies on Stored Procedures in the database. One of the two following options should be selected:

- *Just Queries*: No stored procedures at all. Just selects, inserts, deletes (basic queries).
- *Stored Procedures*: No matter how many, but the system relies on Stored Procedures to function.

This information can be valuable because it can give a view of how hard it would be to perform refactoring or changes in the codebase.

2.16 Examples of Releasable Elements

Table 1 presents the fifteen attributes of each record of the *Releasable Inventory Tool* and shows two examples for two releasable elements.

3 Fostering Agility Through the *Releasable Inventory Tool*

The *Releasable Inventory Tool* can help a company to fulfil the principles of the Agile Manifesto [11] and the Lean Principles (Value, Value Stream, Flow, Pull and Perfection) [12].

Principle 9 of the Agile Manifesto states that "Continuous attention to technical excellence and good design enhances agility". To follow this principle, any member of a software development team and, more specifically, any member of a delivery team should always make sure that all the contributions to any releasable element are the most successful concerning the organization and to the way the company designs the architecture, to ensure that developments meet technical expectations.

Following the principle 11 "The best architectures, requirements, and designs emerge from self-organizing teams", the *Releasable Inventory Tool* has been created to provide the teams with the greatest possible autonomy [15, 16]. By identifying and making visible any dependency that a team has about elements, products, services, equipment or systems when is needed, it is possible to establish a plan to eradicate or minimize these dependencies and to act to achieve the objectives.

The *Releasable Inventory Tool* is based on "Delivering Value Early". There is a direct relationship between having applications that are easy to change, well maintained and with a safe release pipeline with the capacity of delivering working software in production often and early. Following this line, a way to bring value to customers is aligned to the first Agile principle "Our highest priority is to satisfy the customer through early and continuous delivery of valuable software" and to the third Agile principle, that states "delivering working software frequently, from a couple of weeks to a couple of months, with a preference to the shorter timescale". By managing the company knowledge with the *Releasable Inventory Tool*, it will be possible to act the faster way to solve a problem in a product or a system, to provide a service or to widen the functionality of an existing element, and therefore to offer value to the customers as soon as possible.

The *Releasable Inventory Tool* can be seen as a way to help to map the value stream of each releasable element of the company, not from the customer's point of view, but from a way of reducing unnecessary processes or steps and eliminating the generation of wastes [17, 18]. Using the knowledge that this tool provides, customers can obtain what they want and, at the same time, the cost of producing that element is reduced.

The way of creating Flow, the third Lean principle, is according to principle 10 of the Agile manifesto: "Simplicity - the art of maximizing the amount of work not done - is essential". Removing the wastes from the value stream and ensuring that customers receive products and services without delays is crucial. The developed tool can help the team to set the work rules for making the work dynamic, clear and visible to all.

Table 1. Examples of records for two releasable elements in the *Releasable Inventory Tool*.

Attribute	Releasable element 1	Releasable element 2
A1 Business area	Finance	Loyalty program
A2 Application name	Bookings-details-api	Users-back-office
A3 Criticality	*Level 1* This API is critical to be able to generate invoices	*Level 3* Not affecting sales, but a user in the loyalty program could NOT claim for the accumulated points while this application is down
A4 Is part of a bigger solution?	No, it is its own solution and can be released independently	Yes, the loyalty program is part of the whole users-back-office codebase
A5 Release and rollback process	*CI/CD* Every time a commit is performed it gets releases automatically	*Manual* The **.dll* files need to be generated manually and then uploaded by an FTP
A6 Release and rollback documentation	–	company.github.com/users-back-office/documentaion/how-to-release.txt
A7 Infrastructure	*Test environment*: It's a test Kubernetes cluster called appbooking-service-test running in Azure *Production environment*: It's a Kubernetes cluster called appbooking-service running in Azure This is administered by the actual delivery team	*Test environment*: Server Name APPT01 *Production environment*: Server Name PAPPT01 This is administered by an operations team
A8 Git repositories	company.github.combookings-details-api	company.github.com/users-back-office
A9 Local execution	–	company.github.com/users-back-office/documentaion/how-to-run-in-local
A 10 Problem identification place	https://sgkibana.company.com/goto/987987 There is no explanation on how to read them	No dashboards. To see the logs, ssh the server and navigate to/etc./logs
A11 Technology stack	PHP Symphony 4 framework with PHP 7	The backend code is .NET 4.7 and Angular for frontend
A12 Teams working on it	–	The Profiling team from time to time needs to change code
A13 Specialists	Antonia Mas, Toni Mesquida	Marcos Pacheco

(*continued*)

Table 1. (*continued*)

Attribute	Releasable element 1	Releasable element 2
A14 Maturity of the element coverage test	Over 80% coverage	No test coverage at all
A15 Database logic	Stored procedures	Just queries

By identifying the specialists and the teams working on each element [19], and by fostering a pull and non-push work management environment, a delivery team could establish and limit the Work In Progress (WIP) from the moment the team starts working on some element until it is in production. These practices are aligned to the fourth Lean principle (Pull), with the main objective of creating an efficient and effective company.

By using the *Releasable Inventory Tool*, an organization will be able to bring visibility on the state of the different applications. For instance, how easy is to perform changes in the codebase, how safe is the release pipeline, or if there are knowledge silos. Moreover, having such an inventory will set the stage for improvement actions that will help the organization, as mentioned before, to "Deliver Value Early". Promoting a continuous improvement culture is the way of being aligned to the fifth Lean principle focused on pursuing Perfection.

4 Conclusion

The purpose of the *Releasable Inventory Tool* presented in this paper is to provide a single source of consistent information on all releasable elements available in the organization and to ensure that it is available to the target audience. This tool ensures that element descriptions are expressed clearly for the delivery teams to support continuous deployment. It provides a view on the scope of what releasable elements are available, and on what terms. The inventory management is supported by the delivery team roles responsible for managing, editing, and keeping up to date the list of available elements as they are introduced, changed, or retired.

The *Releasable Inventory Tool* enables the creation of value and is used by many different practices within the service value chain. Because of this, it needs to be flexible regarding what attributes it presents, based on its intended purpose. As such, organizations may wish the inventory to take other forms, such as a document. The final goal is to enable the current list of releasable elements to be communicated to the adequate audience. However, for the inventory to be perceived as useful by the organization it must do more than provide a static platform for publishing information about releasable elements.

Having such a kind of inventory facilitates the creation of an adequate environment to support Agility, being involved in all value chain activities:

- The inventory enables strategy and investment decisions by providing details on current releasable elements.
- Releasable element descriptions are constantly monitored and evaluated to support continual improvement, alignment, and value creation.
- The *Releasable Inventory Tool* enables strategic, tactical, and operational relationships with customers and users by enabling and potentially automating processes or practices.
- The inventory ensures both the utility and warranty aspects of releasable elements are considered and published, including the information security policy or service level agreements.

To sum up, the *Releasable Inventory Tool* provides context for how the releasable elements in the organization will be delivered and supported and it publishes expectations related to agreements and performance.

Acknowledgments. This work has been supported by the Spanish Ministry of Science and Technology with ERDF funds under grants TIN2016-76956-C3-3-R.

References

1. Forsgren, N., Smith, D., Humble, J., Frazelle, J.: 2019 Accelerate state of DevOps report. DORA & Google Cloud (2019)
2. Forsgren, N., Humble, J., Kim, G.: Accelerate: The Science of Lean Software and DevOps: Building and Scaling High Performing Technology Organizations. IT Revolution Press, Portland (2018)
3. Shahin, M., Ali Babar, M., Zhu, L.: Continuous integration, delivery and deployment: a systematic review on approaches, tools, challenges and practices. IEEE Access **5**, 3909–3943 (2017)
4. Humble, J., Farley, D.: Continuous Delivery: Reliable Software Releases Through Build, Test, and Deployment Automation. Addison-Wesley Professional, Boston (2010)
5. Eddy, B., et al.: A pilot study on introducing continuous integration and delivery into undergraduate software engineering courses. In: IEEE 30th Conference on Software Engineering Education and Training (CSEE&T) (2017)
6. Chen, L.: Continuous delivery huge benefits, but challenges too. IEEE Softw. **32**(2), 50–54 (2015)
7. Ashfaque, A., Rahman, U., Helms, E., Williams, L., Parnin, C.: Synthesizing continuous deployment practices used in software development. In: 2015 Agile Conference (2015)
8. Leppanen, M., et al.: The highways and country roads to continuous deployment. IEEE Softw. **32**, 64–72 (2015)
9. Fowler, M., Highsmith, J.: The agile manifesto. Softw. Dev. **9**(8), 28–35 (2001)
10. Highsmith, J., Cockburn, A.: Agile software development: The business of innovation. Computer **34**, 120–127 (2001)
11. Poppendieck, M., Poppendieck, T.: Lean Software Development: An Agile Toolkit. Addison-Wesley, Boston (2003)
12. Womack, J.P., Jones, D.T.: Lean Thinking. Simon & Schuster, New York City (1996)

13. Breske, E., Schweigert, T.: The SPI manifesto revisited. In: Stolfa, J., Stolfa, S., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2017. CCIS, vol. 748, pp. 401–410. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64218-5_33
14. Skelton, M., Pais, M.: Team Topologies: Organizing Business and Technology Teams for Fast Flow. IT Revolution Press, Portland (2019)
15. Pacheco, M., Mesquida, A.-L., Mas, A.: Being agile while coaching teams using their own data. In: Larrucea, X., Santamaria, I., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2018. CCIS, vol. 896, pp. 426–436. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-97925-0_36
16. Pacheco, M., Mesquida, A.-L., Mas, A.: Image based diagnosis for agile coaching. In: Walker, A., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 481–494. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_37
17. Senge, P.: The Fifth Discipline: The Art & Practice of the Learning Organization. Currency (2006)
18. Laloux, F.: Reinventing Organizations: A Guide to Creating Organizations Inspired by the Next Stage in Human Consciousness. Nelson Parker (2014)
19. Jovanović, M., Mas, A., Mesquida, A.L., Lalić, B.: Transition of organizational roles in Agile transformation process: a grounded theory approach. J. Syst. Softw. **133**, 174–194 (2017)

Employability Assessment of Agile Methods for Software Quality: *An Empirical Case Study*

Kamran Wadood[1(✉)], Muhammad Kashif Shahzad[1], and Muhammad Iqbal[2]

[1] Power Information Technology Company (PITC), Lahore, Pakistan
{kamran.wadood,Kashif.shahzad}@pitc.com.pk
[2] University of Education, Lahore, Pakistan
drmiqbal@ue.edu.pk

Abstract. The software industry is compelled to adapt with agile methodologies for their inherent software quality. This requires transition from traditional software engineering methods for an extended competitiveness which is successfully demonstrated in private sector; however, its potential benefits in the public sector are still unknown. In this paper, we assess employability of agile methods to improve software quality in large public sector software development companies. The choice of public sector is made due to its key role in sustainable e-Governance models to ensure uninterrupted services to the citizens. A survey of 28 questions, on software quality components (budget, time, faster debugging and correctness, changeability, easy testability and installability), is collected from the IT professionals of nine software development companies at WAPDA (water and power development authority), the largest public sector organization in Pakistan. The results conclude that these factors have more strong and positive relation to the software quality in agile development than traditional methods in public sector when the software product is to be developed, delivered, deployed and sustainably maintained in short span of time due to its inherent capability of easy debugging, testability and installability.

Keywords: Software quality assurance · Software engineering methodologies · Agile development

1 Introduction

The successful delivery of a software project has emerged as a key factor towards organizations' competitive advantage [6]. This becomes more important due to associated significant investment in IT initiatives, which are 5 to 10% of organizations' total revenues [7]. The success of a software project is not only defect free project but also completing it in time and within the estimated budget (IEEE Std. 610.12-1990). The software quality is one of the vital aspects to measure software project's success and is referred as an extent to which customers' expectations are satisfied [11].

Besides, all efforts to ensure software quality, projects still fail [7]. There are a number of software process models which are grouped as 'Prescriptive Process

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 598–614, 2020.
https://doi.org/10.1007/978-3-030-56441-4_45

Models', (e.g. Waterfall, V Models), 'Incremental Process Models' and 'Evolutionary Process Models' (e.g. Spiral Model) [34]. Traditionally, the software development focus has been placed on waterfall that is witnessed in the failure of software project [12]. In traditional methods, documentation is necessary before any development activity. This has proved fatal on account that the customers may want incorporation of either new or changes in existing features or requirements. Consequently, this might lead towards changes in the scope or behavior and pave its way towards other quality problems [4].

At times where the economy was not growing quickly and the client needs were not extremely muddled, the customary development strategies, for example, Waterfall, V-shaped and Spiral were adequate to take into account client requirements. However, as time advanced, there was a requirement for complex systems in a shorter period that led the emergence of agile software development approach (XP, Scrum, and ASD) to accommodate project changes, continuously [5]. These methods focus on high consumer satisfaction, lower defects rates, quicker development times and an answer for quickly evolving necessities. This keeps project cost min, despite of unavoidable changes [10]. Moreover, it has become one of the most commonly used software project management methodologies in recent years [12] and best serves the success of a software project. According to report published by Standish Group, success rate of waterfall and agile projects is 16% and 39%, respectively. Contrary, traditional methodologies, (e.g. Waterfall, Spiral, or CMM-based strategies) concentrate on the predictability, stability and higher assurance. In any case, both traditional and agile methodologies have situation based inadequacies that, if left unaddressed, could lead towards the failure of the project.

According to Juran and Frank [23], it is not only the customer who profit by emphasis on high quality; but also, the organizations that give proper worth to the quality also turn out to be more responsive and innovative, build their competitive edge and extraordinarily decrease their aggregate expense of development and time to market. If an organization fails to deliver high-quality software to its clients, their trust will be lost and it would keep organization away from getting new clients. Consequently, organizations would not have the capacity to survive in dynamic business settings. The software development organizations need to distribute quality software projects in a short time period. It necessitates guided experts, code builders to present powerful, successful and productive development model and in addition, demonstrated software project procedures and software quality methods [9].

In a modern economy, business sector requirement change rapidly, client needs evolve and new competitive dangers rise all of a sudden. To ensure the successful delivery of a software project, the software houses ought to be adequately proficient to incorporate changes in requirements even in a later phase of the development, yet ought to deliver working project in a small-time span. Therefore, majority of software houses are adapting agile development [17]. However, its benefits in large public software development organizations for the successful project delivery, while ensuring quality, are still not known. This is because of hierarchical bureaucratic management style, lack of right set of attitudes and the lack of willingness for change in these organizations [13]. Hence, in this paper we assess employability of agile method to improve software quality and its successful delivery in large public software development companies.

The choice is made due to its key role in sustainable e-Governance models to ensure uninterrupted services to the citizens. This study is carried out in the largest public organization with 9 software development companies across, Pakistan. The results conclude that the agile development methods' employability in these organizations will have stronger positive impact on improving software quality as well as successful delivery of the software projects.

In this paper, 'software house' refers to a private software development company and 'software development company' refers to a software development and support center in public sector organization. We developed a conceptual framework of factors which influence software project quality and evaluated these factors for agile development in software development companies. This paper is organized into five sections. Section 2 presents literature on factors influencing software quality and traditional and agile development methods. The Sect. 3 presents research methodology and Sect. 4 presents results. The findings are presented and discussed in Sect. 5.

2 Related Work

The literature review is carried out into three areas: software quality, models and development using traditional and agile methods. The objective is to identify key factors influencing software quality and develop a conceptual framework, and current state of traditional and agile software development methods usage.

2.1 Software Quality

In literature, term 'quality' is defined with multiple facets as customer, product and value. Customer based view defines quality as extent to which a product, procedure or a service come across the necessities or desires [27]. The product based quality alludes to measure unpriced properties contained in every unit of the priced property [14] whereas value based quality refers to the level of fabulousness at a reasonable cost and control in respect of variability at a reasonable price [16]. The quality of software is more customer centric and is defined as rationally free from defects or errors, supplied on time within budget, fulfills necessities and/or desires and that can be maintained [19]. The ISO 8402 and IEEE also defines software quality as the ability of a software to satisfy the stated and implied needs [20]. This goes one step ahead as software quality assurance that is defined as systematic pattern of all actions that provides confidence of software conformity to the established technical requirements [6, 18, 21]. The level of quality attained is dependent on the quality assessment and assurance processes [22]. The accuracy or consistency of these processes is dependent on the kind and purpose of the software project. At a basic or fundamental level, the processes go for the minimization of the deviations between the predetermined or encoded, planned performance or behavior of the product and the real developed behavior of the software product. At present, software quality is measured from specialized and customer point of views [24]. In former, software quality is dependent on details or attributes and is measured by developer for the sake of provision of guarantee related to specification or attributes in terms of mistakes in code by testing procedure [29]. Whereas, in later it is measured through customer experience in light of the fact, how sound software is in meeting customer desires. The customer disappointment does not only happen due to failure to meet specifications or coding errors; however, this can happen by postponements in delivering the project or exceeding the budget.

For the sake of measurement and understanding quality, distinctive models have been produced for relating quality attributes to one another [11]. The quality model of McCall sorts out quality of product as an order of aspects, metrics as well as criteria [11]. McCall's model has been censured for subjectivity in the measurement of metrics [25]. Besides this, International Organization for Standardization (ISO) moreover added a software quality standard, ISO 9126, to facilitate assessment of software quality [27]. It characterizes six quality attributes that high quality software must exhibit. The characteristics include correctness or accuracy, integrity, usability, reliability, portability and maintainability.

2.2 Software Quality Models

The software quality models are used in combinations with standards to clearly define the characteristics of a high-quality software project. Here we, briefly present prominent models the quality model by McCall, primarily revolves around developers and development processes. But, attempts have been made to cross over any barrier between customers and system developers by bringing about quality aspects, paying consideration to the needs of both i.e. customers and system developer [28]. Boehm's model has resemblances or likenesses to the quality model of McCall. His subjective methodology characterizes quality branches while focusing on three levels that finish with primitive attributes [30]. ISO 9126; Software Product Evaluation, has been released amongst ISO quality standards series. ISO additionally suggests quality features to assess the six areas of importance as mentioned in Fig. 1.

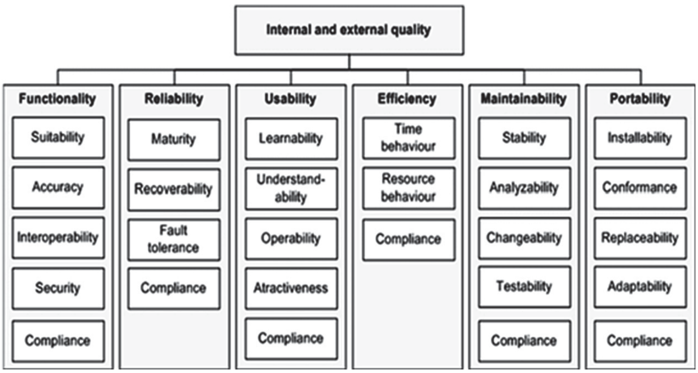


Fig. 1. Quality features or attributes of ISO\IEC 9126

A software quality model similar to McCall and Boehm is presented by [15] as FURPS (functionality, usability, reliability, performance, supportability). This is further divided into two main classes as Functional and Non-Functional classes. The CMMI (capability maturity model for integration), proposed by SEI (software engineering institute), assess the maturity levels of organization through process management, project management, engineering and support maturity levels. This presents five stages while concentrating on maturity levels while considering different software quality models presented above, it has been recognized that quality features described in different models have similarities with each other. It can be seen that all models focus around user requirements, testability, maintainability and portability. The user requirements comprise of not only on the functional requirements, but also the delivery of project within budget and time.

2.3 Traditional vs. Agile Software Development Models

The software process models can be grouped as ‘Prescriptive Process Models’, (Waterfall model, V model), ‘Incremental Process Models’ and ‘Evolutionary Process Models’ (Spiral model) [34]. Among these, the most commonly used traditional model is the Waterfall; however, agile has recently gain its popularity for inherent software quality as well as its emphasis on within budget and time delivery of software [10]. Therefore, this sub-section is limited to agile and waterfall methodologies.

Waterfall model from its initial concept was transformed into an iterative model through response of every stage which influences previous stage [36]. This is comparable to several approaches and includes a progression of definite (6 or 10) stages to run it sequentially with single or multiple substantial deliverables at the conclusion of all stages [36]. The 10-stages (commencement, conception, planning, requirements analysis, design, development, integration, test, deployment, operation, maintenance and disposition) and 6-stage (initiation, requirements analysis, design, development, integration, test and deployment) models are destined to large and small projects, respectively.

The agile method can be traced back to lean manufacturing and agile manufacturing in 1940s and 1990s respectively. This was developed in Agile Alliance (2001) and emerged initially as 12 standards related to software development with focus on customer and within budget and cost software, incorporation of changes throughout the development phases, inclination for short cycles and provision of working condition of software. This also encourages the maintainable development speed for customers, developers and sponsors. These standards can be taken into account as principal philosophies that ought to be inserted in acts of a system related to development of that claim itself to be agile. Besides these standards, the origin of the philosophies and above standards can be traced to iterative procedures [1]. Development of software through agile techniques might be viewed as method for reacting to the ambiguity in respect of development of software as opposed to a method for attaining certainty.

In present era, there are significant discourses, both in support and against agile approaches, in scientific community. Initial agile methods or approaches got criticism for the absence of scientific proof [2] and appropriateness for development of software where group of small numbers of individuals were delivering products having unpredictable requirements [38]. In recent times, organizational and corporate parts in respect of agility are getting extra consideration [31, 32, 35]. So, initial agile techniques as well as procedures are being developed as well old ones are being upgraded.

Recently, most empirical proof on adopting agile methods can be found in adjusting engineering beliefs and strategies of manageable, foreseeable and repeatable procedures with the agile software development, which once have been more self-managing, process adjustment and continuous modifications [26]. Adjusting two methodologies is recommended with a specific end goal to profit by their qualities or strengths and to make up for their shortcomings [8]. The quality with reference to design in development of software through agile can be organized through progressing design done in littler pieces rather than enormous open design of framework [10]. The studies demonstrate that by accepting diverse agile strategies and practices, individual teams practicing agile can achieve a method that hits objectives of CMMI level 2.

Finding of the Pilgrim's study has mirrored that the Agile is more efficient (quality of final software, quantity or number of bugs, time taken for development) than Waterfall because of its versatility and compliance to real world [33]. Waterfall model is right for the projects which are as of now unchanging or change in the design is not required. The agile development is more suitable at circumstances where the requirements evolve quickly [33]. The Waterfall has consecutive/sequential process in software development; whereas, in agile, focus is on 'adaptability'. The agile development method is ideally suitable to projects, which pay attention towards time to showcase, have high degree of variation in requirements and have single distinctive vital deliverable whereas waterfall development method is suitable to projects which are based on commitment in writing, places focus on analysis and have multiple deliverable [3]. Physically scattered teams frequently battle to deliver project well in time or on agreed time line. The use of agile development can reduce software production cycle and increase 37% delivery speed in respect of comprehensive software developed employing agile up to four to eleven months contrary to six to thirteen months for a similar project using traditional method [37].

The software development companies are moving towards agile development with objectives of 'time to showcase or market' and 'feasibility to incorporate changes even in the later phases of development' [38]. This is essential for their survival and literature review is carried out irrespective of type of Software Company which reveals that the public software development organizations have yet to benefit from the agile methods and they currently rely on the traditional approaches. The key role of these software development organizations in sustainable e-Governance model is critical.

2.4 Conceptual Framework

The literature review highlight quality components and variables related to software project quality which is measured as defect free, delivered in time, within budget and meets requirements and/or expectation [19]. Moreover, faster debugging and correctness, easy testability, changeability, easy installability and delivery on time and budget measures are also retained being organizations' characteristics for software project quality in the context of public sector software development companies.

A conceptual framework based on the software quality components as listed above is presented in Fig. 2. In this regard, the influence of these factors will be evaluated on the software quality within agile methodologies for large public sector software development companies.

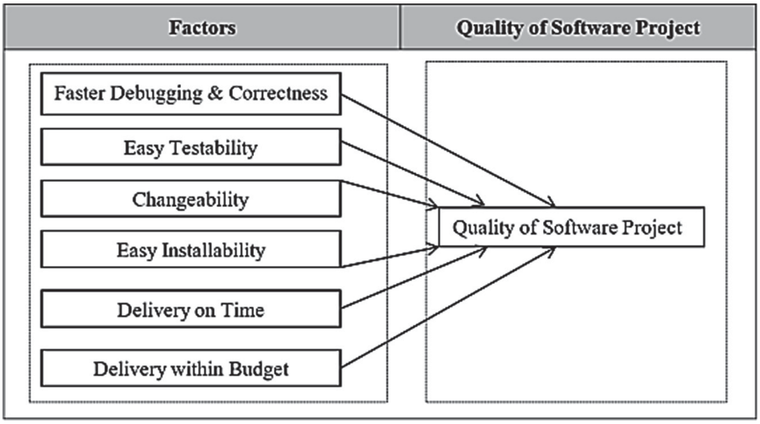


Fig. 2. Conceptual framework

3 Research Methodology

The current research primarily focuses to assert weather employing agile methods could improve software quality in the public-sector software development companies. A survey based quantitative approach is used to evaluate the presented conceptual framework (Fig. 2). For this purpose, power information technology company (PITC) as head quarter and IT centers in nine distribution companies (DISCOs) are selected, that constitutes the largest public sector software development company in Pakistan. As first step, survey of 8 questions was collected from nine DISCO computer centers to find software development organizations and IT professionals that either implemented agile methodologies or are in transition to do so. This ensured likelihood of data for subsequent analysis and conclusions. Following, 20 questions survey was collected from selected IT professionals.

3.1 Population and Sample

The population comprise of project managers, developers, quality assurance team leads and testers from nine DISCO computer centers. The question survey was sent to 216 IT professionals from total 850 employees, across Pakistan (Table 1). The reason for selecting employees with stated designations was that, these employees could give the required data or information for the purpose of this research.

The criteria used for sample is that respondents must not have less than one software project finished in the agile and waterfall methods. The analysis of the 8-question survey shows that 171 respondents qualified the initial criteria (Table 2).

Table 1. Population statistics

Name of the company	Number of employee as				Total
	Project managers	Developers	Quality assurance leads	Testers	
PITC	10	140	16	50	216

Table 2. Sample statistics

Name of the company	Number of employee as				Total
	Project managers	Developers	Quality assurance leads	Testers	
PITC	10	103	14	44	171

3.2 Data Collection (Survey and Interviews)

The data is collected through 24 questions survey on faster debugging/correctness, easy testability, changeability and installability, and agile vs. waterfall methods comparison. The survey is served to 161 IT professionals excluding Managers because with Managers, face to face interviews were held. The 139 valid responses were retained for data analysis in addition to the interview of 10 Managers. The surveys are designed on Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree).

A brain storming session was held with respondents prior to survey completion. However, to collect data for software delivery on time and within budget variables, face to face interviews were held with 10 managers. The responses were recorded and later transcripted prior to data collection. The Managers were asked to provide real time examples to support arguments. The reliability of instruments used for collected data was carried out with SPSS, prior to the conclusions (see Table 3). The value 0.940 in respect of Cronbach Alpha shows a highest internal consistency.

Table 3. KMO and Bartlett test

KMO and Bartlett test		
Kaiser-Meyer-Olkin measure of sampling adequacy		0.740
Bartlett's test of Sphericity	Approx. Chi-Square	1055.832
	df	276
	Sig.	0.000

The Kaiser-Meyer-Olkin measures sampling adequacy ranges from 0 to 1.0 (Table 3) demonstrates that the sum in respect of partial correlations is large relative to the sum of correlations.

The value MKO 0.74 shows compact correlation pattern which means that factor analysis should yield distinct and reliable factors. Moreover, Bartlett's measure tests the null hypothesis that original correlation matrix is an identity matrix. For this data, Bartlett's test is highly significant ($p < 0.05$), therefore, factor analysis is appropriate.

3.3 Hypotheses

To assess employability assessment of Agile development methodology and comparison with traditional methodologies to improve software project quality, following hypothesis are formulated (Table 4):

Table 4. Hypotheses list

No	Type	Description
1.	*H0*	No difference exists in faster debugging and correctness between software developed through agile and waterfall methods ($\mu AC - \mu WC = 0$)
	H1	Difference exists in faster debugging and correctness between software developed through agile and waterfall methods ($\mu AC - \mu WC \neq 0$)
2.	*H0*	No difference exists in easy testability between software developed through agile and waterfall methods ($\mu AT - \mu WT = 0$)
	H1	Difference exists in easy testability between software developed through agile and waterfall methods ($\mu AT - \mu WT \neq 0$)
3.	*H0*	No difference exists in changeability between software developed through agile and waterfall methods ($\mu ACh - \mu WCh = 0$)
	H1	Difference exists in changeability between software developed through agile and waterfall methods ($\mu ACh - \mu WCh \neq 0$)
4.	*H0*	No difference exists in installability between the software developed through agile and waterfall methods ($\mu AI - \mu WI \neq 0$)
	H1	Difference exists in easy installability between the software developed through agile and waterfall methods ($\mu AI - \mu WI \neq 0$)
5.	*H0*	No difference exists in time to deliver or complete between the software developed through agile and waterfall methods ($\mu AD - \mu WD = 0$)
	H1	Difference exists in time to deliver or complete between the software developed through agile and waterfall methods ($\mu AD - \mu WD \neq 0$)
6.	*H0*	No difference exists in estimated budget and actual budget between the software developed through agile and waterfall methods ($\mu AB - \mu WB = 0$)
	H1	Difference exists in estimated budget and actual budget between the software developed through agile and waterfall methods ($\mu AB - \mu WB \neq 0$)
7.	*H0*	No difference exists in software projects developed through agile and waterfall methods ($\mu AQ - \mu WQ = 0$)
	H1	Difference exists in software projects developed through agile and waterfall methods ($\mu AQ - \mu WQ \neq 0$).

The data in its totality is analyzed in the following steps to either support or reject the above formulated hypotheses:

- Keeping in mind the end goal to test derived hypothesis, used means, One Sample T – test on each of the above quality elements and
- *C*hecked hypotheses on Delivery on Time and within Budget, used Chi Square Test.

4 Data Analysis

The percentage of 139 responses is 95% which is presented in Table 5. The acronym S, R and A refers to sample, received and valid questionnaires.

Table 5. Summary of received data

Name of company	Developers			Quality assurance leads			Testers			Total
	S	R	A	S	R	A	S	R	A	
PITC	103	100	95	14	12	12	44	35	32	
Total S	103			14			44			161
Total R		100			12			35		147
Total A			95			12			32	139
R %age										91.30%
A %age (with reference to S)										86.34%
A %age (with reference to R)										94.56%

The summary of %age response rate of 139 respondents is presented in the Table 6.

Table 6. Percentage of response rate respondents category wise

Categories of Respondents		
%Age of Response Rate of Developers	=	97.09%
%Age of Response Rate of Quality Assurance Leads	=	85.71%
%Age of Response Rate of Testers	=	79.55%
%Age of Valid Response Rate of Developers	=	95.00%
%Age of Valid Response Rate of Quality Assurance Leads	=	100.00%
%Age of Valid Response Rate of Testers	=	91.43%

4.1 Faster Debugging and Correctness (H1)

For assessing this quality element, responses of questions (5–9) are subjected to one sample T-Test (developers, QA leads, testers) in respect of Faster Debugging and Correctness and associated hypothesis-1 (Table 7).

Table 7. T-test for Q. No. (5–9)

Sr. No.	Statement / Question	(a) Developers						(b) QA Leads						(c)Testers					
		t	df	Sig. (2-tailed)	Mean Diff.	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)	Mean Diff.	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)	Mean Diff.	95% Confidence Interval of the Difference Lower	Upper
1	User_ Expectation	29.6	94	.000	3.6	3.4	3.9	13.1	11	.000	4.1	3.4	4.8	20.9	31	.000	3.8	3.4	4.2
2	Requirement_ Capture	34.2	94	.000	3.7	3.5	4.0	17.1	11	.000	3.9	3.4	4.4	15.8	31	.000	3.5	3.0	3.9
3	System_Design	32.6	94	.000	3.7	3.5	4.0	15.7	11	.000	3.6	3.1	4.1	20.6	31	.000	3.8	3.4	4.2
4	System_ Implementation	32.3	94	.000	3.6	3.4	3.9	17.8	11	.000	4.1	3.6	4.6	17.5	31	.000	3.7	3.2	4.1
5	Faults_Free	38.7	94	.000	3.8	3.6	4.0	14.2	11	.000	4.1	3.5	4.7	19.4	31	.000	3.7	3.3	4.1

The significant values presented in Table 8 are less than 0.05, so null hypothesis cannot be accepted at 5% significance level. Which means difference exists in faster debugging and correctness between software developed through agile and waterfall methods (H1: $\mu AC - \mu WC \neq 0$).

Table 8. T-test for Q. No. (10–12, 15)

Sr. No.	Statement / Question	(a) Developers						(b) QA Leads						(c)Testers					
		t	df	Sig. (2-tailed)	Mean Diff.	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)	Mean Diff.	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)	Mean Diff.	95% Confidence Interval of the Difference Lower	Upper
1	Execution_Test _Script	29.7	94	.000	3.5	3.2	3.7	11.2	11	0.000	3.5	2.8	4.2	16.6	31	0.000	3.56	3.12	4.0
2	Coding_ Standards	29.1	94	.000	3.5	3.3	3.7	17.1	11	0.000	3.9	3.4	4.4	20.3	31	0.000	3.75	3.37	4.1
3	No_Complex_ Structure	33.1	94	.000	3.8	3.5	4.0	10.4	11	0.000	3.5	2.8	4.2	17.7	31	0.000	3.63	3.21	4.0
4	Low_ Interaction	36.0	94	.000	3.8	3.6	4.0	9.8	11	0.000	3.7	2.8	4.5	16.1	31	0.000	3.66	3.19	4.1

4.2 Easy Testability (H2)

For assessing this quality element, responses of questions (10–12, 15) are subjected to one sample T-Test (developers, QA leads, testers) in respect of easy testability and associated hypothesis-2. The significant values presented in Table 8 are less than 0.05, so null hypothesis cannot be accepted. It is concluded that there does exist difference in easy testability between software developed through agile and waterfall methodologies ($H1 = \mu AT - \mu WT \neq 0$).

4.3 Changeability (H3)

For assessing this quality element, responses of questions (13–14, 16–17) are subjected to one sample T-Test (developers, QA leads, testers) in respect of changeability and associated hypothesis-3. The significant values presented in Table 9 are lesser than 0.05, so null hypothesis cannot be accepted at 5% significance level. This concludes that the difference exists in changeability between software developed through agile and waterfall methods (H1: $\mu ACh - \mu WCh \neq 0$).

Table 9. T-test for Q. No. (13–14, 16–17)

Sr. No.	Statement / Question	(a) Developers						(b) QA Leads						(c)Testers					
		t	df	Sig. (2-tailed)	Mean Diff.	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)	Mean Diff.	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)	Mean Diff.	95% Confidence Interval of the Difference Lower	Upper
1	Easy_Modification	29.0	94	.000	3.5	3.2	3.7	7.1	11	0.000	3.1	2.1	4.0	18.6	31	0.000	3.6	3.2	4.0
2	Easy_Minor_Changes	32.1	94	.000	3.6	3.4	3.8	8.9	11	0.000	3.3	2.5	4.2	20.9	31	0.000	3.8	3.4	4.1
3	Low_Side_Effects	32.1	94	.000	3.7	3.5	3.9	10.7	11	0.000	3.6	2.8	4.3	19.8	31	0.000	3.8	3.4	4.2
4	No_Functional_Issues	29.8	94	.000	3.5	3.3	3.8	9.5	11	0.000	3.6	2.8	4.4	16.7	31	0.000	3.6	3.2	4.0

4.4 Easy Installability (H4)

For assessing this quality element, responses of questions (18–20) are subjected to one sample T-Test (developers, QA leads, testers) in respect of easy installability and associated hypothesis-4. The significant values presented in Table 10 are lesser than 0.05, so null hypothesis cannot be accepted at 5% significance level. This concludes that the difference exists in easy installability between software developed through agile and waterfall methods (H1: $\mu AI - \mu WI \neq 0$).

Table 10. T-test for Q. No. (18–20)

Sr. No.	Statement / Question	(a) Developers						(b) QA Leads						(c)Testers					
		t	df	Sig. (2-tailed)	Mean Diff.	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)	Mean Diff.	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)	Mean Diff.	95% Confidence Interval of the Difference Lower	Upper
1	Installation_Challenges	28.8	94	.000	3.5	3.3	3.7	8.9	11	0.000	3.3	2.5	4.2	16.5	31	0.000	3.5	3.0	3.9
2	Installation_Modify	32.0	94	.000	3.6	3.4	3.9	10.0	11	0.000	3.3	2.6	4.1	24.1	31	0.000	3.8	3.5	4.1
3	Compatibility	30.2	94	.000	3.6	3.4	3.9	10.3	11	0.000	3.7	2.9	4.5	17.7	31	0.000	3.7	3.3	4.1

4.5 Delivery on Time (H5)

For assessing this quality element ‘Delivery on Time’, interviews are conducted. This segment concisely analyses the responses against interviews and Chi Square test is used for the verification of the hypothesis-5. The data is collected from 10 project managers in respect of projects completed by them in last 5 years. The projects completed in time with agile and waterfall methods are 69.23% and 52.00%, respectively.

As per values presented in Tables 11 and 12, Chi Square value is 2.70, whereas critical value is 3.84. As the observed value is less than the critical value so, null hypothesis cannot be rejected at 5% significance level. Therefore, it is concluded that there is no proof or evidence to say that a difference in time delivery or completion of software project in respect of agile and waterfall methods exists (H0: $\mu AD - \mu WD = 0$).

Table 11. Observed and expected values in respect of interviews with project managers for quality factor delivery or completion on time

Observed and expected values				
Method		Agile	Waterfall	Total
Successful	Observed	27	26	53
	Expected	23.22	29.78	53
Unsuccessful	Observed	12	24	36
	Expected	15.78	20.22	36
Total		39	50	89

4.6 Delivery Within Budget (H6)

For assessing this quality element, the responses against interviews and Chi Square test is used for the verification of hypothesis-6. The analysis shows that projects completed in budget with agile and waterfall methods are 58.97% and 48.00%, respectively.

As per values presented in Tables 13 and 14, Chi Square value is 1.05, whereas critical value is 3.84. As the observed value is less than the critical value so, null hypothesis cannot be rejected at 5% significance level. Therefore, it is concluded that there is no proof or evidence to say that a difference in estimated spending plan or budget and actual spending in respect of agile and waterfall methods exists (H0: $\mu AB - \mu WB = 0$).

4.7 Comparison Between Agile and Traditional Methods (H7)

For the sake of comparison between agile and traditional software development methods (e.g. waterfall) 08 questions have been designed and added in the questionnaire (Table 15).

The significant values presented in above tables are lesser than 0.05, so null hypothesis cannot be accepted at 5% significance level. And, it can be concluded that there is an evidence to say that a difference exists in software projects developed through agile and waterfall software development methods (H1: $\mu AQ - \mu WQ \neq 0$).

Table 12. Chi Square test in respect of interviews with project managers for quality factor delivery or completion on time

	Chi Square test results				
	O	E	(O − E)	$(O - E)^2$	$(O - E)^2/E$
χ^2	27	23.22	3.78	14.2884	0.615349
	26	29.78	−3.78	14.2884	0.479799
	12	15.78	−3.78	14.2884	0.905475
	24	20.22	3.78	14.2884	0.706647
	Statistics from the test				2.70727
	Critical value of Chi Square				3.841

Table 13. Observed and expected values in respect of interviews with project managers for quality factor delivery or completion within budget

Observed and expected values				
Method		Agile	Waterfall	Total
Successful	Observed	23	24	47
	Expected	20.60	26.40	47
Unsuccessful	Observed	16	26	42
	Expected	18.40	23.60	42
Total		39	50	89

Table 14. Chi Square test in respect of interviews with project managers for quality factor delivery or completion within budget

	Chi Square test results				
	O	E	(O − E)	$(O - E)^2$	$(O - E)^2/E$
χ^2	23	20.60	2.40	5.7600	0.279612
	24	26.40	−2.40	5.7600	0.218182
	16	18.40	−2.40	5.7600	0.313043
	26	23.60	2.40	5.7600	0.244068
	Statistics from the test				1.054905
	Critical value of Chi Square				3.841

Table 15. T-test for questions on comparison of agile and traditional software development methods

Sr. No.	Statement / Question	(a) Developers						(b) QA Leads						(c)Testers					
		t	df	Sig. (2-tailed)	Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)	Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)	Mean	95% Confidence Interval of the Difference	
					Diff.	Lower	Upper				Diff.	Lower	Upper				Diff.	Lower	Upper
1	Flexibility	26.0	94	.000	3.3	3.0	3.6	8.6	11	0.000	3.1	2.3	3.9	17.5	30	0.000	3.8	3.3	4.2
2	Robustness	28.0	94	.000	3.4	3.2	3.7	10.8	11	0.000	3.3	2.7	4.0	19.0	30	0.000	3.5	3.1	3.9
3	Cost_ Effectiveness	28.1	94	.000	3.4	3.2	3.7	9.0	11	0.000	3.4	2.6	4.3	16.8	30	0.000	3.5	3.1	4.0
4	Reusability	26.2	94	.000	3.3	3.1	3.6	10.2	11	0.000	3.4	2.7	4.2	13.9	30	0.000	3.3	2.8	3.8
5	Reduced_ Risks in_ Success	22.7	94	.000	3.1	2.8	3.4	10.0	11	0.000	3.6	2.8	4.4	14.3	30	0.000	3.4	2.9	3.8
6	End_User_ Involvement	24.6	94	.000	3.2	2.9	3.5	9.8	11	0.000	3.2	2.5	3.9	16.5	30	0.000	3.2	2.8	3.6
7	Delivery_on_ Time	24.0	94	.000	3.1	2.8	3.3	8.1	11	0.000	3.0	2.2	3.8	12.9	30	0.000	3.2	2.7	3.7
8	Interaction_ between_Team _Members	25.2	94	.000	3.0	2.8	3.3	10.2	11	0.000	3.4	2.7	4.2	12.7	30	0.000	3.0	2.5	3.5

5 Recommendations and Future Directions

Through analysis, it is found that for quality components 'faster debugging and correctness', 'changeability', 'easy testability' and 'easy installability', a significant difference in agile and waterfall methods exist; however, agile is found to have superiority quality in public sector software development companies. Based on the results, it is recommended, for public software development organizations, that agile development techniques or methods:

- must be used whenever delivery of project is required in a short time span.
- offers easy debugging and high level of correctness for complex software
- offers inherent easy testability and inclusion of changes in requirements
- suitable for software where high consideration is paid to easy installability.

The research has not separated its findings as per magnitude, scale or size of the project, because of the trouble of information or data accumulation within the given time period. Moreover, it was carried out in Pakistani context, and gathered information from IT professional of nine DISCO computer centers. The generalization is possible only for the public software development organizations from developing countries.

The future research can lead to investigate other project quality components as cost effectiveness, time to market and most suitable strategy between agile and waterfall.

References

1. Abrahamsson, P.: Rethinking the concept of commitment in software process improvement. Scand. J. Inf. Syst. **13**(5), 69–98 (2001)
2. Abrahamsson, P., Salo, O., Ronkainen, J., Warsta, J.: Agile software development methods. VTT Technical Report (2002)
3. Alberto, G.: Waterfall vs. Agile: Can they be Friends (2010). http://agile.dzone.com/articles/combining-agile-waterfall. Accessed 1 Oct 2015
4. Ambler, S.: Quality in an agile world. Softw. Qual. Prof. **7**(4), 30–34 (2005)
5. Strode, D.E., Huff, S., Hope, B., Link, S.: Coordination in co-located agile software development projects. J. Syst. Softw. **85**, 1222–1238 (2012)
6. Owens, D.M., Khazanchi, D.: Software Quality Assurance, pp. 245–263 (2009)
7. Charette, R.N.: Why software fails [software failure]. IEEE Spectr. **42**(9), 42–49 (2005)
8. Boehm, B., Turner, R.: Using risk to balance agile and plan-driven methods. Computer **36** (6), 57–66 (2003)
9. Boehm, B.: A View of 20th and 21st Century Software Engineering. ACM, New York (2006). http://portal.acm.org/citation.cfm?id=1134288. Accessed 15 Aug 2015
10. Cockburn, A., Highsmith, J.: Agile Software Development: The People Factor (2001). http://www.Adaptivesd.com/Articles/IEEEArtical2Final.pdf. Accessed 25 Aug 2015
11. Kitchenham, B., Pfleeger, S.L.: Software quality: the elusive target. IEEE Softw. **13**(1), 12–21 (1996)
12. Kayes, I.: Agile testing: introducing PRAT as a metric of testing quality in scrum. ACM SIGSOFT Softw. Eng. Notes **36**(2), 1–5 (2011)
13. Mullins, L.J.: Management and Organisational Behaviour. Pearson Education (2007)
14. Leffler, K.B.: Ambiguous changes in product quality. Am. Econ. Rev. **7**(3), 129 (1982)
15. Grady, R.B.: Practical Software Metrics for Project Management and Process Improvement. Prentice Hall, New Jersey (1992)
16. Broh, R.A.: Managing Quality for Higher Profits. McGraw Hill (1982)
17. Highsmith, J.: History: The Agile Manifesto (2001). http://agilemanifesto.org/history.html. Accessed 15 July 2015
18. Hoyer, R.W.H., Hoyer, B.B.Y.: What is Quality - Quality Progress? The Global Voice of Quality, vol. 34, pp. 53–62 (n.d)
19. Ramand, G.: Software quality assurance. Qual. Assur. Softw. Test. **11**(2), 32–34 (2009)
20. Bevan, N.: Quality in use: meeting user needs for quality. J. Syst. Softw., 1–14 (1999). http://www.sciencedirect.com/science/article/pii/S0164121299000709. Accessed 12 Aug 2013
21. Runeson, P., Isacsson, P.: Software quality assurance - concepts and misconceptions. In: Proceedings of the 24th EUROMICRO Conference, pp. 853–859. IEEE Computer Society (1998)
22. Osterweil, L., et al.: Strategic directions in software quality. ACM Comput. Surv. **28**(4), 738–750 (1996)
23. Juran, J.M., Frank, G.: Juran's Quality Control Handbook. McGraw-Hill, Boston (1988)
24. Kokol, P., Žumer, V., Stiglic, B.: New evaluation framework for assessing the reliability of engineering software systems design paradigms. In: Brebbia, C.A., Ferrante, A.J. (eds.) Reliability and Robustness of Engineering Software II, pp. 173–184. Springer, Dordrecht (1991). https://doi.org/10.1007/978-94-011-3026-4_14
25. Côté, M.A., Suryn, W., Georgiadou, E.: Softw. Qual. J. **15**(4) (2007)
26. Lycett, M., Marcos, E., Storey, V.: Model-driven systems development: an introduction. Eur. J. Inf. Syst. **16**, 346–348 (2007). https://doi.org/10.1057/palgrave.ejis.3000684

27. Botella López, P., Burgués Illa, X., Carvallo Vega, J.P., Franch Gutiérrez, J., Quer, C.: Using quality models for assessing COTS selection. In: Proceedings of MPEC 2004 and ICSE 2004 (2004)
28. McCall, J.A., Richard, P.K., Walters, G.F.: Factors in software quality. Nat. Tech. Inf. Serv. **1**(3), 2–3 (1977)
29. Musa, J.I.A., Okumoto, K.: Software Reliability. McGraw-Hill, New York (1990)
30. Boehm, B.W.: Characteristics of Software Quality. North-Holland, Amsterdam (1978)
31. Baskerville, R., Pries-Heje, J.: Short cycle time systems development. Inf. Syst. J. **14**(2), 234–264 (2004)
32. Olsen, D.: Interaction in chaos. In: Proceedings of ACM 2nd International Conference on Intelligent User Interfaces, San Francisco, January 1998
33. Pilgrim, G.: Waterfall Model vs. Agile (2010). http://www.buzzle.com/articles/waterfall-model-vs-agile.html. Accessed 01 Sept 2015
34. Pressman, P.S.: Software Engineering: A Practitioners Approach, 7th edn. McGraw-Hill, Boston (2010)
35. Royce, W.W.: Managing the development of large software systems: concepts and techniques. In: Proceedings of the IEEE WESTCON, Los Angeles, pp. 1–9 (1970)
36. Smith, P.: Waterfall SDLC Methodology, 28 June 2011. http://skysigal.xactsolutions.com/Resources/SoftwareDevLifeCycle/WaterfallMethodSDLC/tabid/600/Default.aspx. Accessed 28 Sept 2015
37. Toronto and Boulder: The Agile Impact Report: Proven Performance Metrics from the Agile Enterprise (2008). http://www.pr-inside.com/new-study-shows-that-agile-teams-r737178.htm. Accessed 25 Oct 2015
38. Williams, L., Cockburn, A.: Agile software development: it's about feedback and change. IEEE Comput. **36**(6), 39–43 (2003)

Standards and Assessment Models

Towards a Model Based Process Assessment for Data Analytics: An Exploratory Case Study

Mert Onuralp Gökalp[1(✉)], Kerem Kayabay[1], Ebru Gökalp[2], Altan Koçyiğit[1], and P. Erhan Eren[1]

[1] Informatics Institute, Middle East Technical University, Ankara, Turkey
{gmert, kayabay, kocyigit, ereren}@metu.edu.tr
[2] Technology and Knowledge Management Department, Baskent University, Ankara, Turkey
ebrug@baskent.edu.tr

Abstract. The ability to leverage data analytics can enhance the decision-making process in organizations by generating valuable insights. However, there is a limited understanding of how organizations can adopt data analytics as part of their business processes due to a lack of comprehensive roadmap with a structural approach like a Process Capability Maturity Model (PCMM). In this study, the development of a PCMM based on the ISO/IEC 330xx standard for the data analytics domain is proposed to assist organizations in assessing their data analytics processes capability level and providing a roadmap for improving them continuously. Towards this goal, we conducted an exploratory case study for one data analytics process to evaluate the applicability and usability of the proposed approach.

Keywords: Data analytics · Data-driven organization · Process assessment model · Process improvement · SPI · ISO/IEC 330xx

1 Introduction

Data analytics attracted too much attention from both industry and academia by providing promising solutions to generate valuable insights to support strategic decision making. Accordingly, performing data analytics processes efficiently, effectively as well as in a robust way, has gained significant attention. A recent study of McKinsey [1] predicts that the potential impact of data analytics and artificial intelligence (AI) on the global economy will be around 13 million U.S. dollars by 2030. This study also expects that non-adopters of data analytics and AI will experience a 20% decrease in cash flow by 2030. Hence, organizations need to regard data analytics and data-driven culture as their core assets to exploit its promising potential. Thus, the organizations have increasingly started considering how they can improve their data analytic capabilities to attain a competitive advantage in their businesses [2]. However, there is a limited understanding of how organizations can assess and improve their data analytic capabilities due to a lack of comprehensive roadmap based on a structural approach.

Process capability maturity models (PCMMs), are utilized to evaluate and understand how to implement the value of relatively new technologies and capabilities in an

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 617–628, 2020.
https://doi.org/10.1007/978-3-030-56441-4_46

organizational context [3]. PCMMs describe fundamental patterns in the assessment of process capabilities and provide directions for improvement [4, 5]. PCMMs can be used descriptively to assess the current process capability level. Additionally, it also has prescriptive objectives to provide the steps that organizations should undertake to improve their current level of process capabilities. As a result of the observed benefits of PCMMs, it has been customized to different domains, including automotive [6], medical [7], and government [8]. Thus, PCMMs can be convenient to guide organizations for grasping the potential benefits of data analytics.

The main objective of this research is to explore the needs and requirements of a PCMM developed for the data analytics domain. Accordingly, we aim to propose a data analytics PCMM to assist organizations in assessing their current data analytics processes capabilities, deriving a gap analysis, and creating a comprehensive roadmap for improvement in a structured way.

Towards these goals, we intend to develop a data analytics PCMM based on ISO/IEC 330xx [9], a well-accepted set of standards. The proposed model comprises six data analytics processes, including; business understanding, data understanding, data preparation, model building, evaluation, and deployment. To evaluate the usefulness and applicability of the proposed model, the "business understanding" process is defined and assessed as an exploratory case study.

The paper is organized as follows. In Sect. 2, we give background information about PCMMs and review the existing studies in the scope of this study. Section 3 details the development of the proposed data analytics PCMM. In Sect. 4, we explain exploratory case study development and implementation steps. The results and findings of the conducted exploratory case study are discussed in Sect. 5. Finally, we conclude the paper and state directions for future research in Sect. 6.

2 Background and Related Work

In this section, a brief explanation of PCMMs, and a review of existing assessment models developed for the data analytics domain are given.

2.1 Process Capability Maturity Models (PCMM)

The PCMMs provide a common framework to assist organizations in assessing their capability and continuously improving their processes. They provide a place to start, a common language and a shared vision, a framework for prioritizing actions, and a way to define what improvement means for organizations [10]. The PCMMs provide necessary information for organizations to address their problems and challenges by providing a structured reference model to assess the existing process capabilities and organizational maturity as well as to provide a roadmap for improvement [11].

Capability Maturity Model Integration (CMMI) [12] and ISO/IEC 15504, which has been replaced with ISO/IEC 330xx series [9] are well-accepted PCMMs in the literature. These models can be used as a basis for process assessment and improvement, assuming that a higher level of process capability results in better process efficiency and product as well as service quality. The main observed benefits of these

PCMMs are expense savings, improved process quality, predictable and consistent process outputs, and increased employee productivity. However, these well-accepted PCMMs are mainly developed for the purpose of assessing software development processes. Thus, customizing them to different domains other than software development is the subject of growing interest in the literature. Accordingly, many initiatives based on these well-accepted PCMMs are proposed for various domains: Automotive SPICE [6], Medi-SPICE [7], and government [13].

2.2 Related Work

In the literature, some studies focus on proposing a maturity model for big data and data analytics domain.

The studies by Lukman et al. [14], Cosic et al.[15] and Raber et al. [16] propose a business intelligence maturity model. However, these studies do not specify any comprehensive detail about how to measure the data analytics capability or maturity of an organization objectively. Moreover, none of these models was developed based on an accepted framework or standard and validated across different organizations and industries.

There are also some studies in the grey literature related to data analytics maturity, including the studies presented by Halper et al. [17] and Davenport [18]. These studies were not developed based on a well-accepted framework or standard. Though they focus on assessing the current capability of the data analytics processes, they do not provide any structured roadmap for process improvement. These studies do not guarantee an unbiased academic view since they were proposed by a consulting company or a technology vendor. Moreover, these studies do not include any detail about the development and validation process, which hinders their adoption in practice.

There is a growing research interest in data analytics in recent years. However, as the related works show, there is a research gap due to the limited research in the use of PCMMs for the data analytics domain. Existing models mainly lack an objective assessment method and measurement attributes defined in a standardized way. Besides, none of them are compatible with any well-accepted framework, model, or standard such as CMMI, ISO/IEC 330xx, or ISO/IEC 15004. Toward this end, there is a need to define a well-structured, standardized, and consistent PCMM for data analytics. In this study, we aimed to fill this research gap by providing a PCMM for the data analytics domain based on a well-accepted PCMM standard, ISO/IEC 330xx.

3 Data Analytics Process Capability Maturity Model

The proposed data analytics PCMM includes two main dimensions; capability dimension and process dimension. On the one hand, the capability dimension, adapted from ISO/IEC 33002 [19], includes the definition of capability levels, Process Attributes (PA), Generic Practices (GP), and Generic Practice Indicators (GPI). On the other hand, the process dimension includes the data analytics process reference model, including process definitions of the data analytics processes.

3.1 Process Dimension

Data analytics can be defined as collecting, storing, transforming, and analyzing data statistically and quantitatively to build explanatory, predictive, and prescriptive models to drive business decisions and actions [18]. The data analytics aims to turn raw data into valuable insights. The Cross-Industry Standard Process for Data Mining (CRISP-DM) [20] framework provides a well-established process model for data analytics projects. The standard CRISP-DM data analytics lifecycle, which is depicted in Fig. 1 comprises six critical processes [20], including; business understanding, data understanding, data preparation, model building, evaluation, and deployment.

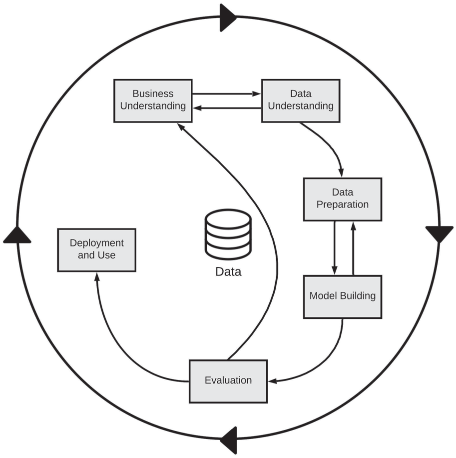


Fig. 1. Standard process lifecycle for data analytics projects

Business Understanding process identifies the business needs and requirements to achieve organizational strategy and vision from a data point of view [20]. In this process, organizations need to conduct a gap analysis to determine how data analytics can provide value in the context of their business strategy and what are the possible opportunities. It also includes formulating organizational bottlenecks and business opportunities into a data analytics problem. In this process, organizations need to conduct a gap analysis to define needs, problem scope, requirements, assumptions, and constraints. Moreover, a draft project plan should be prepared to determine the project needs and to estimate required IT, human resources, and budget throughout the project.

Data Understanding process explores data, data sources, and needs, and verifies data quality [20]. It represents the ability to understand the data with statistical and mathematical thinking. This process includes examining the structure and quality of the dataset, the distributions of individual variables, and the relationships between two or more variables. It also recognizes the strengths and shortcomings of the available data to determine whether the desired business problem can be solved or not. Thus, organizations need to define their data needs, data sources, and data quality requirements and metrics to solve a specified problem in the business understanding process.

Data Preparation process assesses how organizations deal with data cleansing, data wrangling, data quality, and structuring [20]. The success of model building highly depends on the merit and the quality of the data set. Thus, the raw data need to be cleaned from noisy, incomplete, inconsistent, redundant, and duplicate records to yield better results in model building. This process also includes deriving new attributes, transforming and scaling existing attributes to enrich data according to the characteristics of raw data, and with the knowledge of the business domain. Moreover, existing data analytics technologies and techniques may require data in different forms to perform computations and modeling. Thus, data structuring, manipulations, and conversion are also necessary to base practices in the data preparation process.

Model Building process investigates how useful knowledge can be extracted in a timely manner and to make data-driven decision-making capabilities pervasively [20]. This process measures how organizations embrace a diverse set of analytical techniques from descriptive analytics to prescriptive analytics to identify the patterns that form the basis of actionable insights. In this process, modeling requirements, objectives, as well as assumptions, need to be defined. A conceptual architecture needs to be designed by assessing available data analytics tools and techniques to find an optimal solution [21].

Evaluation process validates the success of developed data products or services rigorously [20]. It focuses on understanding how the developed data analytics solution is helpful to improve the business processes and achieve organizational strategy and vision. In this process, organizations need to evaluate the continuity and reliability of their data sources, design testing for intermediate and final products and services, determine key performance indicators and collect feedback from stakeholders to assess improvement opportunities.

Deployment and Use process integrates data analytics solutions provided in the model building process to business processes [20]. In the deployment and use process, organizations need to understand the technical architectures of the organization, such as the data ingestion and collection system, to ensure that the provided solutions in the model building phase are useful in practices. In this process, organizations need to define deployment requirements and policies and to prepare a deployment and use plan, allocate IT resources for making deployment environment ready, and to prepare a monitoring and a maintenance plan.

3.2 Capability Dimension

The process capability level assessment describes the organization's competency relative to a particular process. The measurement framework provided by ISO/IEC 33002 [19] is utilized. The proposed model, which is delineated in Fig. 2, has six capability levels, and each level has its own PAs which is defined as a measurable property of process capability. A capability level represents a well-defined set of PAs that provide a significant improvement in the capability of a process [22]. The defined PAs in each level addresses the specific needs of the capability levels, and they progress through the improvement of the capability of any process. Each process is required to be at least a "Largely Achieved (L.A)" for the corresponding PAs for each capability level and "Fully Achieved (F.A)" for any lower capability levels PAs.

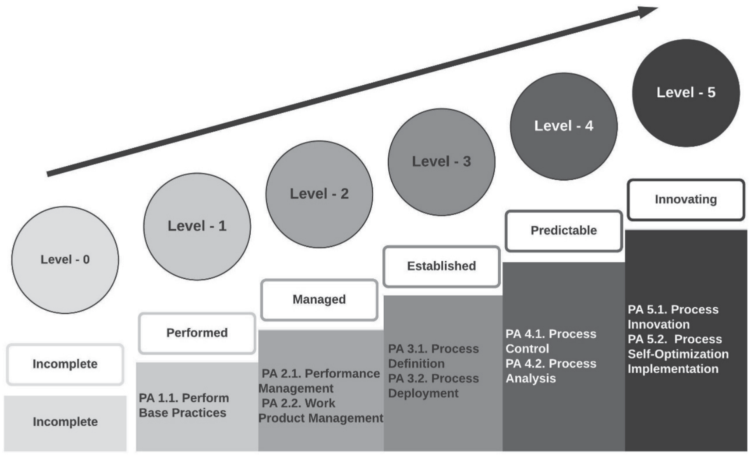


Fig. 2. Capability levels of data analytics processes (Adapted from ISO 3300xx)

Level 0 – Incomplete: In this capability level, the organization does not fully or largely perform base practices.

Level 1 – Performed: The organization largely or fully performs the base practices PA.1.1. Perform Process Attributes, but most of these processes are performed ad-hoc. In other words, there is not a consistent way of performing data analytics processes. At this capability level, the processes are unpredictable, poorly controlled, and reactive.

Level 2 - Managed: The base practices are fully performed. The organization starts to recognize the business value of data analytics and starts focusing on improving process performance by defining performance objectives for each process. Moreover, organizations are expected to identify, manage, and control work products of each process for performing processes consistently.

Level 3 - Established: In this capability level, organizations need to fully achieve the PAs of level 2 and also at least largely perform process definition and process deployment PAs to achieve level 3. In this level, organizations are expected to perform and maintain processes in a standardized way by defining and controlling each process.

Level 4 - Predictable: The organization begins managing its processes through the quantitative data that describes the performance, as well as variations in performing best practices, are reduced by controlling and analyzing processes. A controlled process is planned, performed, and monitored to achieve a quantitatively managed process. To this end, statistical and quantitative measures are collected at this level to control and monitor the process against the plan and to take appropriate corrective action when it is needed.

Level 5 - Innovating: In this capability level, the organization fully realizes defined data analytics PAs and starts self-learning from collected measures to improve the performance of data analytics life-cycle continuously. Moreover, the business model is evolving into an innovative structure with the gained insights from data analytics.

4 Exploratory Case Study

4.1 Exploratory Case Study Design

In this study, we conducted an exploratory case study to assess the applicability and usability of the proposed data analytics PCMM. The measure used in the case study is the capability level of the Business Understanding process. The primary sources of evidence of this exploratory case study are the assessment interviews, follow-up interviews that were conducted with process owners and stakeholders, and the information-gathering documents. The research questions for this exploratory case study are as follows;

- *RQ-1:* How suitable and applicable is the application of ISO/IEC 330xx based capability assessment for data analytics to be used with the purpose of identifying the current state of an organization data analytics process capability?
- *RQ-2:* How well it provides a roadmap for improving the capability level of the organizations' data analytics processes?
- *RQ-3:* What are the strengths and weaknesses of the proposed model for improvement?

In the exploratory case study, we performed a process capability assessment for the process of Business Understanding in a manufacturing organization that operates in the energy industry. The organization aims to improve its operational performance and decision-making capabilities of business units through data analysis. Semi-structured interviews were conducted with two process owners. All of the interview participants have more than five years of work experience. The participants were responsible for the development of data analytics processes and quality assurance. In the assessment process, we followed the ISO/IEC 330xx guidance on performing the assessment. Before conducting interviews, we detailed the proposed model, assessment plan, data collection, and validation methods.

4.2 Process Assessment

The process assessment plan is depicted in Fig. 3. The assessment process started with documentation of the assessment plan that includes defining the assessment team, interview schedule, delivery dates, and outputs. The assessment plan was shared with the organization and participants to be approved. Then, we conducted three-hours semi-structured interviews with participants to collect data, observations, and pieces of evidence to rate PAs. After analyzing the collected evidence, the capability level of the process was rated, and an assessment report, including the current capability level of the assessed process and a roadmap as well as suggestions to improve process capability to the next level, was prepared and shared with participants.

The PA of Level-1 is Process Performance, which is a measure of the extent to which the process purpose is achieved. For the capability assessment of Level-2 to Level-5, we utilized GPs defined in ISO/IEC 330xx, which is also an extended part of ISO/IEC 15504 Part – 5 [22].

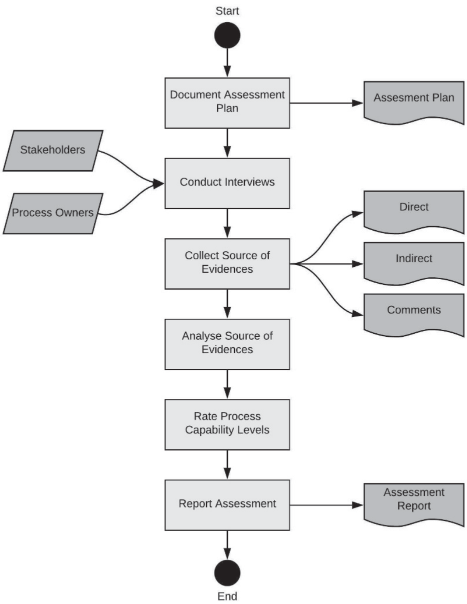


Fig. 3. Assessment process plan

In order to assess the PA of Level-1, we defined the process definition of Business Understanding process by utilizing the template provided by ISO/IEC 330xx standards. The template includes a description of processes in terms of the process name, process purpose, and process outcomes to define data analytics. It also includes performance indicators of the process by defining a set of base practices to provide a definition of the tasks and activities that need to accomplish the process purpose and fulfill the process outcomes. The developed process definition of Business Understanding is given in Table 1.

5 Findings and Discussion

According to the assessment results, the capability level of the Business Understanding process in the organization was determined as Level–2. The findings and results are detailed in Table 2.

In order to improve the capability level of the Business Understanding process to Level–3, the PAs of Performance, and Work Product Management should be improved to F.A. and the PAs of Process Definition and Deployment should be improved at least L.A. We provided some suggestions to improve the process capability level to Level–3 in the assessment report. These suggestions are as follows;

- The performance and success metrics of the process need to be clearly established.
- Quality criteria of the work products should be identified to enable reviewing the quality and content of the work products against defined requirements.

Table 1. Business understanding process definition

Process name	Business understanding
Process purpose	The purpose of Business Understanding process is to analyse the business needs and requirements from a data point of view and formulate data analytics projects to achieve the objectives
Process outcomes	As a result of successful implementation of this process: 1) A gap analysis is conducted and organizational issues are defined and formulated as a data analytics problem; 2) Objectives, scope and business success criteria are defined; 3) Requirements, limitations and assumptions are defined; 4) Project risks are analysed and identified; 5) An action plan for each defined risk is prepared; 6) A business glossary document is published; 7) A sketch project plan is prepared; 8) Cost and benefit analysis are conducted according to the draft project plan; 9) A document that includes all deliverables and steps in business understanding is published and shared with all related stakeholders
Base practices (BP)	**BP1: Conduct a Gap Analysis and Define Needs:** Conduct a gap analysis, define needs and formulate problems to reach business strategy [Outcome: 1] **BP2: Identify Objectives, Scope, and Business Success Criteria.** Define objectives, scope, and business success criteria for planning data products and services [Outcome: 2] **NOTE 1:** The statement of business objectives, scope and business success criteria is controlled using the practices of the Evaluation process **BP3: Define Requirements, Assumptions and Constraints**. Identify all organizational and industrial assumptions, constraints for the data-science project. Define sources of risks, areas of impact and events that might cause to fail the project [Outcome: 3] **BP4: Identify Risks and Prepare Action Plans.** Identify and analyse risks to the project. Define an action plan for each identified risk [Outcomes: 4, 5] **BP5: Develop a Business Glossary.** Define a terminology glossary relevant to industrial terms to enable common understanding and to reduce the risk due to inconsistent understanding of the business concepts [Outcome: 6] **BP6: Prepare a Draft Project Plan.** Establish and maintain a draft project plan to analyse the project needs better and to estimate required IT and Human resources and budget to be needed throughout the project [Outcome: 7] **NOTE 2:** This may include: • Estimating project cost and preparing a draft project time schedule, • Identify human and IT resource requirements, • The sketch of a work breakdown structure and assign responsibilities **BP7: Analyse Cost and Benefits.** Analyse cost and benefits of the project according to the draft project plan. Discuss the project benefits towards costs with all related stakeholders [Outcome: 8] **BP8: Publish a Business Understanding Document.** Prepare and publish a document that includes all steps, deliverables and decisions in business understanding processes [Outcome: 9]
Output work-products	Gap Analysis Results → [Outcome: 1] Requirement Specification → [Outcome: 2, 3] Risk Management Plan → [Outcomes: 4, 5] Business Glossary → [Outcome: 6] Work Breakdown Structure → [Outcome: 7] Cost and Benefit Analysis → [Outcome: 8] Business Understanding Document → [Outcome: ALL]

Table 2. Assessment results

Level	PAs	Evidence	Result
Level 1	Process performance	The process clearly achieved its purpose and outcomes by understanding the business needs and requirements from the data point of view	F.A.
Level 2	Performance management	The performance of the process is managed, but the performance and success metrics are not clearly defined	L.A.
	Work product management	The requirements of work products are identified; however, their quality metrics are not established. Moreover, the work products are not reviewed against the defined requirements	P.A.
Level 3	Process definition	A standard process definition is established; however, the interaction between other processes is not defined. Moreover, the required infrastructure and work environment need for the process are not defined	P.A.
	Process deployment	A process deployment guideline is established in the organization. Required infrastructure and work environment for performing the defined process are available and maintained. However, data required to understand the behavior, suitability, and effectiveness of the defined process are not identified and collected	P.A.

- A terminology glossary relevant to industrial terms needs to be documented, controlled, and maintained. The interaction between processes should be defined.
- The required infrastructure and work environment need for the process should be defined.
- Business needs need to be revised, managed, and controlled systematically.
- Data required to understand the behavior, suitability, and effectiveness of the process need to be clearly identified and collected in real-time.
- Resolving issues arising from work product reviews should be tracked systematically. Training for deploying the process should be performed.

We also conducted another interview with the same participants after presenting the assessment results. In this interview, open-ended structured questions, which are defined below, were utilized to validate the usefulness and applicability of the proposed approach.

- Are measuring the process capability and obtaining a guideline for improvement useful?
- Do you think that applying these suggestions will improve the process performance?
- Do you think that language and terminology in the questions are easy to understand?
- Is there any missing item in the guideline for the improvement list?
- Is there any information you want to add in the process definition?

The participants declared that provided suggestions and guidelines are applicable and useful to improve their current capability of the business understanding process, and they will apply these suggestions in the organization. They also indicated that the language and terminology used in the questions were easy to understand. Moreover, according to them, the process definition fully covers the business understanding process, and suggestions for the process improvement were not deficient.

6 Conclusion

Performing data analytics processes efficiently and effectively grasps the potential to improve operational performance, decision-making capabilities of business units as well as to attain a competitive advantage in their businesses. However, there is a limited understanding of how organizations can measure and improve their data analytic capabilities due to a lack of comprehensive roadmap with a structural approach like a PCMM. To this end, this study proposes a data analytics PCMM based on the well-accepted framework, ISO/IEC 330xx, to determine the current capabilities of data analytics processes and provide a comprehensive roadmap to improve capability levels in a structured way. It aims to provide a variety of values as indicated in the SPI Manifesto [23], including cost savings, more involved employees, improved and predictable quality, as well as productivity, generating a consistency of process capture and use.

The main contributions of this study are proposing a PCMM for data analytics domain based on ISO/IEC 330xx, and conducting an exploratory case study to discuss the applicability and usefulness of the proposed approach. According to the findings of the exploratory case study, the proposed model is applicable and usable to identify the current state of an organization's data analytics process capability and provide suggestions and roadmaps for improvement.

Despite its contributions, this study has some limitations that we need to focus on in future studies. The exploratory case study performed for one data analytics process, Business Understanding, to check the applicability and usefulness of the proposed approach. We need to develop the data analytics process reference model, including process definitions for all data analytics processes. After finalizing the development of the model, we plan to perform multiple case studies across different sectors and organization sizes to understand the applicability, usability, and generalizability of the proposed model.

References

1. Manyika, J., Chui, M., Joshi, R.: Modeling the global economic impact of AI. McKinsey (2018). https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-AI-frontier-modeling-the-impact-of-ai-on-the-world-economy. Accessed 25 Dec 2019
2. Gokalp, M.O., Kayabay, K., Akyol, M.A., et al.: Big data for Industry 4.0: a conceptual framework. In: 2016 International Conference on Computational Science and Computational Intelligence (CSCI), pp. 431–434. IEEE (2016)

3. Hüner, K.M., Ofner, M., Otto, B.: Towards a maturity model for corporate data quality management. In: Proceedings of the ACM Symposium on Applied Computing, pp. 231–238. ACM (2009)
4. Khan, A.A., Keung, J., Niazi, M., et al.: GSEPIM: a roadmap for software process assessment and improvement in the domain of global software development. J. Softw. Evol. Process **31**, e1988 (2019)
5. Barafort, B., Mesquida, A., Mas, A.: ISO 31000-based integrated risk management process assessment model for IT organizations. J. Softw. Evol. Process **31**, e1984 (2019)
6. Automotive SIG: Automotive SPICE process assessment model. Final Release, v4 4:46 (2010)
7. Mc Caffery, F., Dorling, A.: Medi SPICE development. J. Softw. Evol. Process **22**, 255–268 (2010)
8. Gökalp, E., Demirörs, O.: Model based process assessment for public financial and physical resource management processes. Comput. Stand. Interfaces **54**, 186–193 (2017)
9. ISO/IEC: ISO/IEC 33001:2015 information technology – process assessment – concepts and terminology (2015)
10. Korsaa, M., Johansen, J., Schweigert, T., et al.: The people aspects in modern process improvement management approaches. J. Softw. Evol. Process **25**, 381–391 (2013)
11. Varkoi, T., Mäkinen, T., Cameron, F., Nevalainen, R.: Validating effectiveness of safety requirements' compliance evaluation in process assessments. J. Softw. Evol. Process **32**, e2177 (2020)
12. Ahern, D.M., Clouse, A., Turner, R.: CMMI. SEI Ser. Softw. Eng. (2001)
13. Gökalp, E., Demirörs, O.: Government process capability model: an exploratory case study. In: Mitasiunas, A., Rout, T., O'Connor, R.V., Dorling, A. (eds.) SPICE 2014. CCIS, vol. 477, pp. 94–105. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-13036-1_9
14. Lukman, T., Hackney, R., Popovič, A., et al.: Business intelligence maturity: the economic transitional context within Slovenia. Inf. Syst. Manag. **28**, 211–222 (2011)
15. Cosic, R., Shanks, G., Maynard, S.: Towards a business analytics capability maturity model. In: 2012 Proceedings of the 23rd Australasian Conference on Information Systems (ACIS 2012), pp. 1–11. ACIS (2012)
16. Raber, D., Winter, R., Wortmann, F.: Using quantitative analyses to construct a capability maturity model for business intelligence. In: 2012 45th Hawaii International Conference on System Sciences, pp. 4219–4228. IEEE (2012)
17. Halper, B.F., Stodder, D.: A Guide to Achieving Big Data Analytics Maturity (TDWI Benchmark Guide) (2016). https://tdwi.org/whitepapers/2018/01/aa-all-ms-a-guide-to-achieving-big-data-analytics-maturity.aspx. Accessed 6 May 2019
18. Davenport, T.H., Harris, J.G.: Competing on Analytics: The New Science of Winning. Harvard Business Press (2007)
19. ISO/IEC: ISO/IEC 33002:2015 - information technology – process assessment – requirements for performing process assessment (2016)
20. Wirth, R., Hipp, J.: CRISP-DM: towards a standard process model for data mining. In: Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining, pp. 29–39. Citeseer (2000)
21. Gökalp, M.O., Kayabay, K., Zaki, M., et al.: Open-source big data analytics architecture for businesses. In: 2019 1st International Informatics and Software Engineering Conference (UBMYK), pp. 1–6. IEEE (2019)
22. ISO/IEC TR 15504-5:1999: Information technology—software process assessment—part 5: an assessment model and indicator guidance (1999)
23. Pries-Heje, J., Johansen, J.: SPI Manifesto. Eur. Syst. Softw. Process Improv. Innov. (2010)

Experience with the Performance of Online Distributed Assessments – Using Advanced Infrastructure

Damjan Ekert[1(✉)], Richard Messnarz[1], So Norimatsu[2], Tobias Zehetner[1], and Laura Aschbacher[1]

[1] ISCN GesmbH, Graz, Austria
dekert@iscn.com
[2] Norimatsu Lab, Tokyo, Japan
norimatsu@np-lab.com

Abstract. Nowadays many projects are distributed with e.g. the project management in one country, the systems development in another, and software development even on another continent. Travelling restrictions, even more so in Corona virus times, lead to a lot of issues to be solved when performing an assessment in such a largely distributed setting. This paper describes experiences with online distributed assessment and describes case studies of how this was implemented and lessons learned.

In addition the paper describes how the infrastructure helped in applying the Automotive SPICE 3.1 assessment model and the VDA Automotive SPICE Guidelines. The VDA Automotive SPICE Guidelines (blue-gold book [2]) contain rules and recommendations to interpret ASPICE 3.1 and also consider dependencies in the ratings. Assessors are now confronted with hundreds of additional rules and recommendations.

Keywords: Automotive SPICE 3.1 · VDA · Automotive SPICE Guidelines · Distributed and online assessment

1 Introduction - Setting the Scene

1.1 Case Studies

Automotive projects are largely distributed with development teams in different countries and continents and customers e.g. on a different continent. In order to save money, travels restrictions apply in many projects with an impact on the assessments. In the past already some interviews with participants e.g. from India were conducted using online web conference facilities, however the assessors and the majority of the project team was still in one location. The paper discusses 3 case studies, where the whole assessments or parts of the assessment were performed online. With each case experiences and lessons were learned.

Case Study 1: Potential customer group in Germany, large supplier in Japan, assessor team from Austria.

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 629–638, 2020.
https://doi.org/10.1007/978-3-030-56441-4_47

Case Study 2: Assessment of project in Germany, assessor team from Austria and Germany. Assessment according to Automotive SPICE 3.1 without referencing the VDA Guidelines. In order to limit travels, the assessment for Process Attribute 3.1 was performed completely online.

Case Study 3: Planned onsite process review based on Automotive SPICE 3.1 was due to the corona virus and travel restrictions performed completely online, with all participants attending the process review from home.

1.2 Assessment Model

The assessments have been performed using the Automotive SPICE 3.1 [1, 2, 3, 4–6, 9, 11, 12, 14, 26] process assessment model, which require interviews, evidence collection, rating of practices, rating of process attributes and the presentation of capability level and process attribute profiles.

And, when performing the interviews, the newly available VDA Automotive SPICE Guidelines (blue-gold book [2, 27]) provide a lot of additional checklists, rules and recommendations to consider when rating a specific base or generic practice.

2 Setting the Infrastructure

2.1 Assessment Infrastructure Applied in All Case Studies

In order to perform an online assessment some of the following requirements have to meet for the assessment tool:

Accessibility - The assessment tool should be accessible from all over the world without any restrictions which would require the involvement of the organisation's IT department.

Team Rating - Assessors are working in a team and each assessor should be able to perform his/her own rating and documenting his notes. Each assessor should have an own account.

Consolidation - An important part of the assessment is the consolidation of ratings. The assessment tool should support this activity by allowing the assessors to view each other comments and ratings.

With the usage of the Capability Adviser Assessment Tool the above criteria have been fulfilled:

The Capability Adviser is a web-based assessment tool, accessible with a standard browser. No additional installations are required. The secure communication is ensured by using the Hypertext Transfer Protocol Secure (HTTPS) protocol. Each assessor can write his own rating and his own comments, the database stores the ratings and comments from all assessors. During the assessment the assessors can see open questions in the team, differences in rating etc. and can rather ask another question to clarify the deviation instead of trying to consolidate this in a meeting at a later stage.

Comments and ratings from other assessors during the interview can been seen using the consolidation feature.

In addition, the Capability Adviser supports assessment's according to the **VDA 3.1 Guidelines**. The rules (RL) and recommendations (RC) of the VDA guidelines are displayed per BP and GP during the assessment. Moreover, the Capability Adviser database contains all relationships so that rule violations are identified and displayed to the assessor. This allows that during interviews rules, recommendations and rule violations are live displayed and can be handled directly. This then safes a lot of time in complex consolidation sessions after the interviews. In case the project is safety relevant, the **Functional Safety integration with ASPICE can be used**. In the working group SOQRATES a mapping was developed of Parts 2 to 6 of ISO 26262:2018 onto ASPICE 3.1 and this has been integrated as an extended assessment view to the tool. This allows a combination of ISO 26262 audit and ASPICE assessment. In addition to the functional safety integration also a **Cybersecurity Integration with ASPICE** is in work. In the working group SOQRATES a mapping of SAE J3061 onto ASPICE 3.1 has been developed and a mapping of ISO 21434 onto ASPICE 3.1 is in work. Once the best practice mapping has been agreed in the partnership it will be integrated as an extended ASPICE 3.1 cybersecurity view into the Capability Adviser. The SOQRATES initiative was funded in 2003 where leading automotive and electronic suppliers from Germany and Austria establish task force teams to work on new challenges like functional safety, cybersecurity etc. ISCN is the moderator and member of the SOQRATES group.

An important feature is also the possibility to import and export the assessments results between servers. In case the access to the assessment tool is not possible, the rating and comments can still be done using an excel sheet and later imported in the assessment tool (fallback solution) (Figs. 1 and 2).

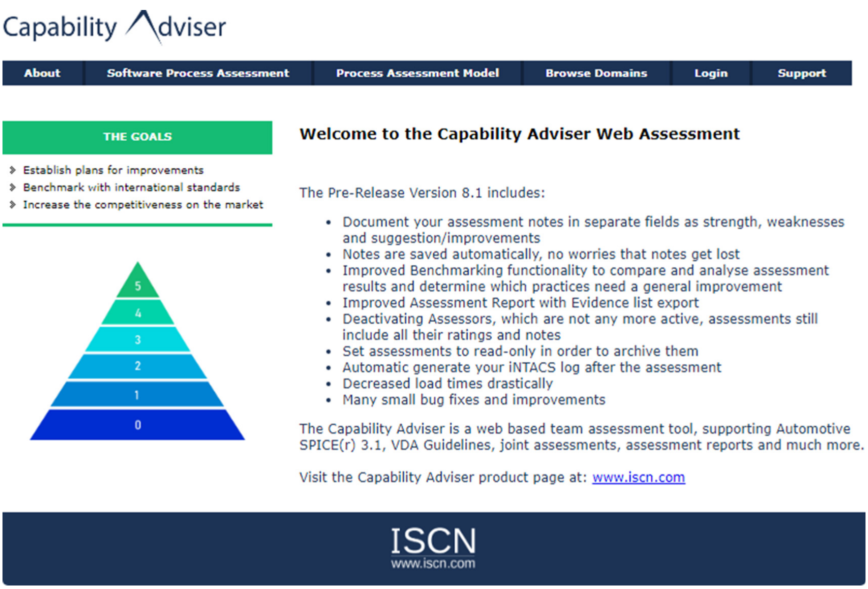
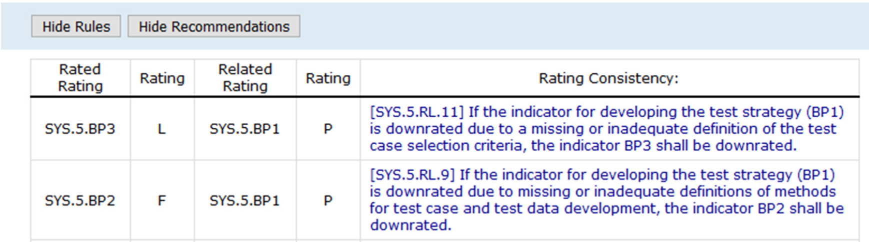


Fig. 1. Online assessment server system



Rated Rating	Rating	Related Rating	Rating	Rating Consistency:
SYS.5.BP3	L	SYS.5.BP1	P	[SYS.5.RL.11] If the indicator for developing the test strategy (BP1) is downrated due to a missing or inadequate definition of the test case selection criteria, the indicator BP3 shall be downrated.
SYS.5.BP2	F	SYS.5.BP1	P	[SYS.5.RL.9] If the indicator for developing the test strategy (BP1) is downrated due to missing or inadequate definitions of methods for test case and test data development, the indicator BP2 shall be downrated.

Fig. 2. VDA consistency view check example (theory, not in case study)

2.2 Communication Infrastructure

Case Study 1:

In case study 1 the communication infrastructure was provided by the large Japanese supplier and was based on Webex and later switched towards Microsoft Teams. The moderator of the meeting was the co-assessor from the Japanese company.

The assessment team consisted of two external assessors (one Austrian, one Japanese) and selected co-assessors from the target company to be assessed. The two external assessors were fully independent and were leading the interviews.

For the exchange of data a consolidated list of evidences was used, the evidence folder due to NDA had to stay inside the assessed company, assessors could review online but not get copies of the NDA protected materials.

Case Study 2:

The assessment was performed using a German Web Conference system called Fastviewer. The Web Conference system was used to share the screen of the presenting participant, a phone conference call was used for performing the interviews. The assessment was moderated by the lead assessor. Evidences were documented directly in the assessment tool.

Case Study 3:

The process review was performed using only Microsoft Teams. Evidences were directly documented in the assessment tool.

3 Experiences and Lessons Learned During the Assessment

Below are some of the experiences and best practices applied in the online assessment.

3.1 General

Some general observations and experiences:

- If possible, use two screens in the assessment. On one screen the presenters screen is shared on the other the assessment tool is displayed. This will reduce the effort of switching between screens.
- Test the web conference system before. Ensure that all participants are able to connect to the system.

- Mute yourself regularly when typing notes or ensure that the noise of the keyboard is not to loud and disturbing. Ask other participants to mute themselves if not speaking.
- Prepare a fallback solution in case your internet connection is down. Test your connection by connecting through the 3G or 4G network through your mobile phone. Almost all modern phones have nowadays the possibility to create an own hotspot.
- Use a comfortable headset as you will use it during the whole day.

3.2 Case Study 1 Approach and Experience

Planning Related. Instead of doing one interview plan for one week the assessment interviews have been split in different blocks along a number of days and weeks. Some days had 4 h, others 6 h, and one week later more hours. In an online setting in an inter-continental context (especially with Japan and 8-h time difference) an assessment day had minimum 2 h and max. 6 h; the duration was longer.

Tool Related. Using a work place tool at one of the assessors is not working, an infrastructure of an assessment database system was required as outlined before.

Interview and Consolidation Related. Interviews needed more time than in an onsite setting. Originally, we tried with 2 h per process, but this did not work. We usually then had 2 h interview, listed open questions, and returned to the same process later again. Open issues were marked with W(eakness) in the comments online. Assessors could see each other and see each other's rating and jointly acted in the interview. The consistency and consolidation helped to manage the assessor team performance (Fig. 3).

PC	Assessor	Score	Note
	Dr. Richard Messnarz	F	SW functions are specified. Parameters are specified. Evidence: [redacted]
	Non disclosed content for paper		
SWE.1.BP1 Specify software requirements.	**Norimatsu Messnarz**	F	There is a [redacted] Software Functional Specification. SW requirements are described per SW function. Example continued from system level. Example checked on system level [redacted] and related spec. in SYS.3 - also related Setting the register. This relates in SW specification to [redacted] on SW level - referring to a table below about which settings are done for different modes e.g. standby, streaming. Evidence: [redacted]

Fig. 3. Consolidation and team view example

Reporting Related. While each assessor had his own account in the system, the joint agreed comments were finally entered in a joint assessor account. The joint assessor in the database was like an assessor with comments to which all co-assessors agreed. The joint assessor account comments and ratings were used to generate reports.

3.3 Case Study 2 Approach and Experience

Planning Related. Automotive SPICE 3.1 assessments with a target capability level 2 with the VDA Scope are usually planned for 5 days. In case of a target capability level 3, the assessment would take additional 1,5 days resulting in a higher traveling costs for all assessors. The assessment would take place on two consecutive weeks with a weekend either onsite or most of the time traveling. Therefore, we agreed with the assessment sponsor to perform the assessment for the process attribute 3.1 online, through web conferences. For each process 45 min were planned, distributed over two weeks. An additional 30 min were planned for the consolidation afterwards, which were reduced to 15 min after the experience from the first two consolidations.

Tool Related. All assessors used the assessment online tool. There were no tool related issues during the assessment.

Interview and Consolidation Related. In most cases the 45 min were enough to perform the assessment for the process attribute 3.1. As the assessors were already checking the other assessors' comments during the assessment, the consolidation phase became very efficient. In the assessment the lead assessor was external and the co-assessors were internal (from another division in the organisation).

Reporting Related. After the assessment an assessment report was prepared by considering all assessors relevant comments. While each assessor had his own account in the system, the joint agreed comments were finally entered in a joint assessor account. The joint assessor in the database was like an assessor with comments to which all co-assessors agreed. The joint assessor account comments and ratings were used to generate and refine the assessment report. All assessors had always the possibility to access the latest version of the comments.

3.4 Case Study 3 Approach and Experience

Planning Related. A process review based on Automotive SPICE 3.1 was performed on a potential supplier. Due to the corona virus, all travels were restricted. The process review was performed online, with all participants connecting from their home. The process review was scheduled for two days, from 8:00 to 17:00 with a 45 min lunch break and regular 15 min breaks in-between.

Tool Related. All assessors used the assessment online tool. There were no tool related issues during the assessment.

Interview and Consolidation Related. For each process a responsible person was selected to present the process, others participants were supporting. The switching between the screens of the participants during the interviews was smoothly and faster than during onsite interviews. The consolidation was performed on the third day, by reviewing all the findings and based on them preparing a management summary.

Reporting Related. The assessment report was prepared in the excel with the major findings. This way the assessed organisation could also easily adjust and import the findings in their internal issue system.

4 Summary from the Case Studies

Case study 1 "Inter-continental Experience with Large Time Difference" related:

Plan a set of interview sessions for a duration longer than a week. Normally it takes about 4 to max. 6 h per day. Some days might be only 4 h.

Use an advanced infrastructure with online assessment portal and communication tools.

Prepare an evidence list together with the supplier beforehand so that there is a road map what to look at during the interviews.

Allow a team view of assessors and integrate the different assessor views into a joint assessor account in the system to be able to generate a report.

In case of functional safety be prepared for more discussions in your process area [4–7, 8, 9, 10, 26, 27].

Case Study 2: "Online Assessment for PA 3.1"

There was no drawback in performing the assessment online. The traveling costs were reduced, the assessors didn't need to travel two times or stay the weekend abroad and for each of the process enough time was dedicated, without any pressure in case of some delays. The assessed organisation was also in favour of the assessment approach.

Case Study 3: "Online Process Review"

The initial fear that an online process review is never the same as onsite process review diminished during the process review. A nice gesture was that at the introduction round everyone shared their web camera to see also the people behind the voice. The process review was a success, the assessment organisation as well as the customer were satisfied with the results. By using the assessment tool, effort and time was saved, usually needed to align comments and ratings in excel sheets.

5 Conclusions

In the VDA Guidelines, section 8.3.1 pure online assessments (phone and/or video conferences) are considered as not appropriate. However there is no such a recommendation or rule in the ISO/IEC 330xx. On the 18th of April intacs published a Guideline for Remote Process Assessments [XX], supporting online assessments during the Corona crisis.

The effort for performing online assessments/process reviews raises is similar to the onsite assessments when counting the net time. The benefit of online/part online assessments is that the dates are more flexible, the team is not bound completely to the 5 days onsite and the assessors have more time to reflect on the assessment findings. The drawback are potential technical issues which could delay or postpone the assessment, missing body language, direct communication and the coffee break discussions with the project team.

Overall, we can recommend performing parts of the assessments online (e.g. PA 3.1) to save travel costs and effort on site.

6 Relationship with the SPI Manifesto

A platform were such new cross-cutting approaches can be discussed is EuroAsiaSPI[2]. Its mission is to develop an experience and knowledge exchange platform for Europe where Software Process Improvement (SPI) practices can be discussed and exchanged and knowledge can be gathered and shared [15–17, 20]. The connected SPI manifesto defines the required values and principles for a most efficient SPI work. One main goal is to support changes by innovation and include all affected stakeholders.

The principle "**Use dynamic and adaptable models as needed**" means that an organisation can also apply a combination of process models in the improvement. Such an assessment system supporting different models and views contributes to that principle.

Another important platform for such new cross-cutting approaches is the European DRIVES project. DRIVES is a BLUEPRINT [18, 22, 24, 25, 28] project for the automotive industry and creates a strategy that support the development of new business visions for 2030. It will also emphasise the combined use of different norms.

Acknowledgements. We are grateful to the European Commission which has funded the BLUEPRINT project DRIVES (2018–2021) [22, 28]. In this case the publications reflect the views only of the author(s), and the Commission cannot be held responsible for any use, which may be made of the information contained therein.

We are grateful to a working party of Automotive suppliers SOQRATES (www.soqrates.de) who exchange knowledge about such assessment strategies. This includes: Böhner Martin (Elektrobit), Brasse Michael(HELLA), Bressau Ernst (BBraun), Dallinger Martin (ZF), Dorociak Rafal (HELLA), Dreves Rainer (Continental Automotive), Ekert Damjan (ISCN), Forster Martin (ZKW), Geipel Thomas (BOSCH), Grave Rudolf (Elektrobit), Griessnig Gerhard (AVL), Gruber Andreas (ZKW), Habel Stephan (Continental Automotive), Hällmayer Frank (Software Factory), Haunert Lutz (Giesecke & Devrient), Karner Christoph (KTM), Kinalzyk Dietmar (AVL), König Frank (ZF), Lichtenberger Christoph (MAGNA ECS), Lindermuth Peter (Magna Powertrain), Macher Georg (TU Graz & ISCN), Mandic Irenka (Magna Powertrain), Maric Dijas (Lorit Consultancy), Mayer Ralf (BOSCH Engineering), Mergen Silvana (TDK/EPCOS), Messnarz Richard (ISCN), Much Alexander (Elektrobit), Nikolov Borislav (msg Plaut), Oehler Couso Daniel (Magna Powertrain), Riel Andreas (Grenoble INP & ISCN), Rieß Armin (BBraun), Santer Christian (AVL), Schlager Christian (Magna ECS), Schmittner Christoph (Austrian Institute of Technology AIT), Schubert Marion (ZKW), Sechser Bernhard (Process Fellows), Sokic Ivan (Continental Automotive), Sporer Harald (Infineon), Stahl Florian (AVL), Wachter Stefan (msg Plaut), Walker Alastair (Lorit Consultancy), Wegner Thomas (ZF).

References

1. Automotive SPICE©3.1, Process Assessment Model, VDA QMC Working Group 13. Automotive SIG, November 2017
2. Automotive SPICE©Guidelines, 2nd edn. VDA QMC Working Group 13, November 2017
3. Höhn, H., Sechser, B., Dussa-Zieger, K., Messnarz, R., Hindel, B.: Software Engineering nach Automotive SPICE: Entwicklungsprozesse in der Praxis-Ein Continental-Projekt auf dem Weg zu Level 3, Kapitel: Systemdesign, dpunkt. Verlag (2015)
4. Christian, K., Messnarz, R., Riel, A., et al.: The AQUA automotive sector skills alliance: best practice in an integrated engineering approach. Softw. Qual. Prof. **17**(3), 35–45 (2015)
5. Messnarz, R., et al.: Integrating functional safety, automotive SPICE and six sigma – the AQUA knowledge base and integration examples. In: Barafort, B., O'Connor, R.V., Poth, A., Messnarz, R. (eds.) EuroSPI 2014. CCIS, vol. 425, pp. 285–295. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-662-43896-1_26
6. Kreiner, C., et al.: Automotive knowledge alliance AQUA – integrating automotive SPICE, six sigma, and functional safety. In: McCaffery, F., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2013. CCIS, vol. 364, pp. 333–344. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-39179-8_30
7. Macher, G., Sporer, H., Brenner, E., Kreiner, C.: Supporting cyber-security based on hardware-software interface definition. In: Kreiner, C., O'Connor, R.V., Poth, A., Messnarz, R. (eds.) EuroSPI 2016. CCIS, vol. 633, pp. 148–159. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-44817-6_12
8. Macher, G., Much, A., Riel, A., Messnarz, R., Kreiner, C.: Automotive SPICE, safety and cybersecurity integration. In: Tonetta, S., Schoitsch, E., Bitsch, F. (eds.) SAFECOMP 2017. LNCS, vol. 10489, pp. 273–285. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66284-8_23
9. Messnarz, R., Kreiner, C., Riel, A.: Integrating automotive SPICE, functional safety, and cybersecurity concepts: a cybersecurity layer model. Softw. Qual. Prof. (2016)
10. Messnarz, R., Spork, G., Riel, A., Tichkiewitch, S.: Dynamic learning organisations supporting knowledge creation for competitive and integrated product design. In: Proceedings of the 19th CIRP Design Conference – Competitive Design. Cranfield University, p. 104, 30–31 March 2009
11. Messnarz, R., König, F., Bachmann, V.O.: Experiences with trial assessments combining automotive SPICE and functional safety standards. In: Winkler, D., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2012. CCIS, vol. 301, pp. 266–275. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-31199-4_23
12. Messnarz, R., Kreiner, C., Riel, A.: Integrating automotive SPICE, functional safety, and cybersecurity concepts: a cybersecurity layer model. Softw. Qual. Prof. **18**(4), 13–23 (2016)
13. Much, A.: Automotive security: challenges, standards and solutions. Softw. Qual. Prof. (2016)
14. Stolfa, J., et al.: Automotive quality universities - AQUA alliance extension to higher education. In: Kreiner, C., O'Connor, R.V., Poth, A., Messnarz, R. (eds.) EuroSPI 2016. CCIS, vol. 633, pp. 176–187. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-44817-6_14
15. Korsaa, M., et al.: The SPI Manifesto and the ECQA SPI manager certification scheme. J. Softw.: Evol. Process **24**(5), 525–540 (2012)
16. Korsaa, M., et al.: The people aspects in modern process improvement management approaches. J. Softw.: Evol. Process **25**(4), 381–391 (2013)

17. Messnarz, R., et al.: Social responsibility aspects supporting the success of SPI. J. Softw.: Evol. Process **26**(3), 284–294 (2014)
18. GEAR 2030: European Commission, Commission launches GEAR 2030 to boost competitiveness and growth in the automotive sector (2016) http://ec.europa.eu/growth/tools-databases/newsroom/cf/itemdetail.cfm?item_id=8640
19. Bratzel, S.: Kautschuk Gummi Kunststoffe, no. 03, pp. 10–11 (2019)
20. http://2018.eurospi.net/index.php/manifesto. Accessed 2 Apr 2019
21. https://www.energie-lexikon.info/rollwiderstand.html. Accessed 6 Apr 2019
22. EU Blueprint Project DRIVES. https://www.project-drives.eu/. Accessed 6 Apr 2020
23. Much, A.: Automotive security: challenges, standards, and solutions. Softw. Qual. Prof. **18**(4) (2016)
24. GEAR 2030: High Level Group on the Competitiveness and Sustainable Growth of the Automotive Industry in the European Union (2017)
25. European Sector Skill Council: Report, Eu Skill Council Automotive Industry (2013)
26. Messnarz, R., Ekert, D.: Assessment-based learning systems - learning from best projects. Softw. Process Improv. Pract. **12**(6), pp. 569–577 (2007). https://doi.org/10.1002/spip.347. (Special Issue: Special Issue on Industrial Experiences in SPI, November/December 2007)
27. Messnarz, R., Ekert, D., Zehetner, T., Aschbacher, L.: Experiences with ASPICE 3.1 and the VDA automotive SPICE guidelines – using advanced assessment systems. In: Walker, A., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 549–562. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_42
28. Stolfa, J., et al.: DRIVES—EU blueprint project for the automotive sector—a literature review of drivers of change in automotive industry. J. Softw.: Evol. Process **32**(3) (2020). (Special Issue: Addressing Evolving Requirements Faced by the Software Industry, March 2020)

Forecasting Completion Deadlines in Software Testing

Thomas M. Fehlmann[1(✉)] and Eberhard Kranich[2]

[1] Euro Project Office AG, Zurich, Switzerland
thomas.fehlmann@e-p-o.com
[2] Euro Project Office, Duisburg, Germany
eberhard.kranich@e-p-o.com

Abstract. During the test phase of a running software development project test managers are encouraged to check continuously whether the test phase will be completed successfully at a predefined deadline at the latest. It is common practice to track the progress of testing by burn charts, especially by burn-up charts in the case when the number of test stories included in a test varies during the time line of the test phase by, e.g., adding a new test story to or removing an existing test story from the actual test where a test story is a collection of test cases. In such situations a burn-up chart and the underlying order statistics enable to estimate or to forecast a new deadline or the number of test stories successfully tested at the predefined deadline. This paper shows that utilizing a burn-up chart combined with a specific Statistical Process Control (SPC) technique leads to an efficient tool for forecasting completion deadlines in software testing.

Keywords: Software testing · Forecasting deadlines · Burn-up charts · Order statistics · Statistical Process Control (SPC) · Q control charts · Exponential smoothing

1 Introduction

When testing software, it is advisable to track and to monitor the completion status of an entire test or parts of it such as test stories or test cases. A popular monitoring tool is a burn-up chart which depicts two graphs where one graph represents the cumulated planned work of testing at points in time $t_1, t_2, \ldots, t_n$ and the other the cumulated completed work at time points preceding point in time t_k plus the completed work of testing at t_k. If $k < n$, a burn-up chart does not reveal at point in time t_k whether the predefined deadline t_n will be hit or whether the actual completion status potentially lead to a schedule overrun or even undercut with reference to t_n.

To remedy this drawback of burn-up charts a traditional approach is to generate a series of forecasts at point in time t_k, with the goal to get some insight into whether the planned deadline t_n will be hit with which variance. Clearly,

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 639–651, 2020.
https://doi.org/10.1007/978-3-030-56441-4_48

depending on the distance between t_k and t_n the forecasts and associated prediction intervals are more or less informative. That means, the more distant t_k is from t_n, the vaguer are the forecasts, or in other words, forecasts are informative when t_k is close to t_n.

It is apparent that the k cumulated sums of completed work visualized by a burn-up chart specify a statistical sample of values which are arranged in ascending order. Such ordered sample values are termed *order statistics*. To fictively complete an actual burn-up chart the k^{th} order statistic value is utilized to estimate or to forecast the missing aggregated sums at points in time $t_{k+1}, t_{k+2}, \ldots, t_n$. Consequently, the difference between the n^{th} order statistic value of the fictively completed work and that of the planned work reveals whether additional time is probably needed to finalize the testing. In this paper approaches to estimate or to forecast the extra time and the work completed at the predefined testing deadline are presented. Furthermore, it is shown how Statistical Process Control (SPC) techniques help to monitor the testing so that the forecasting results are more reliable.

Related to the SPI Manifesto [12] software testing is a prerequisite for releasing software artifacts which are accepted by users and customers and therefore add value to the business (Value B). Monitoring the software testing process combined with predicting actual testing deadlines is a basic means to improve the currently performed testing process (Principle 3).

The paper is organized as follows. Sect. 2 is concerned with order statistics and how they can be applied to solve the German tank problem. In Sect. 3 two approaches are introduced to estimate the aggregated completed work at the predefined deadline t_n. How to monitor the completed work is described in Sect. 4. Exponential smoothing as a set of forecast methods used to predict the cumulative sum of completed work at the deadline is presented in Sect. 5. The paper closes with a conclusion in Sect. 6.

2 Order Statistics

As already mentioned, a burn-up chart visualizes the total amount of work planned to be done and the amount of work completed up to an actual point in time t_k, see, e.g., Dinwiddie [1]. The corresponding data points $(t_k, w_v(t_k))$ are calculated recursively by

$$w_v(t_1) = w(t_1), \quad w_v(t_k) = w_v(t_{k-1}) + w(t_k), \tag{1}$$

for $2 \leq k \leq n$, where $w_v(t_k)$ denotes the cumulative sum of either planned or completed work at point in time t_k, and $w(t_k)$ is equal to $p(t_k)$, the vector of planned work at t_k, or to $c(t_k)$, the vector of completed work at t_k. Obviously, $w_v(t_k) \leq w_v(t_{k+1})$ for $1 \leq k \leq n-1$, since $w(t_k) \geq 0$ for $1 \leq k \leq n$. Each $w_v(t_k)$ is said to be an order statistic.

2.1 Basics of Order Statistics

Consider random variables $X_1, X_2, \ldots, X_N$. Furthermore, assume that a random sample $(X_1, X_2, \ldots, X_n)$ with $n \leq N$ which is drawn without replacement from $X_1, X_2, \ldots, X_N$. For $i \in \{1, 2, \ldots, n\}$, let the random variable $X_{(i)}$ denote the i^{th} smallest element of $\{X_1, X_2, \ldots, X_n\}$, termed the i^{th} *order statistic* for $(X_1, X_2, \ldots, X_n)$, i.e., the i^{th} order statistic is defined by

$$X_{(i)} = \min\{X_j \mid \#\{k \mid X_k \leq X_j\} \geq i\}, \tag{2}$$

yielding that $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ with the extreme order statistics

$$X_{(1)} = \min_{1 \leq i \leq n} \{X_i\} \text{ and } X_{(n)} = \max_{1 \leq i \leq n} \{X_i\}, \tag{3}$$

Siegrist [17] proves that the *probability density function* of an order statistic $X_{(i)}$ is equal to

$$P(X_{(i)} = x) = \frac{\binom{x-1}{i-1}\binom{N-x}{n-i}}{\binom{N}{n}}, \tag{4}$$

where $x \in \{i, i+1, \ldots, N-n+i\}$. Especially in the case of the extreme order statistics,

$$P(X_{(1)} = x) = \frac{\binom{N-x}{n-1}}{\binom{N}{n}}, \quad P(X_{(n)} = x) = \frac{\binom{x-1}{n-1}}{\binom{N}{n}}. \tag{5}$$

In addition, Siegrist [17] shows that the expected value of the order statistic $X_{(i)}$ is given by

$$E(X_{(i)}) = i\,\frac{N+1}{n+1}, \tag{6}$$

and the variance of $X_{(i)}$ by

$$Var(X_{(i)}) = i(n-i+1)\frac{(N+1)(N-n)}{(n+1)^2(n+2)}. \tag{7}$$

Consequently,

$$E(X_{(1)}) = \frac{N+1}{n+1}, \quad E(X_{(n)}) = n\,E(X_{(1)}), \tag{8}$$

and

$$Var(X_{(1)}) = n\,\frac{(N+1)(N-n)}{(n+1)^2(n+2)} = Var(X_{(n)}). \tag{9}$$

The characteristics of ordered statistics presented in this section assume that the number N of the random variables X_i is known. The next section demonstrates how order statistics can be utilized when the number N is unknown.

2.2 An Application: The German Tank Problem

During World War II, the Allies had major interest in reasonably reliable statements on the German war equipment, particularly on the number of tanks being produced, see Ruggles and Brodie [16] for historical notes.

In order to estimate the unknown number N of tanks the Allies gathered the serial numbers of wrecked and captured tanks (in principle they draw a random sample of size n without replacement), analyzed the serial numbers, generated corresponding order statistics and attempted to calculate an estimate, compare Goodman [4], Goodman [5].

The Frequentist Approach. It turned out that the statistic

$$T\left(X_{(i)}\right) = \frac{n+1}{i} X_{(i)} - 1 \qquad (i \in \{1, 2, \ldots, n\}) \tag{10}$$

is well suited to determine an estimate $\hat{N}$ of N with respect to $X_{(i)}$, see, e.g., Johnson [11], Limbore and Gurao [14] or the article in Wikipedia [18].

By (6),

$$E\left(T\left(X_{(i)}\right)\right) = E\left(\frac{n+1}{i} X_{(i)} - 1\right) = \frac{n+1}{i} E\left(X_{(i)}\right) - 1 = N \tag{11}$$

for $i \in \{1, 2, \ldots, n\}$, i.e., $T\left(X_{(i)}\right)$ is an *unbiased estimator* of N.

The quality of each unbiased estimator $T\left(X_{(i)}\right)$ is measured by the corresponding variance $V\left(T\left(X_{(i)}\right)\right)$, i.e., by (7) with $i \in \{1, 2, \ldots, n\}$,

$$\begin{aligned} Var\left(T\left(X_{(i)}\right)\right) &= Var\left(\frac{n+1}{i} X_{(i)} - 1\right) \\ &= \left(\frac{n+1}{i}\right)^2 Var\left(X_i\right) \\ &= \frac{(N+1)(N-n)(n-i+1)}{i(n+2)}. \end{aligned} \tag{12}$$

Clearly, for fixed N and n, $Var\left(T\left(X_{(i)}\right)\right)$ decreases for increasing i. Hence, the order statistic $T\left(X_{(n)}\right)$ provides an estimate with minimum variance and is termed a *uniformly minimum variance unbiased estimate* (UMVUE) for N yielding

$$Var\left(T\left(X_{(n)}\right)\right) = \frac{(N+1)(N-n)}{n(n+2)}, \tag{13}$$

compare (12). It is evident that $Var\left(T\left(X_{(n)}\right)\right) \to 0$ for $n \to N$.

It is common practice to quantify the uncertainty of an estimate by constructing a confidence interval with respect to a given confidence level, say q. As pointed out in the Wikipedia article [18] the one-sided confidence intervals for the UMVUE $T\left(X_{(n)}\right)$ can approximately be calculated by

$$\left[x_{(n)}, \left\lceil (1-q)^{-\frac{1}{n}} x_{(n)} \right\rceil \right], \tag{14}$$

where $x_{(n)}$ denotes the observed value of $T\left(X_{(n)}\right)$ and $\lceil a \rceil$ the smallest integer larger than a. In practice, the confidence level $q = 0.95$ is typically chosen.

The Bayesian Approach. Contrary to the frequentist approach, the Bayesian approach treats the unknown number N of random variables $X_1, X_2, \ldots, X_N$ as a random variable. Hence, N has a probability distribution in the context of the German tank problem.

The probability distribution of N is determined by the *Theorem of Bayes*, see, e.g., Held and Sabanés Bové [6] or Kruschke [13]. The theorem says that the probability of an event A under the condition an event B has been observed is equal to the probability of B under the condition A has been observed multiplied by the probability of A, divided by the total probability of B, i.e.,

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \qquad P(B) > 0, \tag{15}$$

where $P(A|B)$ is called the *posterior probability* of A, $P(B|A)$ the *likelihood* of A and B (that means, the probability that event B occurs under the condition that A occurred), and $P(A)$ the *prior probability* of A. The posterior $P(A|B)$ is often interpreted as the *credibility* of A by taking the occurrence of B into account, whereas the prior $P(A)$ reflects the credibility of A solely.

Considering the UMVUE $T\left(X_{(n)}\right)$ with $T\left(X_{(n)}\right) = x$ the posterior distribution of N is given by

$$P(N|x) = \frac{P(x|N)P(N)}{P(x)} \tag{16}$$

and the likelihood by

$$P(x|N) = \frac{\binom{x-1}{n-1}}{\binom{N}{n}} \qquad \text{for } n \leq x \leq N \tag{17}$$

and 0 otherwise, see Höhle and Held [7]. As prior distribution $P(N)$ Höhle and Held [7] propose to take an improper or non-informative unform distribution, i.e. $P(N) \propto 1$ for $N \in [0, \infty]$ and show that the posterior distribution is equal to

$$P(N|x) = \frac{n-1}{x} \binom{x}{n} \binom{N}{n}^{-1} \qquad \text{for } N = x, x+1, x+2 \ldots \tag{18}$$

and 0 otherwise, see Höhle and Held [7, Section 3.1] for further explanations and a mathematical proof.

In the context of a Bayesian framework, the posterior expected value $E(N|x)$ is termed the *Bayes estimate* of N. Hence,

$$E(N|x) = \frac{n-1}{n-2}(x-1) \qquad \text{for } n > 2, \tag{19}$$

and the posterior variance is given by

$$Var(N|x) = \frac{(n-1)(x-1)(x-n+1)}{(n-2)^2(n-3)} \qquad \text{for}\, n > 3, \tag{20}$$

compare Höhle and Held [7].

The posterior probability distribution (18) contains all actually available information on the parameter N and reflects the credibility of N under the condition that the order statistic $T\left(X_{(n)}\right)$ has been observed. In order to gather the most credible values of the parameter N when $T\left(X_{(n)}\right) = x$ is given, so-called *Highest Posterior Density* (HPD) intervals are calculated. Evidently, parameter values located close to the middle of an HPD interval tend to be more credible than those parameter values which are located at the limits of the HPD interval. Höhle and Held [7] show that the HPD intervals for the parameter N are one-sided intervals of the form $[x, N_q]$, where N_q denotes the q quantile of the posterior distribution (18) and is approximately equal to

$$N_q \approx \left\lceil \left(x - \frac{n}{2}\right)(1-q)^{-\frac{1}{n-1}} + \frac{n-2}{2} \right\rceil \tag{21}$$

with $\lceil a \rceil$ is equal to the smallest integer larger than a. Concerning q, a popular choice is $q = 0.95$.

3 Estimating Actual Completion Deadlines

In order to estimate an actual completion deadline and the aggregated completed work at that point in time two approaches are used in this section.

3.1 Five–Number Summary Estimates

The five–number summary provides information on the distribution of the components of a completed work vector c up to an actual point in time t_k where the components are arranged in ascending order. The five numbers are the minimum, three quartiles and the maximum where the first quartile comprises 25%, the median 50% and the third quartile 75% of the lowest values of completed work, compare Table 1.

Table 1. Five–number summary

i		1	2	3	
q_i	min	25%	Median	75%	max
$c_{[q_i]}$	$c_{\min}$	$c_{[q_1]}$	$c_{[q_2]}$	$c_{[q_3]}$	$c_{\max}$

Consider a quartile $c_{[q_i]}$, $1 \le i \le 3$. Then, for $j = 1, 2, \ldots, n-k$, the estimate $\hat{w}_{c_{[q_i]}}(t_{k+j})$ of the cumulative sum of completed work $w_c(t_{k+j})$ which is based on the actually observed cumulative sum of completed work $w_c(t_k)$ at point in time t_k is given by

$$\hat{w}_{c_{[q_i]}}(t_{k+j}) = w_c(t_k) + \left\lceil (t_{k+j} - t_k)\, c_{[q_i]} \right\rceil, \tag{22}$$

compare (1). Especially, $\hat{w}_{c_{[q_i]}}(t_n) = w_c(t_k) + \left\lceil (t_n - t_k)\, c_{[q_i]} \right\rceil$.

In order to estimate an actual deadline according to a quartile $c_{[q_i]}$ the left-hand side of (22) is set to $w_p(t_n)$ so that

$$\hat{t}_{c_{[q_i]}} = t_k + \left\lceil \frac{1}{c_{[q_i]}} \left(w_p(t_n) - w_c(t_k)\right) \right\rceil \tag{23}$$

yields an estimate for the point in time t_{k+j}.

3.2 Bayes Estimates

As aforementioned, the cumulative sums of completed work are order statistics so that the frequentist approach as well as the Bayesian approach explained in Sect. 2.2 can be used to estimate the cumulative sum of completed work at a predefined deadline t_n.

Starting with the cumulative sum of completed work $w_c(t_k)$ at point in time t_k with $k > 2$ the Bayesian approach yields the estimates

$$\hat{w}_c(t_{k+j}) = \begin{cases} \dfrac{k-1}{k-2}\Big(w_c(t_k) - 1\Big) & \text{for } j = 1, \\ \dfrac{k+j-2}{k+j-3}\Big(\hat{w}_c(t_{k+j-1}) - 1\Big) & \text{otherwise,} \end{cases} \tag{24}$$

compare (19). The corresponding HPD intervals are of the form $[\hat{w}_c(t_{k+j}), N_q]$ with

$$N_q \approx \left\lceil \left(\hat{w}_c(t_{k+j}) - \frac{k+j}{2}\right)(1-q)^{-\frac{1}{k+j-1}} + \frac{k+j-2}{2} \right\rceil, \tag{25}$$

compare (21).

For $j \geq 2$, estimates (24) are calculated by expressing $\hat{w}_c(t_{k+j})$ in terms of the cumulative sum of completed work $w_c(t_k)$ at an actual point in time t_k. For instance, $j = 2, 3, 4$ yields

$$\begin{aligned}
\hat{w}_c(t_{k+2}) &= \frac{k}{k-1}\Big(\hat{w}_c(t_{k+1}) - 1\Big) \\
&= k\left\{\frac{1}{k-2}\Big(w_c(t_k) - 1\Big) - \frac{1}{k-1}\right\}, \\
\hat{w}_c(t_{k+3}) &= \frac{k+1}{k}\Big(\hat{w}_c(t_{k+2}) - 1\Big) \\
&= (k+1)\left\{\frac{1}{k-2}\Big(w_c(t_k) - 1\Big) - \frac{1}{k-1} - \frac{1}{k}\right\}, \\
\hat{w}_c(t_{k+4}) &= \frac{k+2}{k+1}\Big(\hat{w}_c(t_{k+3}) - 1\Big) \\
&= (k+2)\left\{\frac{1}{k-2}\Big(w_c(t_k) - 1\Big) - \frac{1}{k-1} - \frac{1}{k} - \frac{1}{k+1}\right\} \\
&= (k+2)\left\{\frac{1}{k-2}\Big(w_c(t_k) - 1\Big) - \sum_{i=2}^{4}\frac{1}{k+i-3}\right\}.
\end{aligned} \tag{26}$$

Consequently, for $j \geq 2$ and $k > 2$,

$$\hat{w}_c(t_{k+j}) = \big(k+j-2\big)\left\{\frac{1}{k-2}\Big(w_c(t_k)-1\Big) - \sum_{i=2}^{j}\frac{1}{k+i-3}\right\}. \tag{27}$$

The difference in curly brackets is not significantly affected by the summation term so that

$$\hat{w}_c(t_{k+j}) \approx \big(k+j-2\big)\left\{\frac{1}{k-2}\Big(w_c(t_k)-1\Big)\right\}. \tag{28}$$

Hence, dividing (28) by that difference and replacing $\hat{w}_c(t_{k+j})$ with $w_p(t_n)$ results in a deadline estimate $\hat{t}_{k+j}$, i.e.,

$$\hat{t}_{k+j} \approx \left\lceil \frac{w_p(t_n)}{\frac{1}{k-2}\left(w_c(t_k)-1\right)} + 2 \right\rceil = \left\lceil \frac{(k-2)\, w_p(t_n)}{w_c(t_k)-1} + 2 \right\rceil. \tag{29}$$

4 Monitoring the Completed Work

Consider a current point in time t_k with $1 \leq k \leq n$. Then, the *completion gap* at point in time t_k is defined by

$$\Delta_{t_k} = c(t_k) - p(t_k). \tag{30}$$

Clearly, if $\Delta_{t_k} < 0$, then additional effort has to be invested during at least one of the subsequent points in time $t_{k+1}, t_{k+2}, \ldots, t_n$ to compensate the gap Δ_{t_k} in order to achieve $w_c(t_n) = w_p(t_n)$ at the deadline t_n.

The more interesting case is when $\Delta_{t_k} > 0$, since either planned work at points in time $t_1, t_2, \ldots, t_{k-1}$ has only partially been completed at point in time t_k or some of the planned work at points in time $t_{k+1}, t_{k+1}, \ldots, t_n$ has already been completed at t_k.

In any case it makes no sense to utilize the completion gap (30) at a specific point in time as an indicator for the overall testing status, since it does not take information about the completion gaps Δ_{t_j} with $1 \leq j < k$ into account.

4.1 Monitoring the Completion Gaps

Since $c(t_k) = p(t_k) + \Delta_{t_k}$ by (30), the recursive formula (1) to calculate $w_c(t_k)$ can easily be converted into a non-recursive one. Hence,

$$w_c(t_k) = \sum_{j=1}^{k} c(t_j) = \sum_{j=1}^{k}\big(p(t_j) + \Delta_{t_j}\big) = w_p(t_k) + \sum_{j=1}^{k}\Delta_{t_j}. \tag{31}$$

Consequently, for $1 \leq i \leq n-k$,

$$w_c(t_{k+i}) = w_p(t_{k+i}) + \sum_{j=1}^{k}\Delta_{t_j}, \tag{32}$$

implying that the cumulative sum of completion gaps is a capable indicator for the testing status or progress. Thus, the value of $w_c(t_n)$ may be viewed as a prediction of the cumulative completed work at the deadline, under the assumption that $\Delta_{t_{k+i}} = c(t_{k+i}) - p(t_{k+i}) = 0$ for $1 \leq i \leq n-k$.

Applying Statistical Process Control (SPC) techniques such as Shewhart–type control charts to the cumulative sum of completion gaps helps to detect whether the actual testing is under (statistical) control or even not, compare Fehlmann and Kranich [2]. In order to identify an out-of-control situation a set of rules is adopted, see e.g., Fehlmann and Kranich [2] or Hoyer and Ellis [8]. From an SPC point of view, the key purpose of deploying such rules is to signal a potential out-of-control situation before the considered testing process is really out of control.

***Q*-Statistics and *Q* Control Charts.** Quesenberry [15] considers a sequence of k independent, identically and normally distributed (i.i.d.) random variable $X_1, X_2, \ldots, X_k$ with mean μ and variance σ^2, i.e., $X_j \sim N(\mu, \sigma^2)$ for $1 \leq j \leq k$. The observed values of the random variables are denoted by $x_1, x_2, \ldots, x_k$.

In practice, the parameters μ and σ^2 are not known. Quesenberry [15] therefore proposes to apply the so-called *Q-Statistic* defined by

$$Q_k(x_k) = \Phi^{-1}\left\{G_{k-2}\left[\sqrt{\frac{k-1}{k}}\left(\frac{x_k - \bar{x}_{k-1}}{s_{k-1}}\right)\right]\right\}, \quad k \geq 3, \tag{33}$$

where $G(\cdot)$ denotes the Student t cumulative distribution function with $k-2$ degrees of freedom, and $\Phi^{-1}\{\cdot\}$ the inverse of the standard normal distribution function $N(0,1)$, and where $\bar{x}_k$ is the sample mean with

$$\bar{x}_k = \frac{1}{k}\sum_{j=1}^{k} x_j = \bar{x}_{k-1} + \frac{1}{k}\left(x_k - \bar{x}_{k-1}\right) \tag{34}$$

for $k \geq 2$ and $\bar{x}_1 = x_1$, and where s_k^2 is the sample variance with

$$s_k^2 = \frac{1}{k-1}\sum_{j=1}^{k}\left(x_j - \bar{x}_k\right)^2 = \left(\frac{k-2}{k-1}\right)s_{k-1}^2 + \frac{1}{k}\left(x_k - \bar{x}_{k-1}\right) \tag{35}$$

for $k \geq 3$ and $s_2^2 = \frac{1}{2}(x_2 - x_1)$.

Since each Q-Statistic $Q_k(x_k)$ in (33) is normally distributed with mean equal to 0 and with variance equal to 1, the 3σ upper control limit, the center line (CL) and the 3σ lower control limit (LCL) of a so-called *Q control chart* are constant, i.e.,

$$UCL = +3, \quad CL = 0, \quad LCL = -3. \tag{36}$$

In the best case, the Q-Statistics (33) oscillates closely around the center line, i.e., the considered process is under statistical control. For further details see Fehlmann and Kranich [3] and the references given therein.

Monitoring the Cumulative Sum of Completion Gaps. The Q-Statistic (33) applied to the cumulative sum of completion gaps Δ_{t_k} defined in (30) and the associated Q control chart give insight into the trend of the cumulative sum of completion gaps in the course of testing.

For instance, the Q-Statistic $Q_k(x_k)$ with $x_k = \sum_{j=1}^{k} \Delta_{t_j}$ signals an outlier at point in time t_k when $Q_k(x_k) < LCL$ or $Q_k(x_k) > UCL$. Since an outlier may strongly impact the cumulative sum of completion gaps at subsequent points in time t_j, $j > k$, it is advisable to take action. An action could be to stop the actual Q-Statistic at the point in time of the outlier occurrence and to start a new Q-Statistic from scratch at the subsequent point in time. Fehlmann and Kranich [3] exemplify in an other context that this strategy is reasonable since ignoring an actually detected outlier may mask another outlier potentially occurring at a future point in time.

Applying the previously mentioned SPC rules referenced by Fehlmann and Kranich [2] and Hoyer and Ellis [8] as often as possible to Q control charts is strongly recommended since these rules caution against a potential out-of-control situation at future points in time.

5 Forecasting Actual Completion Deadlines

It is pointed out in Sect. 4 that the cumulative sum of completion gaps calculated at a point in time t_k can be used to predict the cumulative sum of completed work at the deadline t_n, compare (32). Hence it is of interest to forecast either the completion gaps defined by (30) or the cumulative sum of completion gaps in terms of (32) at points in time t_{k+i} for $1 \leq i \leq n-k$. This section focuses on the second case solely.

5.1 Forecasting by Exponential Smoothing

Exponential smoothing refers to a class of forecasting methods. Methods based on exponential smoothing produce reliable forecasts which are weighted averages of past observations where the weights diminishes exponentially as the corresponding observations move from the most recent point in time to an earlier point in time, see, e.g., Hyndman and Athanasopoulos [10] for further details on exponential smoothing methods and Hyndman [9] for an implementation of these methods.

Simple Exponential Smoothing. This type of exponential smoothing is particularly suitable to forecast the next i or i-step-ahead values of a considered time series. Related to (32) the forecasts of the cumulative sums of completed work at points in time t_{k+i} are calculated by

$$\hat{w}_c(t_{k+i}) = w_p(t_{k+i}) + \hat{\Delta}_{t_{k+i}}, \quad i = 1, 2, \ldots . \tag{37}$$

In general, the forecasts $\hat{\Delta}_{t_{k+i}}$ are computed recursively by

$$\hat{\Delta}_{t_{k+i}} = \hat{\Delta}_{t_{k+i-1}} + \lambda \left(\sum_{j=1}^{k} \Delta_{t_j} - \hat{\Delta}_{t_{k+i-1}} \right), \tag{38}$$

where $\lambda \in [0,1]$ is the so-called smoothing parameter. It is apparent that when a small value of λ is chosen, the (initial) forecast $\hat{\Delta}_{t_{k+i-1}}$ affects subsequent forecasts more than a large value of λ. If the previous forecast $\hat{\Delta}_{t_{k+i-1}}$ is not known, it can be set to the first completion gap Δ_{t_1}.

The `R forecast` package developed by Hyndman [9] implements simple exponential smoothing by means of the function `ses` which return among others the limits of a prediction interval with respect to given confidence level.

Double Exponential Smoothing. In general, data of times series are subject to random variations. But it may be the case that the data reveal a shift to lower or higher values over a certain time period, i.e., a trend pattern of the data is present. A trend reflects the long-term direction of observed data.

In exponential smoothing a trend is of the form

$$tr(t_{k+i}) = \nu \left(\hat{\Delta}_{t_{k+i}} - \hat{\Delta}_{t_{k+i-1}} \right) + (1-\nu) tr(t_{k+i-1}), \tag{39}$$

where $\nu \in [0,1]$ is a smoothing parameter which plays an analogous role as the parameter λ in simple exponential smoothing.

Incorporating the trend (39) into (38) results in

$$\hat{\Delta}_{t_{k+i}} = \hat{\Delta}_{t_{k+i-1}} + tr(t_{k+i-1}) + \lambda \left(\sum_{j=1}^{k} \Delta_{t_j} - \hat{\Delta}_{t_{k+i-1}} - tr(t_{k+i-1}) \right), \tag{40}$$

so that $\hat{w}_c(t_{k+i})$ in (37) is a forecast with trend of the cumulative sum of completed gaps at points in time t_{k+i} in this case.

Since in (40) the smoothing parameters λ and ν are given explicitly and implicitly the procedure to calculate forecasts is termed double exponential smoothing. The `R forecast package` [9] implements double exponential smoothing by means of the functions `holt` and `hw`.

***Q*-Statistics and Exponentional Smoothing.** In Sect. 4 it is described how the cumulative sums of completion gaps can be monitored by means of Q-Statistics and Q control charts. Hence it may be of interest to calculate forecasts for the Q-Statistics at points in time subsequent to an actual point in time t_k. Such forecasts are of the form

$$\hat{Q}_{k+i}(x_k) = \hat{Q}_{k+i-1}(x_k) + \hat{\Delta}_{t_{k+i}} \quad \text{with } x_k = \sum_{j=1}^{k} \Delta_{t_j}, \tag{41}$$

where $\hat{\Delta}_{t_{k+i}}$ is equal to either (38) or (40) but the cumulative sums of completion gaps $\sum \Delta_{t_j}$ is replaced by $Q_k(x_k)$. Consequently, the forecasts $\hat{Q}_{k+i}(x_k)$ provide information whether the actual testing process may be under statistical control at points in time t_{k+i}, i.e., whether $LCL \leq \hat{Q}_{k+i}(x_k) \leq UCL$ at points in time t_{k+i}, compare Sect. 4.

5.2 Forecasting Completion Deadlines

Clearly, the calculation of forecasts $\hat{w}_c(t_{k+i})$ defined in (37) where $\hat{\Delta}_{t_{k+i}}$ is equal to either (38) or (40) is continued until the most recent forecast exceeds the cumulative sum of planned work $w_p(t_n)$ at the deadline t_n.

That means, $\hat{w}_c(t_{k+i_0}) \geq w_p(t_n)$ at a specific point in time t_{k+i_0}. From the perspective of the current point in time t_k, the pre-calculated value of i_0 indicates how much additional effort (measured in points in time) has roughly to be invested to achieve the goal $w_p(t_n)$. Consequently, if $t_{k+i_0} < t_n$, then the testing process may be finalized before the deadline; the equality $t_{k+i_0} = t_n$ predicts that the predefined deadline will be hit, i.e., the cumulative sum of completion gaps is zero at point in time t_n; otherwise, the forecast procedure is continued.

6 Conclusion

This paper describes the utilization of order statistics for estimating and forecasting an actual completion deadline in a software testing environment from the perspective of a current point in time. If the solely interest is to get a feeling whether the planned work may be completed at the predefined deadline it is sufficient to estimate the amount of completed work at that line.

Applying forecasting in conjunction with Statistical Process Control techniques such as Q-Statistics and Q control charts is beneficial when controlling and monitoring the testing progress is a further issue. Concerning the visualization of the testing progress a best as possible information is provided when a burn-up chart consisting of two graphs which are related to the cumulative sums of planned work and to the actual and predicted cumulative sums of completed work is combined with a Q control chart which is based on the cumulative sums of actual and predicted completion gaps.

Acknowledgement. The authors would like to thank the anonymous referees for their valuable comments and suggestions which led to an improvement of the paper.

References

1. Dinwiddie, G.: Feel the burn - getting the most out of burn charts. Better Softw. **11**(5), 26–31 (2009)

2. Fehlmann, Th.M., Kranich, E.: Introducing short-run control charts for monitoring the software development process. In: MetriKon 2013, Tagungsband des Software Metrik Kongresses. Fraunhofer IESE, Kaiserslautern, Germany, 14–15 November 2013, pp. 213–234. Shaker Verlag, Aachen (2013)
3. Fehlmann, Th.M., Kranich, E.: Exponentially weighted moving average (EWMA) predictions in the software development process. In: Joint Conference of the International Workshop on Software Measurement and the International Conference on Software Process and Product Measurement (IWSM-MENSURA), 6–8 October 2014, Rotterdam, The Netherlands, pp. 263–270. Institute of Electrical and Electronics Engineers (IEEE), New York (2014). https://www.researchgate.net/publication/301409874_Exponentially_Weighted_Moving_Average_EWMA_Prediction_in_the_Software_Development_Process. Accessed 31 Mar 2020
4. Goodman, L.A.: Serial number analysis. J. Am. Stat. Assoc. **47**(260), 622–634 (1952)
5. Goodman, L.A.: Serial number analysis. J. Am. Stat. Assoc. **49**(265), 97–112 (1952)
6. Held, L., Sabanés Bové, D.: Applied Statistical Inference - Likelihood and Bayes. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-642-37887-4
7. Höhle, M., Held, L.: Bayesian estimation of the size of a population. Technical paper 499. Ludwig-Maximilians-Universität München, Institut für Statistik, Munich, Germany (2006). https://epub.ub.uni-muenchen.de/2094/1/paper_499.pdf. Accessed 31 Mar 2020
8. Hoyer, R.W., Ellis, W.C.: A graphical exploration of SPC (part 2). Qual. Progr. **29**(6), 57–64 (1996)
9. Hyndman, R.J.: Package Forecast Version 8.11 - Forecasting Functions for Time Series and Linear Models (2020). https://cran.r-project.org/web/packages/forecast/. Accessed 31 Mar 2020
10. Hyndman, R.J., Athanasopoulos, G.: Forecasting - Principles and Practice. OTexts.com, Melbourne (2018). https://www.otexts.org/fpp
11. Johnson, R.W.: Estimating the size of a population. Teach. Stat. **16**(2), 50–52 (1994)
12. Korsaa, M., et al.: The SPI Manifesto and the ECQA SPI manager certification scheme. J. Softw.: Evol. Process **24**(5), 525–540 (2012)
13. Kruschke, J.K.: Doing Bayesian Data Analysis - A Tutorial with R, JAGS, and Stan, 2nd edn. Academic Press, London (2014)
14. Limbore, J., Gurao, R.: Estimation of population size using sample maximum. Int. J. Sci. Res. **5**(5), 2086–2088 (2016)
15. Quesenberry, C.P.: SQP Q charts for start-up processes and short or long runs. J. Qual. Technol. **23**(3), 213–224 (1991)
16. Ruggles, R., Brodie, H.: An empirical approach to economic intelligence in world war II. J. Am. Stat. Assoc. **42**(237), 72–91 (1947)
17. Siegrist, K.: Order statistics (2018). http://www.randomservices.org/random/urn/OrderStatistics.html. Accessed 31 Mar 2020
18. Wikipedia. German Tank Problem (2017). https://en.wikipedia.org/wiki/German_tank_problem.html. Accessed 31 Mar 2020

Metrics and Dashboard for Level 2 – Experience

Bernhardt Steger[1(✉)], Damjan Ekert[1], Richard Messnarz[1], Jakub Stolfa[2], Svatopluk Stolfa[2], and Zdenek Velart[3]

[1] ISCN GesmbH, Graz, Austria
{bsteger,dekert,rmess}@iscn.com
[2] VSB-TUO Ostrava, Ostrava, Czech Republic
{jakub.stolfa,svatopluk.stolfa}@vsb.cz
[3] SCOVECO, s.r.o., Ostrava, Czech Republic
zdenek.velart@scoveco.cz

Abstract. In the recent years, the amount of metrics collected and monitored in a project has been grown. This increases the time and effort required to create and track the success of an organization or of a particular activity, but often does not lead to an improvement in project management. In this article we will show a different approach to finding those essential metrics that should not be missing in any project status report, especially for smaller teams or application projects. We will show how to obtain a complete list of generally useful metrics per process (taking into account the VDA scope) and how to extract from this list the two or three key metrics per process based on a risk evaluation to finally arrive at a lean metric framework. The processes Change Request Management (SUP.10) and Software Requirements Analysis (SWE.1) are used as examples described in detail.

Keywords: Best practices · ASPICE 3.1 · Capability level 2 · Dashboard · Metrics · Lean quality management

1 Introduction

The experiences of the last few years show that more and more companies in the automotive sector try to establish countless statistics and key metrics in their development organizations under the influence of Automotive SPICE®. Every attribute, every little detail is measured, evaluated and tracked. This increases the effort required to determine the metrics - and yet this amount of metrics often does not lead to a clear picture of where in the project there is a need for action, because you can no longer see the wood for the trees. In the following we will show how to get a clearer view of the maturity of a project again - by choosing a lean approach for a metric framework.

1.1 Metrics and Automotive SPICE®

In the automotive sector, most developments of mechatronic systems are done according to the V-Model [16]. For each of the development processes as well as

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 652–672, 2020.
https://doi.org/10.1007/978-3-030-56441-4_49

supporting and management processes objectives and goals are defined and documented as metrics. To measure if the process goals have been achieved (or not) metrics are used. They tell you if the project is on track from various viewpoints and can give you a forecast of the performance based on the existing measures. Usually a combination of metrics is used, as there is not one metric which can cover all aspects of a development project.

Automotive SPICE© 3.1, Process Assessment Model [1–5] is used to assess the process capability of automotive development projects containing embedded systems. Metrics play an improvement role for the achievement of the capability level 1 and capability level 2 in an Automotive SPICE® assessment.

Especially in assessments performed with the purpose "process-related product risk" metrics are essential to show the expected project progress.

E.g., for the engineering processes metrics are used to measure the bidirectional traceability between the different work products. Typically, the customer requirements coverage is determined by measuring the number of links from the customer requirements specification to the system requirements specification. The test coverage is also an important metric, measuring the degree requirements are tested [6].

With Automotive SPICE® 3.1 the consistency between work products of the engineering processes became more highlighted. It addresses the content and semantics and ensures that the relevant work products are not in contradiction to each other. An evidence for ensuring the consistency between the work products is a performed review. The review coverage is therefore often measured, by counting the number of requirements or test cases which were reviewed vs. all requirements.

The static verification of software units is also typically performed using static code analysis and checks against coding standards and guidelines. The "HIS Scope Metrics" [17] are used in many projects to evaluate the software unit characteristics.

Metrics are also used to analyse problem trends (see SUP.9.BP9).

For the achievement of Level 2, metrics are an essential tool to monitor the performance of the process (GP 2.1.3) by collecting data and continuously comparing the actual status with the planned status. In order to know what to measure, the objectives for the performance of the process (GP 2.1.1) need to be defined.

As most projects in the automotive sector are assessed according to the VDA scope:

Engineering processes	Management and supporting processes
System Requirements Analysis (SYS.2)	Project Management (MAN.3)
System Architectural Design (SYS.3)	Quality Assurance (SUP.1)
System Integration and Integration Test (SYS.4)	Configuration Management (SUP.8)
System Qualification Test (SYS.5)	Problem Resolution Management (SUP.9)
Software Requirements Analysis (SWE.1)	Change Request Management (SUP.10)
Software Architectural Design (SWE.2)	Supplier Monitoring (ACQ.4)

(*continued*)

(*continued*)

Engineering processes	Management and supporting processes
Software Detailed Design and Unit Construction (SWE.3)	
Software Unit Verification (SWE.4)	
Software Integration and Integration Test (SWE.5)	
Software Qualification Test (SWE.6)	

The MAN.6 Measurement [1] process is often overlooked.

The purpose of the Measurement Process is to collect and analyse data relating to the products developed and processes implemented within the organization and its projects, to support effective management of the processes and to objectively demonstrate the quality of the products.

The outcomes of the process are defined as:

As a result of successful implementation of this process:

1) *organizational commitment is established and sustained to implement the measurement process;*
2) *the measurement information needs of organizational and management processes are identified;*
3) *an appropriate set of measures, driven by the information needs are identified and/or developed;*
4) *measurement activities are identified and performed;*
5) *the required data is collected, stored, analysed, and the results interpreted;*
6) *information products are used to support decisions and provide an objective basis for communication; and*
7) *the measurement process and measures are evaluated and communicated to the process owner.*

Therefore the process Measurement forms a good starting point to start defining and measuring metrics.

1.2 Issues and Challenges When Defining Metrics

In a first step many companies start to define/implement metrics (either triggered by Automotive SPICE® or a customer requested dashboard) for each and every working step of a process [12]. This often results in defining metrics which cannot be measured with the existing tools, additional effort in measuring manually and in the end collecting metrics which no one reads. Experience has also shown that metrics, if they are not well defined, argued, easy to access and relevant, show no impact on project or quality management and projects members rarely use them to adjust their work behaviour or prioritise tasks. [7, 8, 13, 14]. The project manager usually focuses only on the metrics related to costs and effort and is not interested in the development/ engineering related metrics.

The main challenge therefore is to define a lean framework with a set of few metrics, which on the other hand give the best feedback to the project manager about the status of the development project. From our experience not more than two, maximum three metrics should be defined per process. If we take the VDA scope for Automotive SPICE® this leads already to 30–40 metrics for all processes.

1.3 Approach How to Develop a Lean Framework of Metrics

The following approach has been taken in order to prepare a lean framework for metrics:

1. Metrics are collected for each development process based on the methods defined in Sect. 2. The goal is to define as many metrics as possible.
2. From the list of metrics, we selected the two or three metrics which from our point of view are essential. A risk evaluation was performed by asking our self "What is the risk if we don't measure it". The risk is documented and evaluated for the impact on the project objectives. If the metric is measured implicitly by another metric, we referenced it.
3. As a result of step 2, a set of metrics will remain with a high impact on project objectives and not covered by any other metric. This metrics are the essential metrics we were searching for and should be implemented.

2 Definition of Metrics

To obtain a lean framework for metrics, a list of possible metrics per development process (according to VDA Scope) must first be collected. The following methodical approach was chosen for this:

1. Questioning project team members about desired or required metrics from their point of view.
2. Analysing existing customer requirements for metrics or customer required dashboards.
3. Deriving metrics from the process goals.
4. Deriving metrics from the analysis of all base practices according to Automotive SPICE© 3.1, Process Assessment Model [1].

When collecting the possible metrics, an evaluation or sorting was deliberately avoided, each metric was included in the list.

2.1 Collecting Metrics by Using Questionnaires and Interviews

For a first list of possible metrics, different automotive suppliers were selected. On the basis of the development process landscape implemented there, roles and employees in mechatronic projects were determined. Subsequently, they were asked which metrics they would like to have for and about their daily work.

The following methods were applied:

- questionnaire by mail or individual interview
- group interview

The time horizon for answering the questionnaire was 2 working days, individual interviews lasted between 15 and 20 min. For group interviews, the groups were mainly grouped according to roles for a duration of 20 to 30 min, e.g. all system requirement managers. For agile teams, the daily stand up meeting was used, whereby at least 2 were selected in order not to exceed the maximum duration of a stand up meeting.

Experience has shown that a mixed approach - i.e. both individual and group interviews – brought the most feedback, as the project members were able to first think about their ideas and then find additional desired metrics based on the suggestions of the others. However, sometimes this approach did not lead to any feedback - there are apparently also project managers who can manage projects without metrics.

All proposals were documented in the list of possible metrics per process, no proposal was evaluated or rejected - regardless of whether or not an evaluation of the metric is possible.

2.2 Collecting Metrics by Analysing (Known) Dashboards and Customer Requirements

In a second step, the existing customer requirements were analysed in various mechatronic projects of the above selected automotive suppliers to determine whether certain metrics are required by the OEM. These were added to the list of possible metrics.

Today it is common practice of many OEMs to require their automotive suppliers to fill a template, a so-called dashboard. Using these metrics, the OEM wants to track the status and progress of each project (Fig. 1).

All metrics from the dashboards were added to the list of possible metrics.

2.3 Collecting Metrics by Analysing the Goals of the Development Processes (According to VDA Scope)

In a further step, all processes from the Automotive SPICE© 3.1, Process Assessment Model [1], were analysed. Based on the purpose and the outcomes of a process, the objectives of the respective process were defined.

From our experience as consultants in the design and establishment of a development process landscape, the goals defined there were also taken into account.

For example, from the purpose of process Software Requirements Analysis (SWE.1), "*The purpose of the Software Requirements Analysis Process is to transform the software related parts of the system requirements into a set of software requirements.*", the process goal "*Software requirements are derived from system requirements, internal requirements and the system architecture in the software requirement specification.*" is derived. The following goals were defined for the process Software Requirements Analysis (list not complete):

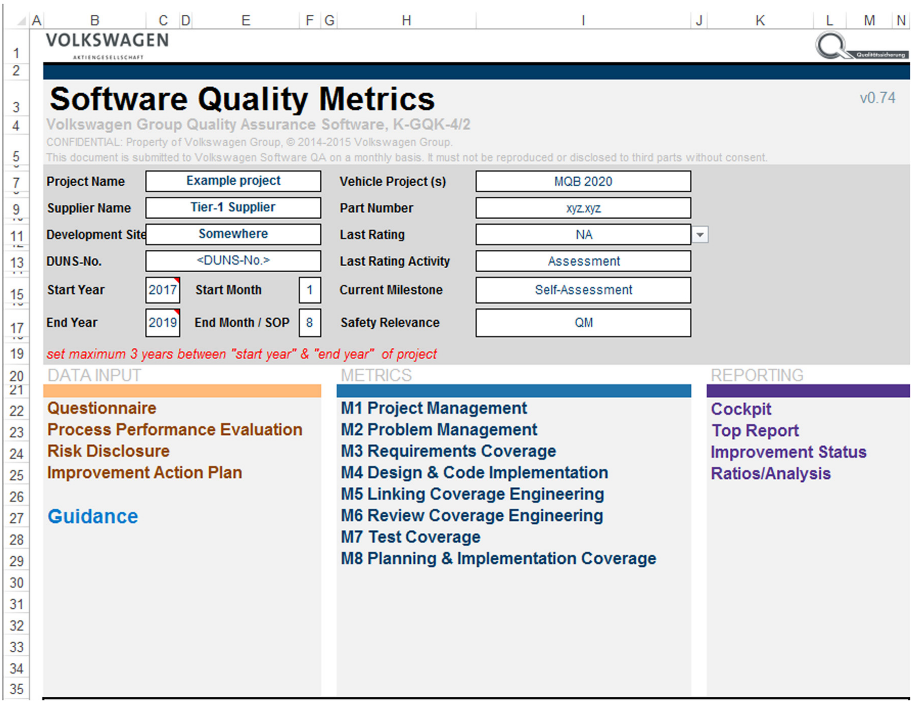


Fig. 1. Example dashboard of Volkswagen AG

- 100% of software requirements (related to current development phase) have a link from the system requirements with the domain set to software, system architecture or have the attribute "Internal Requirement" set.
- All system requirements software have been analysed and assessed and all accepted system requirements software are covered by software requirements.
- All software requirements have been reviewed and their analysis status corresponds with the review results.
- All accepted software requirements are traced via tasks in ALM tool and the direct traceability is ensured.

And the following goals were defined for the process Change Request Management (list not complete):

- No change request is longer than two weeks in status "New" or "Specified".
- All changes that can have effect on the product are referenced to the relevant change request.
- The status of the change requests and releases, and therefore the status of the project, are transparent and can be checked at all times.

Then it was defined how the individual objectives could be measured. These indicators were added to the list of possible metrics.

2.4 Collecting Metrics by Analysing Each Base Practise of Each Development Process (According to VDA Scope)

Next, the Automotive SPICE© 3.1, Process Assessment Model [1], was used to analyse all base practices of the processes and, as far as possible, metrics were derived from them.

For example, from the base practise 3 of Software Requirements Analysis (SWE.1. BP3), "*Analyze software requirements.*", the goal "*100% of software requirements (related to current development phase) have the following attributes set: Verification Criteria, Domain, Phase, TestLevel, and Safety (if applicable).*" is derived. The following goals were defined for the process Software Requirements Analysis (list not complete):

- 100% of suspect links between system and software requirements are analysed.
- 100% of software requirements related to current development phase have the status reviewed, implemented or tested.

And the following goals were defined for the process Change Request Management (list not complete):

- All change request are positively tested.
- The resolution of all changes requests of type blocker is planned. No blocker change requests are opened for current release.

2.5 Summary

Finally, the list of possible metrics was re-worked: Duplicates were eliminated, tool-specific information was generalized, and similar metrics were grouped. The list could be extended, if the project is functional safety [10, 11, 15] or cybersecurity [9] relevant. Metrics are also often additionally grouped by source (e.g. internal or customer) and affected domain (e.g. system, subsystem, software, gearbox).

As a result, the following list was created for the process Change Request Management:

ID	Possible metrics (what can be measured)
SUP.10_M01	Number of change requests not analysed within 2 weeks
SUP.10_M02	Number of change requests with a status set to New/Submitted/Open/Draft (and older 2 weeks)
SUP.10_M03	Number of all (external) change requests
SUP.10_M04	Number of change requests with a status set to New/Submitted/Open/Draft
SUP.10_M05	Number of change requests with a status set to Analysed
SUP.10_M06	Number of change requests with a status set to Approved
SUP.10_M07	Number of change requests with a status set to [Planned/in Work/Implemented/Tested]
SUP.10_M08	Number of change requests with a status set to Rejected
SUP.10_M09	Number of change requests planned for the current release

(*continued*)

(*continued*)

ID	Possible metrics (what can be measured)
SUP.10_M10	Number of change requests planned for the current release with a status implemented/Number of change requests planned for the current release
SUP.10_M11	Number of change requests planned for the current release and tested/Number of change requests planned for the current release
SUP.10_M12	Number of change requests planned for the current release and reviewed/Number of change requests planned for the current release
SUP.10_M13	Number of change requests planned for the current release and positively tested/Number of change requests planned for the current release
SUP.10_M14	Number of change requests planned for the current release and ready for delivery (status solved)/Number of change requests planned for the current release
SUP.10_M15	Number of change requests planned for the current release with links to affected work products/Number of change requests planned for the current release
SUP.10_M16	Number of change requests closed by requester

And the following list was created for the process Software Requirements Analysis:

ID	Possible metrics (what can be measured)
SWE.1_M01	Number of system architecture requirements assigned to software and linked to software requirements/Number of all system architecture requirements assigned to software
SWE.1_M02	Number of system requirements assigned to Software and linked to software requirements/Number of all system requirements assigned to software
SWE.1_M03	Number of customer requirements assigned to software and linked to software requirements/Number of all customer requirements assigned to software
SWE.1_M04	Number of system architecture requirements assigned to software and agreed by software/Number of all system architecture requirements assigned to software
SWE.1_M05	Number of system requirements assigned to software and agreed by software/Number of all system architecture requirements assigned to software
SWE.1_M06	Number of customer requirements assigned to software and agreed by software/Number of all system architecture requirements assigned to software
SWE.1_M07	Number of software requirements linked to system requirements, system architecture, marked as internal requirement or customer requirements/Number of software requirements
SWE.1_M08	Number of software requirements with filled out attribute XY/Number of all software requirements
SWE.1_M09	Number of software requirements with the functional safety relevance set/Number of all software requirements
SWE.1_M10	Number of software requirements with verification criteria defined/Number of all software requirements

(*continued*)

(*continued*)

ID	Possible metrics (what can be measured)
SWE.1_M11	Number of software requirements with release (phase) assignment/Number of all software requirements
SWE.1_M12	Number of software requirements with assigned test level/Number of all software requirements
SWE.1_M13	Number of software requirements with a status set to in work/Number of all software requirements
SWE.1_M14	Number of software requirements with a status set to changed/Number of all software requirements
SWE.1_M15	Number of software requirements with a status set to be reviewed/Number of all software requirements
SWE.1_M16	Number of software requirements with a status set to reviewed/Number of all software requirements
SWE.1_M17	Number of software requirements with a status set to implemented/Number of all software requirements
SWE.1_M18	Number of software requirements with a status set to (positively) tested/Number of all software requirements
SWE.1_M19	Number of software requirements aligned with all relevant parties/Number of all software requirements
SWE.1_M20	Number of software requirements with a status set to delivered/Number of all software requirements
SWE.1_M21	Number of functional/non-functional/process software requirements/Number of all software requirements

* The list could be extended, if the metrics would be additionally classified by release or affected software component.

3 Risk Evaluation

After a (complete) list of possible metrics was collected for each process, the risk to the project that would arise if they were not measured and tracked was evaluated for each individual metric. The risk was then assessed according to the impact on the project objectives:

- high
- medium
- low

In a further step it was examined to what extent one metric could be covered by another. If coverage is possible, the risk was mitigated accordingly.

All metrics with a high impact on the project objectives that cannot be covered by other metrics were selected for the lean framework.

3.1 Change Management Metrics

ID	Possible metrics (what can be measured)	Risk, if not measured	Impact
SUP.10_M01	Number of change requests not analysed within 2 weeks	No information how the process for analysing change requests has been performed Is covered by metric SUP.10_M02	Low
SUP.10_M02	Number of change requests with a status set to New/Submitted/Open/Draft (and older 2 weeks)	No indicator if change requests are analyzed. This could lead to an unsatisfied change requester as well as change implemented in a later stage in the project, resulting in higher development costs and also impacting project objectives	High
SUP.10_M03	Number of all (external) change requests	No ratio between external and internal change request. Number of external change requests can give an information on the degree of changes/uncertainty/stability in the project	Low
SUP.10_M04	Number of change requests with a status set to New/Submitted/Open/Draft	A high number of change requests can delay the regular development activities as resources are required to analyse the change request (MAN.3 Is covered by metric SUP.10_M02	Medium
SUP.10_M05	Number of change requests with a status set to Analyzed	No indicator on the progress of handling change requests Is covered by metric SUP.10_M14	Low
SUP.10_M06	Number of change requests with a status set to Approved	No indicator on the progress of handling change requests Is covered by metric SUP.10_M14	Low
SUP.10_M07	Number of change requests with a status set to [Planned/in Work/Implemented/Tested]	No indicator on the progress of handling change requests Is covered by metric SUP.10_M14	Low
SUP.10_M08	Number of change requests with a status set to Rejected	No indicator on the progress of handling change requests Is covered by metric SUP.10_M14	Low
SUP.10_M09	Number of change requests planned for the current release	Basis for metrics SUP.10_M10 to SUP.10_M15 Is covered by metric SUP.10_M14	High

(*continued*)

(*continued*)

ID	Possible metrics (what can be measured)	Risk, if not measured	Impact
SUP.10_M10	Number of change requests planned for the current release with a status implemented/Number of change requests planned for the current release	No information if all planned change requests will be implemented Is covered by metric SUP.10_M14	Medium
SUP.10_M11	Number of change requests planned for the current release and tested/Number of change requests planned for the current release	No information if all planned change requests will be implemented Is covered by metric SUP.10_M14	Medium
SUP.10_M12	Number of change requests planned for the current release and reviewed/Number of change requests planned for the current release*	No information if the change requests were reviewed and therefore all required information is provided and the implementation also covers the change	High
SUP.10_M13	Number of change requests planned for the current release and positively tested/Number of change requests planned for the current release	Missing indicator on the progress for the current release, can lead to customer dissatisfaction Is covered by metric SUP.10_M14	High
SUP.10_M14	Number of change requests planned for the current release and ready for delivery (status solved)/Number of change requests planned for the current release	Missing indicator on the progress for the current release, can lead to customer dissatisfaction	High
SUP.10_M15	Number of change requests planned for the current release with links to affected work products/Number of change requests planned for the current release	Not information if the change requests are completely implemented and documented Is covered by metric SUP.10_M12	Medium
SUP.10_M16	Number of change requests closed by requester	Missing information how many changes were rejected/accepted by the custommer. Could lead to customer dissatisfaction	Medium

* Not relevant if a change management workflow is implemented where reviews are required before closing/delivering the change request.

3.2 Software Requirements Analysis

For the process Software Requirements Analysis, the risk evaluation resulted in the following:

ID	Possible metrics (what can be measured)	Risk, if not measured	Impact
SWE.1_M01	Number of system architecture requirements assigned to software and linked to software requirements/Number of all system architecture requirements assigned to software	No indicator, how many system architecture requirements are already covered - the maturity of the software is therefore unclear. A (high) number of uncovered system architecture requirements can lead to a technical risk and a delay in the project as resources are required to define, (maybe) change, and link software requirements to system architecture requirements (MAN.3) Is covered by metric(s) of process System Architectural Design (SYS.3)	High
SWE.1_M02	Number of system requirements assigned to Software and linked to software requirements/Number of all system requirements assigned to software	No indicator, how many system requirements are already covered - the maturity of the software is therefore unclear. A (high) number of uncovered system requirements can lead to a technical risk and a delay in the project as resources are required to define, (maybe) change, and link software requirements to system requirements (MAN.3) Is covered by metric(s) of process System Requirements Analysis (SYS.2)	High
SWE.1_M03	Number of customer requirements assigned to software and linked to software requirements/Number of all customer requirements assigned to software	No indicator, how many customer requirements are already covered - the maturity of the software is therefore unclear. A (high) number of uncovered customer requirements can lead to a technical risk and a delay in the project as resources are required to define, (maybe) change, and link software requirements to customer requirements (MAN.3) Is covered by metric(s) of process Requirements Elicitation (SYS.1) and System Requirements Analysis (SYS.2)	High

(*continued*)

(*continued*)

ID	Possible metrics (what can be measured)	Risk, if not measured	Impact
SWE.1_M04	Number of system architecture requirements assigned to software and agreed by software/Number of all system architecture requirements assigned to software	No indicator, how many system architecture requirements are already agreed - the maturity of the software is therefore unclear. A (high) number of not agreed system architecture requirements can lead to a technical risk and a delay in the project as resources are required to define, (maybe) change, and link software requirements to system architecture requirements (MAN.3) Is covered by metric(s) of process System Architectural Design (SYS.3)	High
SWE.1_M05	Number of system requirements assigned to software and agreed by software/Number of all system architecture requirements assigned to software	No indicator, how many system requirements are already agreed - the maturity of the software is therefore unclear. A (high) number of not agreed system requirements can lead to a technical risk and a delay in the project as resources are required to define, (maybe) change, and link software requirements to system requirements (MAN.3) Is covered by metric(s) of process System Requirements Analysis (SYS.2)	High
SWE.1_M06	Number of customer requirements assigned to software and agreed by software/Number of all system architecture requirements assigned to software	No indicator, how many customer requirements are already agreed - the maturity of the software is therefore unclear. A (high) number of not agreed customer requirements can lead to a technical risk and a delay in the project as resources are required to define, (maybe) change, and link software requirements to customer requirements (MAN.3) Is covered by metric(s) of process Requirements Elicitation (SYS.1) and System Requirements Analysis (SYS.2)	High
SWE.1_M07	Number of software requirements linked to system requirements, system architecture, marked as internal requirement or customer requirements/Number of software requirements	Maybe software is implementing a feature that is not desired or even worse contradicts a defined system feature* Is covered by metric SWE.1_M16	High

(*continued*)

(*continued*)

ID	Possible metrics (what can be measured)	Risk, if not measured	Impact
SWE.1_M08	Number of software requirements with filled out attribute XY/Number of all software requirements	Unclear maturity/status of software requirements. The following processes SWE.2-6 may have no information whether implementation/testing may start or not Not selected as (SWE.1_M16) will identify missing values	High
SWE.1_M09	Number of software requirements with the functional safety relevance set/Number of all software requirements	Unclear maturity/status of software requirements. Are the software requirements analysed according to functional safety? Not selected as (SWE.1_M16) will identify missing values	High
SWE.1_M10	Number of software requirements with verification criteria defined/Number of all software requirements	No indicator, how many software requirements have a verification criteria. A (high) number of undefined verification criteria can lead to a delay in the project as resources are required to define and align verification criteria (MAN.3) Not selected as (SWE.1_M16) will identify missing values	Medium
SWE.1_M11	Number of software requirements with release (phase) assignment/Number of all software requirements	Unclear, what shall be implemented in current release - no progress tracking possible for a single software release Not selected as (SWE.1_M16) will identify missing values	High
SWE.1_M12	Number of software requirements with assigned test level/Number of all software requirements	These information makes the work just easier	Low
SWE.1_M13	Number of software requirements with a status set to in work/Number of all software requirements	These information makes the work just easier	Low
SWE.1_M14	Number of software requirements with a status set to changed/Number of all software requirements	These information makes the work just easier	Low

(*continued*)

(*continued*)

ID	Possible metrics (what can be measured)	Risk, if not measured	Impact
SWE.1_M15	Number of software requirements with a status set to be reviewed/Number of all software requirements	Unclear maturity of software requirements - how many software requirements can be reviewed? A (high) number of not reviewed software requirements can lead to a technical risk and a delay in the project as resources are required to define and (maybe) change software requirements (MAN.3). Additionally, to review software requirements that are not ready to be reviewed will also lead to a delay in the project as more findings will be documented (and have to be solved)	Low
SWE.1_M16	Number of software requirements with a status set to reviewed/Number of all software requirements	Unclear, whether the software requirements meet the specification at all - thus it is not clear whether the later implementation will meet the specifications (system/customer) at all**	High
SWE.1_M17	Number of software requirements with a status set to implemented/Number of all software requirements	Unclear status regarding the maturity/functionality of the software: Which features are implemented, which features are missing? Software qualification test has no information what to test, therefore maybe tests are performed which do not work yet (because the feature is not implemented yet). (These tests probably fail anyhow.) Is covered by metric SWE.1_M18 (and software qualification tests knows what to test because of SWE.1_M16)	High
SWE.1_M18	Number of software requirements with a status set to (positively) tested/Number of all software requirements	Unclear status of software, whether software is ready for delivery and which features are correctly implemented (and do not have to be changed) Missing information of maturity of software for the customer. Could lead to customer dissatisfaction	High

(*continued*)

(*continued*)

ID	Possible metrics (what can be measured)	Risk, if not measured	Impact
SWE.1_M19	Number of software requirements aligned with all relevant parties/Number of all software requirements	Unclear maturity of software requirements - are they clear or still unclear for any of the affected parties? A (high) number of unaligned software requirements can lead to a delay in the project as resources are required to define and align software requirements (MAN.3)	Medium
SWE.1_M20	Number of software requirements with a status set to delivered/Number of all software requirements	Unclear, which features are delivered to customer Most likely covered by Metric SWE.1_M18 as successfully tested software features are delivered to customer	Low
SWE.1_M21	Number of functional/non-functional/process software requirements/Number of all software requirements	These information makes the work just easier	Low

* Review is possible without traceability (it is just more complex and may lead to more effort consumption) - approach grouping by feature is supporting this approach [18].

** Constraints: Review shall include consistency according to traceability, whether the release is set, consistency to source, verification criteria is defined (where needed), functional safety classification (if product is ASIL > QM) given, cyber security classification given, and software architecture/affected software component is set. All relevant parties are participants of these reviews. Period of these reviews shall be defined properly (not only 1 review per year).

4 List of Lean Metrics

4.1 SUP.10 Change Request Management

The following metrics have been identified for the lean framework:

ID	Necessary metrics of lean framework	Risk, if not measured
SUP.10_M02	Number of change requests with a status set to New/Submitted/Open/Draft (and older 2 weeks)	No indicator if change requests are analyzed. This could lead to a unsatisfied change requester as well as change implemented in a later stage in the project, resulting in higher development costs and also impacting project objectives

(*continued*)

(*continued*)

ID	Necessary metrics of lean framework	Risk, if not measured
SUP.10_M12	Number of change requests planned for the current release and reviewed/Number of change requests planned for the current release	No information if the change requests were reviewed and therefore all required information is provided and the implementation also covers the change
SUP.10_M14	Number of change requests planned for the current release and ready for delivery (status solved)/Number of change requests planned for the current release	Missing indicator on the progress for the current release, can lead to customer dissatisfactio

One of the tools that automotive development companies are using for task and change management nowadays is IBM Jazz platform Rational Team Concert (RTC) – a complete tooling environment for task, change and source configuration management [19]. With regard to this fact an example implementation is shown on RTC Jazz cloud platform using BIRT reporting tool [20] that is integrated in the Jazz platform.

Reports can be implemented as trend or numerical value graphs by using Jazz Eclipse Client (BIRT report). Reports can have a form of dashboard widget (for e.g. overview purpose) or classical report (for full report purpose) and is published for general usage to Jazz server by updating the process configuration.

For example, SUP.10_M02 was implemented as a trend graph using integrated BIRT tool and used by users in their dashboard directly.

Simple explanation of the real implementation can be described by the following pseudo SQL query:

Timestamp, COUNT(Timestamp) FROM WorkItems WHERE Type = "Change Request" AND State IN ("New", "Submitted", "Open", "Draft") AND (Timestamp – Created > 2 weeks)

Where *Timestamp* is the actual date of the snapshot when the data were taken and *Created* is the date when the change request was created. Jazz platform performs a snapshot of database every day and data is collected from these snapshots for requested date interval filtered by defined projects, iterations and other possible parameters. Final data set is presented as a line graph (see Fig. 2).

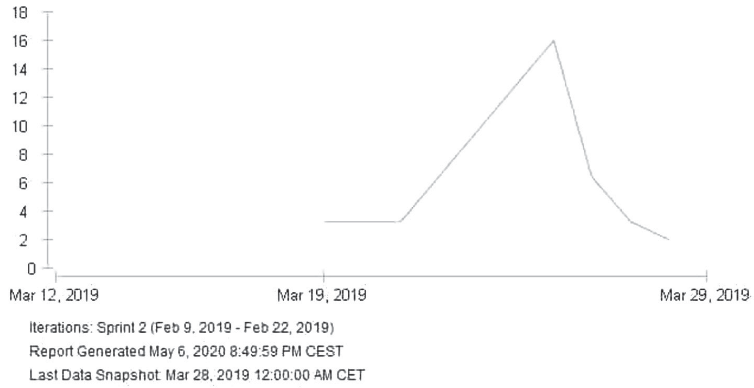


Fig. 2. Example trend chart for SUP.10 M02 (Number of change requests with a status set to New/Submitted/Open/Draft [and older 2 weeks])

4.2 SWE.1 Software Requirements Analysis

The following metrics have been identified for the lean framework:

ID	Necessary metrics of lean framework	Risk, if not measured
SWE.1_M16	Number of software requirements with a status set to reviewed/Number of all software requirements	Unclear, whether the software requirements meet the specification at all - thus it is not clear whether the later implementation will meet the specifications (system/customer) at all
SWE.1_M18	Number of software requirements with a status set to (positively) tested/Number of all software requirements	Unclear status of software, whether software is ready for delivery and which features are correctly implemented (and do not have to be changed) Missing information of maturity of software for the customer. Could lead to customer dissatisfaction

Requirements management is a crucial part of the projects. E.g., IBM Engineering Requirements Management DOORS Next (DNG) [21] is a part of the Jazz platform and allows to manage requirements lifecycle.

DNG uses a Jazz Reporting Service (JRS) as a source of reports. Unfortunately, at the time when requirements metrics were implemented, Jazz platform did not provide a snapshot service for the DNG database, therefore another approach of presenting trend graphs has to be choosen. JRS report can be executed by external applications and can be used as source for manually or automatically selected data snapshots. In our case, the request for presenting requirements management metrics was implemented by Microsoft Excel.

For example, SWE.1_M16 was implemented in excel as a trend graph with a

- sheet for snapshot data

Review Coverage: SW Reqs.			
Example Project			
Date	SWRS	SWRS reviewed	Coverage
21.01.2019	205	15	7,32%
04.02.2019	205	32	15,61%
18.02.2019	314	31	9,87%
04.03.2019	317	70	22,08%
18.03.2019	270	65	24,07%
01.04.2019	298	120	40,27%
15.04.2019	311	231	74,28%
29.04.2019	309	303	98,06%
13.05.2019	309	303	98,06%

- sheet for graph representation and minigraph in overview dashboard (see Fig. 3) with on demand data snapshot update.

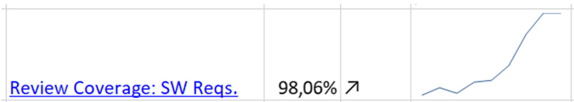


Fig. 3. Example implementation of for SWE.1_M16 (Number of software requirements with a status set to reviewed/Number of all software requirements)

Simple explanation of the real implementation can be described by the following pseudo SQL queries:

COUNT(*) FROM SWRS Modules WHERE Type = "Requirement" AND Status IN ("Reviewed", …)
COUNT(*) FROM SWRS Modules WHERE Type = "Requirement"

Result data are processed by Excel's Visual Basic for Applications programming language and corresponding data and graph sheets are updated.

5 Conclusion

Based on the experience that a large number of metrics does not necessarily lead to better management in a project, we have chosen a new approach to identify those metrics for the respective processes that are essential for the success of a project, based on a risk evaluation. As an example we have selected two processes to illustrate the approach and the result. We were able to downgrade the priority of more than 80% of the possible metrics in each process - two to a maximum of three metrics could be identified per process to be part of a lean framework (although each of the possible metrics obviously has its meaning and usefulness). This not only reduces the effort, but can also focus attention on the essentials in project status reports, as the metrics no longer disappear in the quantity of the entire metric – the focus can be set on the key performance indicators (i.e. the chosen metrics from the lean framework).

IBM Jazz platform was used as a reference tool for change request and requirements management purposes. Two possible metrics implementations were presented on selected metrics. First one using an integrated environment of RTC Jazz, where all support for required metric was present in the platform itself. Second solution used an external access to the source data of DNG to be able to collect necessary snapshots. Regarding the implementation of the metric dashboard/s, it always depends on customer requests and decisions which imply what specific solution or combination of them will be used.

E.g. in case that data from more than one tool need to be integrated and analyzed, it might be beneficial and we can use a complete external presentation and metrics analysis solution application to create desired reports for developers and management.

Acknowledgements. Work is partially supported by Grant of SGS No. SP2020/62, VŠB - Technical University of Ostrava, Czech Republic.

References

1. Automotive SPICE©3.1, Process Assessment Model, VDA QMC Working Group
2. Automotive SPICE©Guidelines, 2nd edn., VDA QMC Working Group 13, November 2017
3. Miklos, B., Richard, M., Davidson, A.G.: The impact of national cultural factors on the effectiveness of process improvement methods: the 3rd dimension. In: Proceedings of the 11th ICSQ Conference, ASQ, USA (2001)
4. Biró, M., Messnarz, R., Colomo-Palacios, R.: Software process improvement leveraged in various application domains. J. Softw.: Evol. Process **26**(5), 465–467 (2014). (Special Issue: Software process improvement leveraged in various application domains)
5. Biró, M., Messnarz, R.: Key success factors for business based improvement. In: Proceedings of the EuroSPI 1999 Conference, Pori (1999). (Pori School of Technology and Economics. Ser. A., 25.)
6. Messnarz, R., Sehr, M., Wüstemann, I., Humpohl, J., Ekert, D.: Experiences with SQIL – SW quality improvement leadership approach from Volkswagen. In: Stolfa, J., Stolfa, S., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2017. CCIS, vol. 748, pp. 421–435. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64218-5_35
7. Korsaa, M., et al.: The SPI manifesto and the ECQA SPI manager certification scheme. J. Softw.: Evol. Process **24**(5), 525–540 (2012)
8. Korsaa, M., et al.: The people aspects in modern process improvement management approaches. J. Softw.: Evol. Process **25**(4), 381–391 (2013). (Special Issue: Selected Industrial Experience Papers of EuroSPI 2010)
9. Macher, G., Sporer, H., Brenner, E., Kreiner, C.: Supporting cyber-security based on hardware-software interface definition. In: Kreiner, C., O'Connor, R.V., Poth, A., Messnarz, R. (eds.) EuroSPI 2016. CCIS, vol. 633, pp. 148–159. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-44817-6_12

10. Macher, G., Messnarz, R., Kreiner, C., et. al.: Integrated safety and security development in the automotive domain. Working Group 17AE-0252/2017-01-1661, SAE International, June 2017
11. Macher, G., Much, A., Riel, A., Messnarz, R., Kreiner, C.: Automotive SPICE, safety and cybersecurity integration. In: Tonetta, S., Schoitsch, E., Bitsch, F. (eds.) SAFECOMP 2017. LNCS, vol. 10489, pp. 273–285. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66284-8_23
12. Messnarz, R., Tully, C.J., Biro, M.: Better Software Practice for Business Benefit: Principles and Experiences. IEEE Computer Society Press, Los Alamitos (1999)
13. Messnarz, R., Ekert, D., Reiner, M., O'Suilleabhain, G.: Human resources based improvement strategies—the learning factor. J. Softw.: Evol. Process **13**(4), 355–362 (2008)
14. Messnarz, R., Spork, G., Riel A., Tichkiewitch, S.: Dynamic learning organisations supporting knowledge creation for competitive and integrated product design. In: Proceedings of the 19th CIRP Design Conference – Competitive Design, Cranfield University, 30–31 March 2009, p. 104 (2009)
15. Messnarz, R., Kreiner, C., Riel, A.: Integrating automotive SPICE, functional safety, and cybersecurity concepts: a cybersecurity layer model. Softw. Qual. Prof. (2016)
16. Messnarz, R., Much, A., Kreiner, C., Biro, M., Gorner, J.: Need for the continuous evolution of systems engineering practices for modern vehicle engineering. In: Stolfa, J., Stolfa, S., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2017. CCIS, vol. 748, pp. 439–452. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64218-5_36
17. Helmar, K.: HIS Source Code Metrics (2008). https://emenda.com/wp-content/uploads/2017/07/HIS-sc-metriken.1.3.1_e.pdf
18. Ekert, D., Steger, B., Messnarz, R.: Functional roadmap - scheduling in complex mechatronic projects. In: Barafort, B., O'Connor, R.V., Poth, A., Messnarz, R. (eds.) Systems, Software and Services Process Improvement (EuroSPI 2014). Communications in Computer and Information Science. Springer, Cham (2014)
19. IBM Workflow Management. https://www.ibm.com/products/ibm-engineering-workflow-management. Accessed 30 Apr 2020
20. BIRT. https://www.eclipse.org/birt/. Accessed 30 Apr 2020
21. IBM Engineering Requirements Management DOORS Next. https://www.ibm.com/cz-en/marketplace/requirements-management-doors-next. Accessed 30 Apr 2020

Application of Mutually Integrated International Standards (A-SPICE PAM 3.1 & IATF 16949/2016)

Samer Sameh(✉) and Sherif EL-Toukhi(✉)

Valeo, Giza, Egypt
{samer.sameh, sherif.eltoukhi}@valeo.com

Abstract. In automotive industry, different standards such as Automotive SPICE PAM 3.1 and IATF 16949:2016 are always followed to ensure that organizations will deliver outcomes with high quality and maturity level considering that the automotive industry is one of the top industries worldwide in the quality standards. Application of both standards proved to have common requirements and practices to meet their defined objects, but with different elaboration of the same requirement in across the different standard. In real life, organization has to define to different internal entities to ensure compliance to the both standards, which leads to deficiency in terms of efforts and cost. This paper is an extension to a previous paper [17, 18] that suggested a new automotive SPICE practices that enables assessments to check compliance of the two standards in an integrated way and map the non-conformities detected to requirements of both standards. In this paper, demonstration will be presented through a case study performed with a certain methodology to illustrate the effectiveness and usability of application of new proposed practices to detect some gaps common between the two standards by only conducting one project assessment.

Keywords: Automotive SPICE PAM 3.1 · Integrated international standards · Automotive software · Improved assessment models

1 Introduction

Automotive industry is well known for high quality standards among the whole supply chain from Suppliers until final OEM and even final consumers. For many years, the automotive industry always required a high co-operation between all supply chain stakeholders considering engineering specialities, mechanical, Optics, software, hardware, production, etc. This led to processes having to be governed by a robust framework to validate that all necessary steps taken to design, verify, and produce parts to meet the requirements for the products and services of the customer in mass production with consideration of supporting and management functions. ISO 9001:2008

S. Sameh—Competent ASPICE Assessor, Process Improvement Expert & DBA Candidate.
S. EL-Toukhi—Provisional ASPICE Assessor & Principal Quality Assurance Engineer.

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 673–683, 2020.
https://doi.org/10.1007/978-3-030-56441-4_50

for Quality management systems Requirements was used to derive ISO/TS 16949:2009 to be as the global standard of quality in the automotive industry. Recently many automotive suppliers went through a transition phase to move from ISO/TS 16949:2009 to IATF 16949:2016 that was updated upon new ISO 9001:2015 release since IATF 16949 is considered the most important and famous global standard of quality in the automotive industry. By the paradigm shift happened in the in industry in the recent years by the rising of new fields like autonomous driving, connected cars and electrical vehicles. Complexity of electronics and software content dramatically increase. As in such complex products and systems, multiple hardware with wide variety of electronics, software and mechatronics components are becoming more merged and integrated together with safety considerations to control. Systems complexity and criticality due to safety requirements lead to a raise in the number of anomalies and failures making it critical to put a major focus on quality, safety and security on all levels such as (system, processes, quality, etc. …). Considering the original automotive experiences and awareness of the importance of non-conformance detection as early as possible in defined operations flow to challenge costs competencies, avoid delays to plan, ensure operations excellence and eventually have the ability to compete in market. The challenge is that IATF 16949 was not originally developed to target software development organization. Accordingly, different software assessments methodologies are used to evaluate the software processes, such as Automotive SPICE and CMMI [6]. With more attention to the automotive industry nature, Automotive SPICE process assessment model Version 3.1 had become a well-known and familiar method for automotive software development companies and car manufactures to use to assess the software projects. Automotive SPICE is not only a standard used as a to determine maturity and capability levels of processes but also to improve processes and operations. It was developed as a customization from the initial ISO/IEC. Generally, SPICE® applies to the development of mechatronic systems focusing on the software and system parts of the product [1] below in addition to other engineering disciplines such as hardware engineering and mechanical engineering like the SPICE for mechanical engineering that was released for review in December 2018. Moreover, to address the challenges in rapidly growing new fields in the industry, many quality-related standards and process assessment models/frameworks in the automotive industry are being improved regularly to empower different organizations to achieve the expected level of Quality.

1.1 Scope

This paper is many addressing the automotive tier 1 suppliers in general and automotive software suppliers are providing software solutions in the automotive field in particular. As currently, most of the OEMs require suppliers to be certified to both IATF 16949 and Automotive SPICE as a prerequisite for becoming and remaining a supplier to these OEMs. Also, OEMs need evidences that the previous project for the same OEM has achieved Automotive SPICE specific CL2 (Capability Level 2) or CL3 (Capability Level 3) in some cases. So, for software development organizations, the compliance to Automotive SPICE and the IATF 16949 becomes a must to stay existing and competing in tough market.

1.2 Problem

In the recent years, automotive suppliers and automotive software companies are working hard to achieve the highest possible maturity and capability levels in operation to be able to deliver high end products. A significant effort in terms of budget, time, work force and even tools, are invested to build knowledge in the two standards and implement these standards requirements. Later the organizations have to perform internal audits and assessments to ensure full compliance and overcome these gaps with protective, corrective and preventive actions. So the suppliers in the automotive supply chain are working to apply both models separately. An entity focusing on compliance with the IATF 16949 standard for audit and certification purposes. On the other side, another entity with more technical focus are focusing to achieve the compliance with the Automotive SPICE. This add a huge efforts over the organization to meet the needed objectives. So as an example there is a total of 281 mandatory requirements that need to be fulfilled in IATF 16949.

From experience, switching the context from one standard to another and preparing for internal audits versus the two standards is not an easy mission at all. Imagine that after each internal check cycle, the process owners and the different stakeholders needs to define improvement action plan to tackle the detected gaps which might be similar in sometimes and totally not relevant in another the audit expectations.

After a deep observation & analysis of both standards needs, it can become easy and obvious to find out that both of them aim, at the high level, to produce consistently quality services and products, encouraging continuous improvement culture and support decision making. One can notice common areas between the two standards on the requirements level. IATF 16949 introduced some new requirements for automotive suppliers focusing on management accountability, risk management and assessment of embedded software. Automotive SPICE on the other hand is focusing in its process areas (management, Supporting and software engineering processes) [1].This means that it is expected to find many common practices to achieve the required quality objectives. According to Automotive SPICE PAM v3.1, Automotive SPICE only can not be only considered as a standalone frame work.

Yet, it can be part of an integrated solution of frameworks including variant development methods and an organizational culture that encourages initiatives, communication and decision decentralization and this is exactly what IATF 16949 is tackling. As the IATF 16949 is talking about the organization culture, business context, communication, management and leadership and the operations. To determine the process capability in Automotive SPICE assessments, what is examined is typically a project (for up to level 2). For capability level 3 and above, the assessors additionally evaluate organizational standard processes which mean that the area of coordination and collaboration between Automotive SPICE and IATF 16949 will be much bigger. As shown in (Fig. 1) IATF 16949 can be considered as a framework (more oriented into “WHAT”) in which Automotive SPICE will be the process assessment model for the embedded software processes of the organization (addresses “WHAT” and slightly “HOW” in the software part in terms of best practices). The suggested solution will help us to identify the common requirements (“WHAT”) between the two standards.

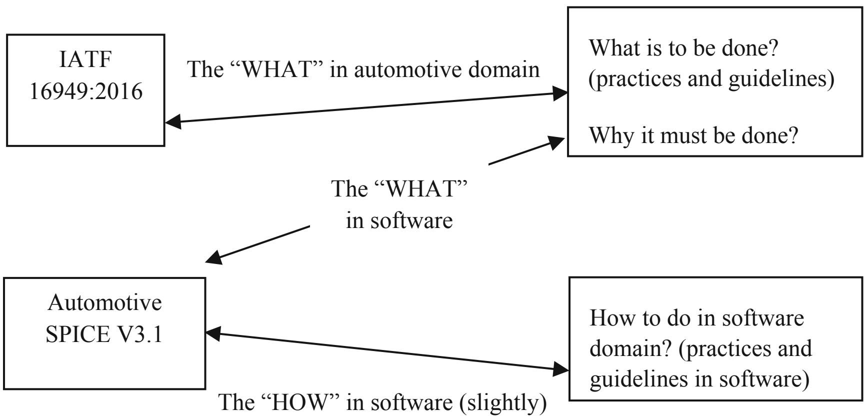


Fig. 1. Big picture of the scope difference between IATF 16949 and Automotive SPICE

1.3 Relationship with the EUROSPI Manifesto

Based on the principle "Use dynamic and adaptable models as needed" from EURO Software Process improvement (EURO SPI) Manifesto [7], which aims to drive organizations improvements for software development processes through applying a combination of process models.

Also, based on the principal "Apply risk management" that enforces the organizations to consider and follow the risk based thinking methodology which is aligned global direction of the IATF 16949-2016 requirements that were derived in line with ISO9001-2015 requirements.

Accordingly, the following solution was suggested as to improve the organization ability for defects detection with the use of a new improved assessment model for software development processes in the project.

1.4 Suggested Solution

Accordingly, IATF 16949 and Automotive SPICE can be integrated and mapped on the requirements level to increase the organizations efficiency and operations level of excellence saving time and efforts to achieve and meet the required conformance target of both standards.

Not only mapping between the requirements of standards helps a lot in finding the requirements that are expressing same concepts and achieving the same purpose, yet it also helps project to detect different new gaps that can be a drive to process improvements.

This improvement proposal was suggested by a previous paper called "Automotive Standards Compliance Cost Optimization by Mutual Integration between Standards" published on July 2019 that suggests a new updates to Automotive SPICE Base practices in certain process areas in addition to some generic practices to have the mapping with IATF 16949 requirements. With focus on application of this mapping, the is paper will be demonstrating a case study of an automotive software project assessment using new proposed Automotive SPICE practices as follow: Sect. 2. Case study methodology and scope, Sect. 3. Case Study Observations and Results Consolidation, Sect. 4. Study Results, Sect. 5. Conclusion and Recommendations For Future Work.

2 Case Study Methodology and Scope

A pilot software project was created especially for the purpose of this case study. That project was initially assessed by Automotive SPICE Process assessment model version 3.1 targeting capability level 2 for following process areas:

- System Engineering Process Group (SYS)
  - SYS.2 System Requirements Analysis
  - SYS.3 System Architectural Design
  - SYS.4 System Integration and Integration Test
  - SYS.5 System Qualification Test
- Software Engineering Process Group (SWE)
  - SWE.1 Software Requirements Analysis
  - SWE.2 Software Architectural Design
  - SWE.3 Software Detailed Design and Unit Construction
  - SWE.4 Software Unit Verification
  - SWE.5 Software Integration and Integration Test
  - SWE.6 Software Qualification Test
- Supporting Process Group (SUP)
  - SUP.1 Quality Assurance
  - SUP.8 Configuration Management
  - SUP.9 Problem Resolution Management
  - SUP.10 Change Request Management
- Management process group (MAN)
  - MAN.3 Project Management

The study methodology followed in this work was based on following steps (Fig. 2):

Fig. 2. Methodology followed in the case study presented in this paper

Data Collection: In this step, initial assessment package and documentation were collected to be used as a main input in case study. In addition to the assessment results were provided to be used in the further steps. Main focus process areas for study were determined.

Analysis and Assessment: In this step, the collected previous documentation was re-assessed based on the new suggested automotive practices taking into consideration the original assessment findings.

Results Consolidation and Investigation: In this step analysis and assessment obtained data from previous step are consolidated. An investigation was carried out to determine the differences between the remarks from the original assessment and new assessment.

Conclusion: In this step a conclusion is given based on application of new proposed practices in the case study to well-known project.

3 Case Study Observations and Results Consolidation

A case study was conducted to study the impact on the BPs/GPs and the final rating of Process Areas [PAs] on a specially created project for the study purpose and was assessed in order to be able to record the observations about the differences between assessment of some practices of Usual A-SPICE model and proposed updated A-SPICE practices.

Based on the historical data available for the case study, not all proposed practices were applicable to use for different reasons such as for an example: The original assessment scope (selected process areas), insufficient data from the original assessment report (some process areas did not show any remarks that were reflected in the assessment report to compare the new observation with) and finally targeted capability level as per original assessment scope.

The focus was given mainly for capability level 1. No detailed ratings were provided considering only the recording of findings to mention. The following Table 1 demonstrates the detailed observation to the assessed case showing the original ASPICE Practices verses the proposed ASPICE Practices, Newly detected gaps based on proposed ASPICE practices and finally a conclusion for each practice:

Although the case study initial conditions did not allow having complete piloting for all proposed practices in the ASPICE across different capability levels (which is referred to later in Sect. 5. Conclusion and recommendations For Future Work). Yet, it gives an overview about the main differences between applying the two practices to assess projects in the organization. Even if this was applied to an imaginary project (specially created for study purpose) which no applicability to check the closure of new weakness detected, this remains a direct application to the concepts of quality continuous improvements in all targeted organization activities.

Table 1. Application of the new proposed PAM & some findings

Current PAM wording	Proposed changes in the PAM	New weaknesses detected based on modified PAM	Conclusion
MAN.03-BP1 Define the scope of work. Identify the project's goals, motivation and boundaries	**MAN.03-BP1** Define the scope of work. Identify the ***Realistic & Measurable*** project's goals, motivation and boundaries	The scope of work according to releases, defined in the project plans was not aligned on SW commitment. 2- Project goals are not realistic given the current number of resources	Although **no new findings were detected** by the new BP & Rating was not impacted. Yet, **the gap became more descriptive & more detailed**
MAN.03-BP5 Define, monitor & adjust project estimates & resources. Define, monitor & adjust project estimates of effort & resources based on project's goals, project risks, motivation & boundaries	**MAN.03-BP5** Define, monitor & adjust project estimates & resources. Define, monitor & adjust project estimates of effort & resources based on project's goals, project risks & ***opportunities analysis***, motivation & boundaries	Feasibility of the project deliveries, is not demonstrated **2-Opportunities for the project are not assessed in the feasibility of the project deliveries**	**A new gap was introduced** for the evaluation of the opportunities. No impact to the project. **Yet can drive a process improvements in the organization which can in future impact ASPICE capability level 3, 4 & 5**
SYS.2-BP.8 Communicate agreed system requirements	**SYS.2-BP.8** Communicate agreed system requirements ***Note: Agreement on the work product could be by obtaining the customer approval on the work product***	As System architecture is not completely defined, consistency cannot be ensured, and communication to other teams	**New gap was detected with major impact to process area rating**
SYS.3-BP1 Specify functional and non-functional system requirements in a system requirements specification	**SYS.3-BP1** Specify functional, non-functional, ***Regulatory, Product Safety, Standards,*** system requirements in a system requirements specification	Regulations for different regions are not handled	**New gap was detected with major impact on the project**
SWE.1-BP.3	**SWE.1-BP.3**		Although **no new findings were**

(*continued*)

Table 1. (*continued*)

Current PAM wording	Proposed changes in the PAM	New weaknesses detected based on modified PAM	Conclusion
Analyze software requirements	Analyze software requirements, ***following a multidisciplinary approach (different Stakeholders) in requirements analysis to determine requirements feasibility***	1-Sw requirements attributes are not complete 2-Requirements check by different stakeholders is not completed during requirements elicitation process	**detected** by the new BP & Rating was not impacted. Yet, **the gap became more descriptive & more detailed. The proposal can lead to detection of new defects**
SUP.01-BP1 Develop a project quality assurance strategy	**SUP.01-BP1** Develop a project quality assurance strategy ***Define QA Plan with Prioritized QA activities based upon risk analysis, criticality of the process(es) and the frequency of QA activities***	Strategy used to determine frequency & priority of performing the defined quality activities is not defined or based on risk analysis for processes criticality in the project	**New gap was detected**. The **proposal highlights the importance of considering the risk analysis in the project planning in all process areas**
SUP.09-BP.6 Raise alert notifications	**SUP.09-BP.6** Raise alert notifications ***to relevant involved people as per defined project organization & communication scheme*** ***NOTE: ensure that the relevant people were informed***	Not all relevant involved parties are informed about the Sw project problems (Some Key stakeholders from management were not informed about resources issues needed for the project)	**New gap was detected**. The **proposal highlights the importance of project's internal communications especially with multisite teams**
SUP.09-BP.7 Initiate problem resolution	**SUP.09-BP.7** Initiate problem resolution ***with suitable possible ways to deal with non-conformances, ex: correction; segregation,***	No authorization was received about some of the deviations/problems	**New gap was detected**. The **proposal highlights the importance of project's internal communications especially with multisite teams**

(*continued*)

Table 1. (*continued*)

Current PAM wording	Proposed changes in the PAM	New weaknesses detected based on modified PAM	Conclusion
	containment, informing the customer; obtaining authorization for acceptance		
SUP.10-BP.1 Develop a change request management strategy	**SUP.10-BP.1** Develop a change request management strategy ***NOTE: Clarify the sources of the change (internal, customer, external supplier)***	No definition for the sources of the change request in the project strategy on sub system level	**New gap was detected. Although the gap detected would not impact the process area or the rating**. Yet, **It can drive further process improvements in the organization which can in future impact ASPICE capability level 3, 4 & 5**

4 Study Results

According to the applied scope of the improved (proposed) ASPICE Practices, the study results showed how can such approach as suggested by EUROSPI Manifesto [7] (to have improved assessment models which in our case was ASPICE PAM V3.1 modified with some of the IATF 16949-2016 requirements and ISO9001-2015 requirements) be useful for projects qualification in the organizations.

This approach helps organizations to be more efficient in different terms such as but not least:

- Budget: as the cost of detecting and fixing non-compliances in early phases would be cheaper.
- Resources: as some of the detected gaps were related to strategy definition, which would involve fewer resources to perform unlike if detected late in the project that would involve more stakeholders to consider the needed updates chain.
- Planning/Scheduling: as covering needs from the projects start ensure less issues rising and hence less impact to projects' plans or schedules.

Although the case study carried out as demonstrated in this paper, was unable to continue the observation for the action plan definition that would target the detected weaknesses elimination or even for the gaps closure verification due to that the case study was only allowed to be carried out on an imaginary project (specially created project for purpose of the study). Yet It provides a clear overview to organizations management about the empowerment that could be added to the assessments carried out in the future. Also, it gives stakeholders in all levels an overview regarding some of the needed weaknesses to consider in software projects to ensure full compliance to IATF 16949-2016 requirements with right fist time mindset with zero defects target approach.

5 Conclusion and Recommendations for Future Work

The case study results obtained demonstrated the importance for organizations to adapt process improvements methodology to empower and increase the efficiency of day-to-day operations (Which is "internal assessment process" this paper). Not only to be fully abide by guidelines and process models like in Automotive SPICE V3.1 PAM [1], but also to have advanced processes with more superseding requirements that can be derived from multiple sources (like IATF 16949-2016 [2], ISO 9001-2015 [4], etc.) to target operations excellence in different domains in the organization itself.

As a future step, A new case study is recommended to be carried out on an ongoing project. Also, Case study is recommended to be carried out on full project assessment scope to study for weaknesses observations over other process attributes that were not applicable to cover in this paper. Case Study shall observe also the whole improvement cycle according to Plan-Do-Check-Act concept which can be translated into terms of observing action plan definition, actions closure and verification and finally the lessons learned dissemination across the organization in terms of standardization of good practices in organization Quality management System through strategies, standards, processes, guidelines, templates, etc.… to ensure sustainability of value added from this activity that can be obtained by the organization.

Another point to consider in future studies is the ISO/IEC TS 33073:2017 (Information technology-Process assessment-Process capability assessment model for quality management), which defines an integrated PRM (Process Reference Model) and PAM (Process Assessment Model) that supports the performance of an assessment by providing indicators for guidance on the interpretation of the process purposes and outcomes and the process attributes.

References

1. Automotive SPICE® Process Reference Model, Process Assessment Model Version 3.1, 1 November 2017. http://www.automotivespice.com/
2. The Automotive Quality Management System Standard, known as IATF 16949:2016, 1 October 2016
3. ISO/TS IS 16949:2009, Quality management systems - Particular requirements for the application of ISO 9001:2008 for automotive production and relevant service part organizations, International Organization for Standardization, June 2009
4. ISO 9001:2015, Quality management systems Requirements, September 2015
5. ISO/IEC 12207:2008, Systems and software engineering – Software life cycle processes (2008)

6. CMMI for development V3.1
7. SPI MANIFESTO Version A.1.2 (2010)
8. IATF 16949 global website. www.iatfglobalozsversight.org. Accessed Jan 2019
9. Automotive SPICE pocket guide version 3 by Kuglermaag, January 2019
10. Sabar, S.: Software process improvement and lifecycle models in automotive industry. Final thesis, Linkopings Universitet, Sweden, June 2011
11. Peldzius, S., Ragaisis, S.: Comparison of maturity levels in CMMI–DEV and ISO/IEC 15504. In: Proceedings of the "Applications of Mathematics and Computer Engineering" (CEMATH 2011) Conference. ISBN 978-960-474-270-7
12. Hoermann, K., Mueller, M., Dittmann, L., Zimmer, J.: Automotive SPICE in Practice: Surviving Implementation and Assessment, Rocky Noor (2008). ISBN 978--1933952291
13. Sassenburg, H., Kitson, D.: A Comparative Analysis of CMMI and Automotive SPICE, Presentation, June 2006
14. https://sites.google.com/a/valeo.com/automotive-quality-standards/. Accessed Jan 2019
15. IATF 16949 Gap analysis Tool by AIAG, January 2019
16. Process Assessment Model SPICE for Mechanical Engineering V1.6 by INTACS
17. Oliveira, P., Ferreira, André L., Dias, D., Pereira, T., Monteiro, P., Machado, R.J.: An analysis of the commonality and differences between ASPICE and ISO26262 in the context of software development. In: Stolfa, J., Stolfa, S., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2017. CCIS, vol. 748, pp. 216–227. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64218-5_17
18. Makkar, S.S., Abdelfatah, A.M., Yousef, A.H.: Automotive standards compliance cost reduction by mutual integration between automotive SPICE and IATF 16949:2016. In: 2019 IEEE International Conference on Vehicular Electronics and Safety (ICVES) (2019)

An Interpretation and Implementation of Automotive Hardware SPICE

Jaka Ivančič[1], Andreas Riel[2,3](✉), and Damjan Ekert[2]

[1] Kolektor Group s.o.o, 5280 Idrija, Slovenia
jaka.ivancic@kolektor.com
[2] ISCN GmbH, 8010 Graz, Austria
andreas.riel@grenoble-inp.fr, {ariel,dekert}@iscn.com
[3] Grenoble Alps University, Grenoble INP, G-SCOP, CNRS, 38031 Grenoble, France

Abstract. In the context of the Automotive SPICE® standard process reference and assessment models, many works have been published about the standard's possible and actual interpretations in several industrial environments and applications. Very few of them, however, deal with the question of what detailed electronic hardware development practices can or shall achieve, and in which order they should be performed. Particularly in safety-relevant systems, however, hardware design quality is vital. In this paper, we present an interpretation of Automotive Hardware SPICE as it is practiced within a global tier-2 supplier. We also show its compatibility with the Hardware SPICE process reference and assessment model plug-in released by the intacs™ at the end of 2019.

Keywords: Automotive SPICE® · Hardware development · Hardware design · Hardware validation and testing

1 Introduction and Related Work

Automotive SPICE® (ASPICE) is the worldwide standard for a process reference model (PRM) and process assessment model (PAM) for the development of automotive embedded systems [1]. Traditionally focused on automotive systems and embedded software development, the version 3.0 of the standard (released in July 2015) came up with a plug-in concept in order to open the path for extending the ASPICE process model with practices for hardware engineering (HWE) as well as mechanical engineering (MEE) [1]. In the automotive sector, formalizing and assessing hardware development processes and practices has become a key necessity with more and more automotive subsystems becoming safety-relevant and therefore subject to the ISO 26262 [2], the standard for handling the functional safety of electronics and software in road vehicles. The ISO 26262 relies on ASPICE processes being in place for both system, hardware and software development, and extends it most notably with specific clauses and methods defining the state-of-the-art of how to determine the safety integrity level of automotive items, as well as how to achieve them by design.

Hardware engineering has a vital role in this because functional safety, i.e., the absence of unreasonable risk due to malfunctions of electric/electronic subsystems [2],

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 684–695, 2020.
https://doi.org/10.1007/978-3-030-56441-4_51

starts at the electronic signal level. Software that has been tested successfully with a 100% test coverage will likely produce errors if the underlying hardware gets defect. This is why the intacs™ Working Group "Hardware Engineering Processes" released an initial version of the HWE plug-in process specification in November 2019 [3], strongly based on previous original works performed in the German SOQRATES initiative [4], as well as on [5]. Specific practical implementations of this process reference model extension to ASPICE have not been published so far, although automotive OEMs keep pushing their suppliers towards higher quality hardware development processes.

This article gives an insight into one of the worldwide first fully blown complete interpretation of the HWE process reference and assessment model extension. It elaborates on the overall model, as well as on the key characteristics of the individual processes, as well as their relationships and rationales.

2 Basic Definitions

The definition of fundamental hardware items in our process is based on the fundamental specification given in [3] and depicted in Fig. 1 below.

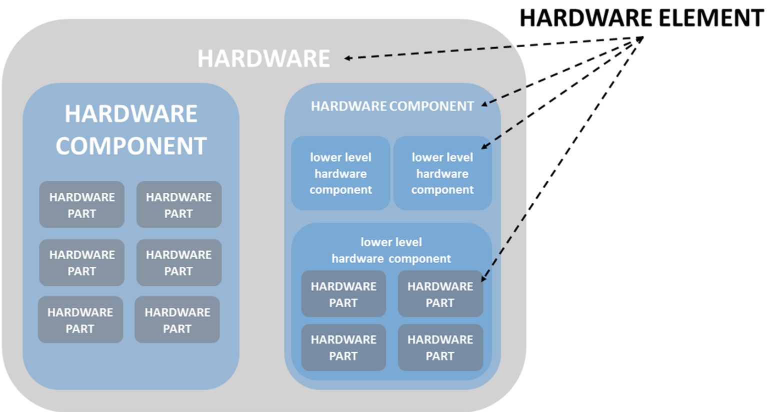


Fig. 1. Graphical specification of hardware elements according to [3]

In our process, "Hardware" is mapped to one or several electronic boards. "Hardware component" is mapped to the item "Hardware Assembly". "Lower level hardware component" is mapped to "Hardware Unit". "Hardware Parts" are considered the basic building blocks of hardware units and are therefore not represented as a dedicated work item. In compliance with the ISO 26262, in our process, a hardware unit is defined as the smallest block of hardware to which a specific hardware subfunction can be assigned. Hardware units are composed of hardware parts, the lowest (indivisible in the application scope) level of hardware (e.g. resistors, capacitors, IC's). Figure 2 shows a power supply hardware unit whose function is to convert

230 V input voltage to a circuit supply of 9 V. This includes different hardware parts and integrates them to a hardware unit implementing a function which can be mapped onto unit testing including e.g. an equivalence class test: Testing the output above 1 A, the fuses F1 shall fire, and there shall be no voltage output. Testing in the 0,8 A ± 0,2 V range, the system operates and delivers current at 9 V. Hardware units can also include diagnose outputs such as a connection to an ADC that writes the measured voltage to a register.

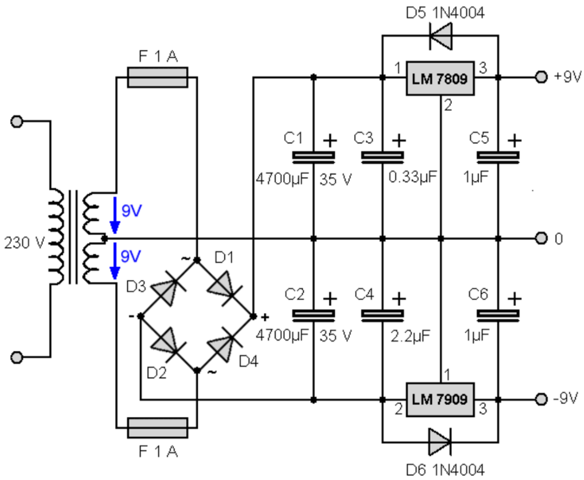


Fig. 2. Example of a hardware unit (detailed schematic design)

Figure 3 depicts the workflow we defined for the *Hardware Unit* work item. The explicit distinction between the schematic design and the layout design for the specification of the hardware unit's detailed design is of particular importance here.

A hardware functional block is defined as a set of hardware units that are related to each other in that they implement a higher functionality on integrated hardware level (e.g., power supply, input-circuit, filtering, micro-controller, driver circuits, etc.). Another example is a power stage for e-motor control which is typically composed of several hardware units implementing functions needed for the power stage (e.g., charge pump, current sensing, temperature sensing). Hardware functional block definitions shall cover the specification of any required hardware units as black boxes with their interfaces and interconnections. In particular, the specification shall cover the hardware block's interface to other functional blocks and units. In Polarion, hardware functional blocks are defined in work items of subtype *HW Functional Block*. The workflow and attributes are the same as for the work item *Hardware Unit*.

A hardware PCB layout represents the integration of any or a particular set of hardware functional blocks and units to a functional hardware circuit. In Polarion, PCBs are represented as work items of type *PCB* with the associated workflow depicted in Fig. 4. A hardware release work item represents the integration of one or

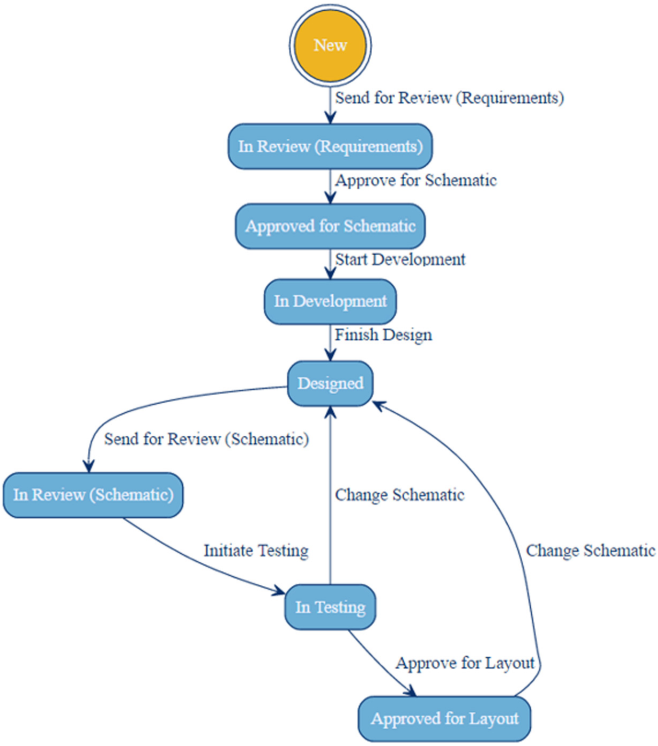


Fig. 3. The hardware unit item workflow

several hardware PCB's and hardware units. The release work item further serves as an interface between teams from development, production, and testing.

The traceability between the different hardware specification levels is facilitated by the Polarion link model and the Polarion work items *Hardware Requirement*, *Hardware Item (subtype Unit or Functional Block), Hardware Release*, and *PCB*, see Fig. 5 below.

Each of these work items includes all the standard field and attributes we defined for system-, software-, hardware-, as well as mechanics-related work items to be managed during their development life cycles in an ASPICE compliant way. Since there are no hardware-related specifics, we will not elaborate on them in this article.

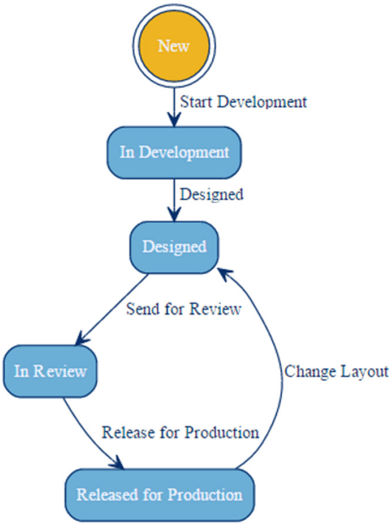


Fig. 4. The PCB item workflow

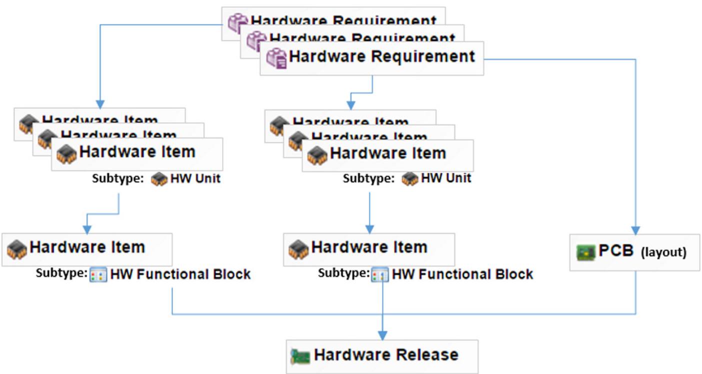


Fig. 5. The hardware work item structure and traceability model in Polarion

3 Hardware Development Process Design

Hardware engineering requires the integration of design, manufacturing and design verification in a way that is different from software engineering (as software does not have to be "manufactured"). In particular, design steps are divided into schematic and layout design, with static and dynamic verification in a simulation environment to be done after each of these two major design steps. The layout design covers both hardware units and hardware functional blocks. Its verification is partly done in a simulation environment before the PCB is actually manufactured and ultimately verified by physical tests.

More specifically, and as shown in Fig. 6, hardware design up to design verification is executed in the following step sequence, and over iterations as required:

1. Create hardware architectural design in the form of a hardware schematic design (of hardware functional blocks).
2. Create hardware detailed design in the form of hardware schematic designs (of hardware units).
3. Perform static/dynamic verification of the hardware schematic designs (schematic design verification through simulation).
4. Create the layout design including any hardware units and hardware architectural design elements (functional blocks).
5. Send the layout design to production (manufacture PCB with all hardware units and functional blocks).

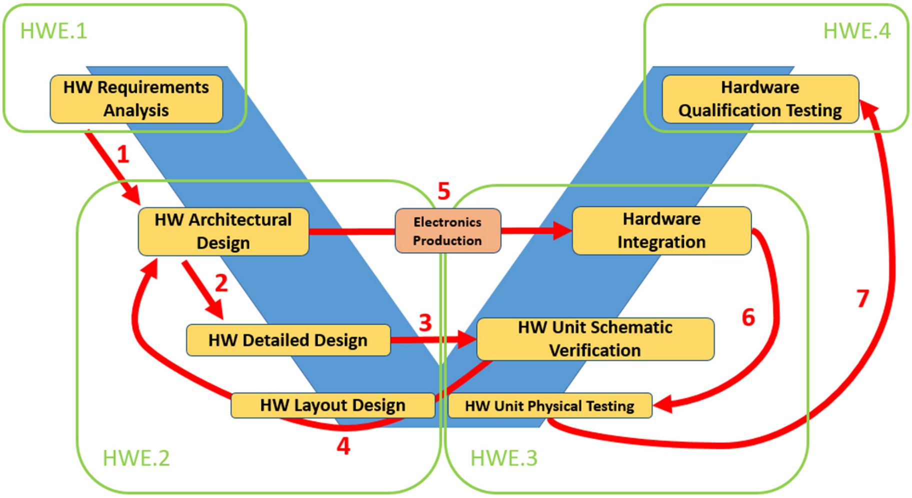


Fig. 6. Integrated hardware design, manufacturing and design verification strategy

6. Perform hardware unit physical tests on the fabricated PCB according to unit requirements.
7. Perform physical hardware qualification testing on the PCB according to hardware requirements.

This sequence of steps clearly shows some specific interpretations of the HWE V-cycle plug-in specified in [3] that we have made:

1. HWE.1 (HW Requirements Analysis) is covered by HW Requirements Analysis.
2. HWE.2 (HW Design) is covered by HW Architectural Design, HW Detailed (Schematic) Design, and HW Layout Design.
3. HWE.3 (Verification against HW Design) is covered by HW Unit Schematic Verification, HW Integration, HW Physical Testing.

4. HWE.4 (Verification against HW Requirements) is covered by Hardware Qualification Testing.

In the following subsections, we will describe the key sub-steps and activities of each of these steps, focusing on those that differ most from related software development process steps.

3.1 HW Requirements Analysis

The purpose of the Hardware Requirements Analysis process is to derive hardware requirements from the system requirements and system architectural design specifications, hardware design standards, as well as internal hardware requirements specifications. Hardware requirements shall define the functional hardware behavior in terms of hardware functions as well as non-functional hardware properties in a solution-neutral way. These hardware requirements subsequently guide the architectural and detailed design of the hardware. We will not describe related process steps and actions further because there is nothing really special compared to software requirements analysis.

3.2 HW Architectural Design

The purpose of the Hardware Architectural Design is to establish the highest level architectural design specification of the hardware to be developed according to the hardware requirements specification. The process identifies which hardware requirements are to be allocated to which hardware items and evaluates the according Hardware Architectural Design against defined criteria. Hardware unit requirements arising from Hardware Architectural Design are mapped to the Hardware Detailed Design level.

The hardware architectural design shall cover the mapping from functional and non-functional requirements to design decisions and hardware items in terms of hardware units and functional blocks. Moreover, it shall specify the internal and external unit and functional block interfaces as well as their interconnections in a way that hardware developers have a very clear and unambiguous specification for the unit detailed design and construction. The specification shall cover both static and dynamic aspects of the hardware architecture as well as specific constraints, solutions and requirements from standards, application guidelines and risk mitigation activities (e.g. dFEMA). Consequently, the following steps are in the scope of hardware architectural design activities:

1. evaluate possible hardware architectural design approaches and justify the choice of the chosen approach,
2. establish a static architecture diagram showing all hardware units and components (set of units) and their signal interfaces and connections,
3. create a hardware unit/functional block work items in Polarion having names corresponding to the names in the diagram and fill out all mandatory attributes,
4. establish dynamic diagrams as considered relevant,

5. consider objectives for resource consumption, reliability, safety, etc. on an architectural level,
6. establish general hardware design decisions as work items of subtype Design Decision in the hardware requirements specification,
7. define relevant criteria for the static hardware verification (e.g., FIT-rates, layout metrics, voltage and temperature levels, etc.),
8. review all the created work products.

Hardware requirements and conceptual architectural design is a basis for static hardware architecture, that consists of:

- identification of hardware units and functional blocks,
- unit functional, non-functional and dynamic requirement specification,
- unit verification criteria,
- interfaces between units and blocks,
- external interfaces,
- applicable PCB layout requirements, mechanical properties or constraints (e.g. positioning of power stage).

Static architecture is given in a graphical representation of units and additional hardware requirements (decisions) that can either become a part of hardware requirements specification or unit requirements specification.

3.3 HW Detailed Design

The purpose of the Hardware Detailed Design process is to establish a detailed design of the hardware units such that they implement and are consistent with the requirements and architectural design. The detailed design specification of these hardware units in terms a documented schematic design is the lowest specification level and therefore provides the basis for the hardware unit construction by developing the hardware layout design. The hardware detailed design and schematic design must implement the hardware requirements in compliance with the hardware architectural design (unit requirements). The hardware unit construction must implement the detailed design in the layout design on the basis of the schematic designs of hardware units, functional blocks, and their interconnections. A standard set of attributes and links are used to support analysis.

The Hardware Detailed Design shall cover the schematic design of the entire hardware based on the hardware architectural design, as well as the hardware requirements. Consequently, the following steps are in the scope of hardware detailed design activities:

1. evaluate hardware schematic design approaches and justify the choice of the chosen approach,
2. establish detailed schematic designs of hardware components, including a detailed specification of component interfaces, as well as the component's units and their interconnections,
3. establish detailed schematic designs of hardware units, including a detailed specification if the unit interfaces and any relevant non-functional static and dynamic

properties (in compliance with related design objectives, e.g. resource consumption, timing behaviour, etc.),
4. establish detailed schematic designs of hardware PCB's in compliance with the hardware architectural design and in full consistency and traceability (by name) to the hardware units and functional blocks,
5. review and release the schematic designs for unit schematic verification.

Schematic design guidelines define which hardware part libraries are allowed, give instructions for creating new components, specify design patterns and rules, as well as naming conventions.

3.4 HW Layout Design

To purpose of the Hardware Layout Design is to develop the hardware units as a hardware layout covering all the units, functional blocks as well as their interconnections on one or several PCB. The hardware layout design covers the entire PCB layout including any unit and functional block. Consequently, the following steps are in the scope of the hardware unit construction activities:

1. per PCB, establish layout designs in compliance with the respective schematic designs for any hardware unit and functional block that shall be integrated on this PCB,
2. apply real-time design rule checking,
3. release the PCB layout design for manufacturing and physical testing.

Just like with the schematic design, PCB layout design guidelines shall assure coherent design practices and decisions throughout the organisation.

3.5 HW Unit Schematic Verification

In order to assure that the design decisions made during the schematic design are correct and to avoid any defects propagating to the layout design and the physical unit and hardware assembly construction, the unit verification steps as described below shall be carried out based on the schematic design and before the PCB layout design. The hardware unit verification shall cover all the hardware units and components assuring their functional completeness (i.e., requirements coverage), correctness (i.e., correct implementation of the requirements before the layout design) and compliance with defined verification criteria. The verification is typically performed by simulation. Consequently, the following steps are in the scope of the hardware unit verification of *functional completeness*:

1. Define test cases with regard to unit requirements and verification criteria resulting in detailed description of test steps and pass fail criteria,
2. set up the simulation environment for test cases,
3. specify tests in the simulation environment,
4. execute these tests,
5. record and analyze the test results,
6. improve the tested units and test cases if required,

7. release the schematic designs for layout design (i.e., unit construction).

These steps aim at verifying hardware units for *correctness* using the defined criteria for verification, and comprise the following activities:

- Review of the hardware unit design.
- Review of the hardware unit specification sheet.
- Review of the hardware unit FIT rate target.
- Integration of the FIT rate and target profile of the hardware unit, including its diagnose capability, in the FMEDA [2, 6, 7] and FTA [8].
- Analysis of the consistency of the assigned ASIL and the determined FIT rate.

Based on this, testing activities comprise circuit simulation in recognized tools, most notably LTspice (by Analog Devices), and based on electronic component simulation models (typically provided by the component manufacturer). Simulation test cases shall include at least the following cases:

- Voltage sweep (verifying operating voltage, continuous and transient voltage levels across the entire circuit),
- Temperature sweep (verifying temperature ranges across the entire circuit),
- Power range check,
- Tolerance calculation (using worst-case analysis, Monte Carlo simulation, etc.).

3.6 HW Physical Testing

Physical hardware unit tests basically involve the same sequence of steps as the static testing. The required activities depend on the hardware item classification according to ISO 26262-5 [2], i.e., either class I (e.g. resistor, capacitor, transistor, diode, quartz, resonator), class II (e.g. fuel pressure sensor, temperature sensor, stand-alone ADC) or class III (e.g. microprocessor, microcontroller, DSP, accelerator) element depending on the item's properties. These classes reflect the difficulty of the verification of the safety-related functionality and the role of the hardware element within the safety concept in safety-relevant systems.

- Class I, II, III: Tables 10–12 in ISO 26262-5 apply.
- Class II and III:
  - Qualification of hardware units for their expected target usage profile.
  - Execution and documentation of the lab tests - a set of electrical tests with measurement points on the ECU layout must be planned to check the function of the hardware unit and its behavior in case of failure conditions. In case of processors performing a low level HIL test with the processor pins as the interface.

It is common that physical unit tests cannot be performed without some level of integration with software or other functional blocks. To overcome this, a higher level testing is allowed as long as the traceability to the unit, as well as a complete unit requirements coverage is assured and maintained.

3.7 HW Integration

The purpose of the Hardware Integration is to integrate and produce (interface to the manufacturing process) the complete hardware following a defined hardware integration checklist to assure the integration of the correct version of all the hardware elements (including required configurations) and to thereby create a defined and documented hardware release. The purpose of the Hardware Integration Test is to verify critical communication interfaces and functional blocks of the produced hardware based on the specified hardware architecture. We will not further elaborate on this process here, since we feel that there are no practices or activities to be particularly highlighted.

3.8 HW Qualification Testing

The purpose of the Hardware Qualification Testing process is to ensure that the integrated hardware is tested to provide evidence for compliance with the hardware requirements and that the hardware is ready for integration into the system. The only thing we want to highlight here is the importance of achieving full coverage of customer-specific hardware qualification/acceptance tests.

Similar to HW Unit Physical tests, it is common that HW Qualification tests are performed on a higher integration level (System Integration or System Qualification) as the testing is not possible without embedded software unless the stimulus is generated by a test setup (e.g. HIL).

4 Conclusion and Outlook

We presented a company-specific interpretation of the Hardware Engineering PRM/PAM of Automotive SPICE. Although for reasons of confidentiality we did not elaborate on the content of the individual processes, we have shown that hardware engineering requires a V-cycle implementation that is different from software engineering. The key differences are in the importance of schematic design for static and dynamic hardware verification, as well as the need of hardware integration on a PCB before the execution of tests on the real hardware from unit test level to qualification test level. This implies a special sequence of running through the V-cycle, as well as clearly defined interface to hardware production on hardware integration level. All the processes along the hardware V-cycle have been defined in detail in terms of process steps, associated roles, activities, and interdependencies. These specifications include special clauses and design patterns for safety-relevant projects that are aligned with the ISO 26262 [7].

The presented process has been applied successfully to several projects and is expected to undergo an ASPICE assessment at the end of 2020. Its key impacts and benefits for our company are

1. an overall improvement of hardware design and development, as well as the quality of related work products;

2. more efficient and complete testing due to well defined requirement-test case pairs complemented with a joint review and common understanding between designer and tester;
3. earlier systematic design error prevention and recognition thanks to a clear link between system requirements, hardware requirements and corresponding design elements (schematics, layout);
4. a more narrowed and comprehensive change, risk and problem resolution management during the implementation of HW units;
5. re-use and cross-project information exchange of hardware requirements, design solutions, test cases as well as information on previous issues, risks and important changes.

In the near future, priority will be given to the improvement of the schematic and layout design guidelines in order to better capitalize on proven design patterns and move towards modular kits and libraries for facilitating variant management. Furthermore, the integration of cybersecurity related practices and aspects is envisaged [9, 10].

References

1. VDA QMC: Automotive SPICE® Process Assessment/Reference Model, Version 3.0 (2017)
2. ISO - International Organization for Standardization: ISO 26262 Road vehicles Functional Safety Part 1-10 (2011)
3. intacs™ Working Group 'HW Engineering Processes': PRM/PAM Hardware Engineering (HWE), Version 1.0 (2019)
4. Schlager, C., Messnarz, R., Sporer, H., Riess, A., Mayer, R., Bernhardt, S.: Hardware SPICE extension for automotive SPICE 3.1. In: Larrucea, X., Santamaria, I., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2018. CCIS, vol. 896, pp. 480–491. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-97925-0_41
5. Sporer, H.: Mechatronic system development: an automotive industry approach for small teams. Ph.D. Dissertation at Grez University of Technology (2016)
6. ISO - International Organization for Standardization: IEC 60812 Analysis techniques for system reliability - Procedure for failure mode and effects analysis (FMEA) (2006)
7. Messnarz, R., Kreiner, C., Riel, A., et al.: Implementing functional safety standards (SafEUr). ASQ Softw. Qual. Prof. **17**(3), 4 (2015)
8. ISO - International Organization for Standardization: IEC 61025 Fault tree analysis (FTA) (2006)
9. Messnarz, R., Kreiner, C., Riel, A.: Integrating automotive SPICE, functional safety, and cybersecurity concepts: a cybersecurity layer model. Softw. Qual. Prof. **18**(4), 13 (2016)
10. Macher, G., Sporer, H., Brenner, E., Kreiner, C.: Supporting cyber-security based on hardware-software interface definition. In: Kreiner, C., O'Connor, R.V., Poth, A., Messnarz, R. (eds.) EuroSPI 2016. CCIS, vol. 633, pp. 148–159. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-44817-6_12

A Barbell Strategy-oriented Regulatory Framework and Compliance Management

Barbara Gallina(✉)

IDT, Mälardalen University, Box 883, 721 23 Västerås, Sweden
barbara.gallina@mdh.se

Abstract. As we are witnessing nowadays with self-driving vehicles and robots performing collaborative tasks, and as it has been witnessed in the past, a disruptive technological innovation may cause a regulatory disruption. As a consequence, established regulatory frameworks need to evolve and, in this evolution process, regulatory excellence is paramount for guaranteeing appropriate risk taking, enabling innovators to innovate while efficiently complying to proportionate regulatory frameworks. In risk management theory, or better, in risk taking theory, Taleb has proposed the Barbell strategy, characterised by maximal certainty/low risk on one extreme of the barbell, and maximal uncertainty on the other extreme, as a means for developing anti-fragility, i.e., ability to gain from disorder. In this paper, we propose our vision consisting of a Barbell strategy-oriented regulatory framework and compliance management, where regulators and manufacturers/innovators, both, adopt a Barbell strategy for reaching an optimal solution in regulatory making, the former, and compliance management, the latter. Finally, conclusion and perspectives for future work are also drawn.

Keywords: Disruptive technological innovation · Regulatory framework · Compliance · Regulatory disruption · Barbell strategy

1 Introduction

As we are witnessing with nowadays disruptive innovations, e.g., with self-driving vehicles, but also with the development and adoption of robots capable of performing collaborative tasks, in which they work side-by-side (located within the same safeguarded space) with human operator(s) rather than replacing them, and as it has been witnessed in the past, a disruptive technological innovation may cause a regulatory disruption. When regulatory disruption occurs, established regulatory frameworks need to evolve. During such evolution process, regulatory excellence is paramount for guaranteeing appropriate risk taking, enabling innovators to innovate while efficiently complying to proportionate regulatory frameworks.

As depicted in Fig. 1, during such evolution process, two actors play a crucial role: regulators/assessors/standardisation bodies on one side and manufacturers/suppliers on the other side. Both actors analyse the disruptive innovations

Partially funded by Sweden's Knowledge Foundation via the SACSys [31] project.

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 696–705, 2020.
https://doi.org/10.1007/978-3-030-56441-4_52

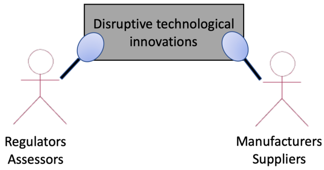


Fig. 1. Actors analysing disruptive innovations.

to contribute to designing the needed changes for its acceptable deployment within society. As Fig. 1 depicts, the two actors compete but also cooperate for the identification of a proportionate regulatory framework.

In risk management theory, or better, in risk taking theory, Taleb [34] has proposed the Barbell strategy, characterised by maximal certainty/low risk on one extreme of the barbell, and maximal uncertainty on the other extreme, as a means for developing anti-fragility, i.e., ability to gain from disorder. In this paper, we propose our vision for handling the impact of disruptive technological innovations on regulatory frameworks and compliance management. Our vision consists of a Barbell strategy-oriented regulatory framework and compliance, where regulators and manufacturers, both, adopt a Barbell strategy for reaching an optimal solution in regulatory making (the former) and compliance management (the latter). Then, we sketch a roadmap for future research. The rest of the paper is organised as follows. In Sect. 2, we provide essential background information. In Sect. 3, we propose our vision. In Sect. 4, we explain the synergy with the SPI Manifesto. In Sect. 6, we briefly discuss related work. In Sect. 6, we propose our intended research agenda. Finally, in Sect. 7, we present our concluding remarks.

2 Background

In this section, we present essential background information on regulatory frameworks and excellence, regulatory compliance management, disruptive technological innovations and regulatory disruption.

2.1 Regulatory Frameworks and Excellence

Regulators need to adopt a pro-innovation culture, i.e., rather than adopting an inflexible and pernicious "by-the-book" mentality, regulators are expected to design regulatory processes which mirror the phases of growth of technological innovations. In [8], key facets of a regulator's organisation and actions that affect its performance are identified. These key facets are internal management; external engagement; priority-setting/decision-making; and problem- solving. In the context of this paper, we recall:

– "Priority-setting/decision-making. An excellent regulator should rely consistently on careful, evidence-based decision making and should set priorities informed by consideration of risks. An excellent regulator should also be clear about which policy principles it uses when choosing how to prioritise different risks and make decisions.

– Problem-solving. Flexible regulatory approaches promise more cost-effective outcomes, as they give regulated entities the opportunity to choose lower-cost means of achieving regulatory outcomes. [...]. But flexible instruments will not work well under all circumstances. The ultimate test for problem-solving lies in finding the right tool for the purposes and circumstances at hand."

2.2 Regulatory Compliance Management

In this subsection, we recall fundamental definitions concerning regulatory compliance management.

Compliance - this term refers to conformance to a rule, such as a specification, policy, standard or law [28].

Regulation - this term refers to the diverse set of instruments by which governments establish requirements for enterprises and citizens. Regulations include laws, formal and informal orders, subordinate rules issued by all levels of government, and rules issued by non-governmental or self-regulatory bodies to whom governments have delegated regulatory powers [28].

Regulatory Compliance - this term refers to obedience by the target population(s) with regulation(s) [28].

Compliance Management - refer to the consistency between the actual development process and the normative reference model embedded in the standard [10]. As presented in [14], compliance management can be handled in various ways. In the reminder of this section, we recall automated compliance by design (compliance checking) and argumentation.

Compliance by Design - approach that provides the capability of capturing compliance requirements through a generic modelling framework, and subsequently facilitate the propagation of these requirements into the process models [32]. The automation of this approach (automated compliance by design, i.e., compliance checking) has been designed [18] and implemented [17] and demonstrated for business processes. Compliance by design has also been further developed and demonstrated in the context of safety-critical systems engineering [2,4–7]. A personal opinion survey about its relevance has also been conducted [3].

Argumentation about compliance consists in formulating an argument for justifying the claims about compliance. The argument shows the reasoning that connects the claims with the supporting evidence. Automatic generation of arguments is an active research topic (the reader may refer to [11–13,35] in case of interest in automatic generation of arguments for process compliance). It however requires stable and broadly accepted reasoning strategies. In the absence of broadly accepted reasoning strategies, arguments are typically the result of manual elaboration.

2.3 Disruptive Technological Innovations and Regulatory Disruption

Certain technological innovations can undermine existing products, firms, or even entire industries. When this occur, they are labelled as disruptive. As a consequence of disruptive technological innovations, a regulatory disruption may follow. Specifically, this may happen when the innovation may fall within an agency's jurisdiction but not square well with the agency's existing regulatory framework [9]. Regulatory disruption occurs when the "disruptee" is the regulatory framework itself [9]. The introduction of Internet is an example of technological innovation, which caused a regulatory disruption. Nowadays, we are witnessing important technological disruptions. In the automotive domain, the autonomous driving is causing a regulatory disruption and is calling for an urgent solution [1]. A set of standards is being planned [26], while UL 4600 [36] has recently reached a consensus for publication. The key question is understanding what functional safety may mean in case of self-driving vehicles, which, due to the current limitations of machine learning-based algorithmic solutions, may experience black swans [33]. Another key question is understanding how assurance shall be provided [1]. Similar situation can be identified in the context of robots performing collaborative tasks in the domain of collaborative robotics. By 2020, revised versions of ISO 10218-1 [25] and ISO 10218-2 [24] are expected to be published. These revised versions, which build on top of the lessons learnt drawn via the application of the ISO technical specification ISO/TS 15066:2016 [23], define the safety requirements for the sphere of robots expected to perform collaborative tasks (named in non-standardised terminology as COBOTS, collaborative robots). As pointed in [20], besides the robot itself, the robot in this context includes the end effector, i.e. the tool adapted on the robot arm with which the robot performs tasks, and the objects moved by it. Close or direct contact between the robot and the human operator gives rise by definition to a possibility of collision. The robot manufacturer's risk assessment must therefore also cover the intended industrial workplace. During the transition period triggered by the occurrence of a disruption, as discussed in [9], to face regulatory disruptions, regulators have various means at disposal: inflexible policy making, threats, sunsets. In the first case, regulators persist adopting traditional policy making and this turns out to be premature. In the second case, regulators (representative of a regulatory agency) might announce via guidance document, warning letter, or press release that they will take action against companies that employ novel technologies or business practices in a certain way. Experience has shown that inflexibility and flexibility via threats are both non viable solutions. Regulators need to decide upon: timing (when to intervene?), form (how to intervene? via adjudication? via guidance? new forms?), durability (how long should the intervention endure? should it be permanent? volatile? e.g., by using sunsets, temporary legislations with finite durations?), and enforcement (how rigorous should be the monitoring of the compliance?). Experience has shown that regulators shall opt for sunset that are binding and enforceable. Sunsets shall be deliberated as early as possible to let them act as stopgap. Sunsets possibly shall be wrapped as regulations or as guidance, practically perceived as binding.

2.4 Taleb's Barbell Strategy

In his book on anti-fragility (property that characterise systems able to gain from disorder, including disorder generated by the occurrence of black swans [33]), Taleb mentions the Barbell-strategy as a means for developing anti-fragility. The Barbell strategy consists of a bimodal strategy, a combination of two extremes, one safe and one speculative. Such bimodal strategy is deemed more robust than a "mono-modal" strategy. Taleb exemplifies the strategy in the financial domain and then illustrates its application in various domains. In the financial domain, a barbell strategy is characterised by maximal certainty/low risk in one set of holdings, maximal uncertainty in another. More precisely, it consists of the mixing of two extreme properties in a portfolio such as a linear combination of maximal conservatism for a fraction of the portfolio, on one hand and maximal (or high) risk on the remaining fraction. In [15], the Barbell strategy is formalised and it is demonstrated that it turns out to be an optimal solution.

3 A Barbell Strategy-Oriented Vision

In this section, we present our vision for an optimal solution consisting of a barbell strategy-oriented regulatory framework and compliance.

3.1 A Barbell Strategy-Oriented Regulatory Framework

As shown in Fig. 1, both actors, regulators/assessors/standardisation bodies and manufacturers/innovators cooperate/compete to identify a proportionate regulatory framework and compliance management. Based on what recalled in the background, to face disruptions, sunset-based enforceable regulations seem to be the appropriate regulatory tool to face unknown. Thus, in our vision, regulators/assessors/standardisation bodies shall cooperate to propose a barbell strategy aimed at facing known knowns (i.e., innovation-related risk that we know that we know) via traditional portions of regulatory frameworks and facing unknown (i.e., innovation-related risk that we do not know we do not know, the kernel of the disruption) with subset-based regulations or binding guidance. This vision is depicted in Fig. 2.

As depicted in Fig. 3, our vision can shed lights on the current development in the automotive and robotics domains. In the automotive domain, road vehicles that have a degree of automation that still squares well with the agency's existing regulatory framework need to comply ISO 26262 [21]; while road vehicles that have a disruptive degree of automation are expected to comply with UL4600, a binding guidance, result of a consensus-based deliberation process.

In a similar manner, in the robotics domain, robots that have a zero-degree of cooperation (are not located within the same safeguarded space as human operators/are not expected to perform collaborative tasks) still squares well with the agency's existing regulatory framework and need to comply with ISO 10218-1:2011 [22]; while robots that have a disruptive degree of cooperation (are

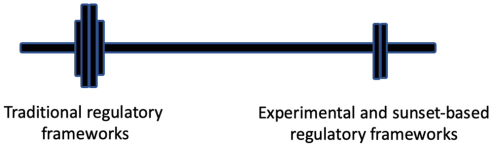


Fig. 2. Barbell strategy-oriented Regulatory Framework.

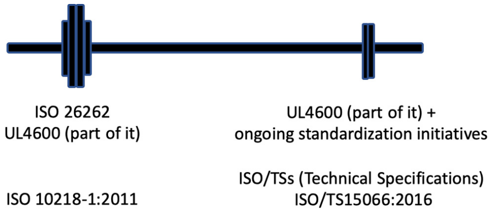


Fig. 3. Regulatory framework vision for robotics and automotive.

located within the same safeguarded space as human operators/are expected to perform collaborative tasks) are expected to comply with ISO/TS 15066 [23] and later on with the currently under development ISO 10218-1 [25] and ISO 10218-2 [24], a binding guidance result of a consensus-based deliberation process. These dual strategies are aligned with our vision.

3.2 A Barbell Strategy-Oriented Compliance Management

As recalled in Sect. 2, compliance can be managed via different approaches with different degrees of automation. On one hand, compliance by design is time-consuming given the considerable effort required during the formalisation process. However, this solution seems promising in terms of return on investment especially within stable regulatory frameworks, where rules are known and also the corresponding allowed interpretations. On the other hand, argumentation, mainly formulated in natural language, is mainly conducted manually. Argumentation offers a possibility for negotiating via the elaboration of compelling arguments which may go beyond traditional ways of doing. Thus, our Barbell strategy-oriented compliance management vision is depicted in Fig. 4.

3.3 Discussion

Our vision based on a double Barbell strategy applied to regulatory frameworks and compliance management has the potential to increment a dialogue between the parties and ease the exploitation of the best practices allowing parties taking advantage of the opportunities inherent proven/mature best practices, while actively searching for solution on how to handle the disruption.

Fig. 4. Barbell strategy-oriented Compliance Management.

4 Synergy with the SPI Manifesto

The SPI Manifesto targets software. Our vision is not limited to software, it embraces disruptive innovations in general. Software process improvement is highly connected to the regulatory framework, which may constrain it. In our vision, we do not only focus on the improvement of the process adopted by the innovator(s) to develop their software but also on the process adopted by the regulator. Our vision could be seen as a generalisation of the SPI Manifesto to a PI. Specifically, we see that our Barbell strategy-oriented vision pursues a bimodal strategy that preserves useful portions of regulatory frameworks and discard what is not applicable anymore, replacing it with risky but temporary and dynamic solutions. Thus, it can be seen in synergy with principle 6 of the SPI Manifesto [29], which states "use dynamic and adaptable models as needed". As quoted in Sect. 6.1 of the SPI Manifesto, "All models are wrong - some are useful". The best practices can be replaced by new practices, recognised as potential new best practices or pioneering practices. In addition, our vision is strongly related with principle 7 of the SPI Manifesto, which states "apply risk management". Our vision is an exploration of the risk management, which includes risk taking towards an optimal solution for handling disruptive innovations, in line with the recently proposed manifesto for antifragile software [30]. Finally, our vision is also strongly related with principle 8 of the SPI Manifesto, which states "Manage the organisational change in your improvement effort". A disruptive innovation represents a change.

5 Related Work

To the best of our knowledge, our work represents a novelty since no research so far has been conducted to elaborate a Barbell strategy-oriented perspective on regulatory frameworks and compliance management to handle disruptions. A conceptual exploration of a Barbell strategy has been conducted in the context of agile development and education. Specifically, in [19], it is proposed that the agile enterprise portfolio executes on proven business opportunities while actively embracing change and uncertainty. In [27], a Barbell strategy is explored in the context of Asian countries, where students aim for academic excellence, risking psychological diseases. The strategy allows for maintaining a trajectory of academic excellence while avoiding the psychological pressures, which usually accompany Asian students.

6 Research Roadmap

In this section, we present our intended research roadmap, expected to be addressed within the SACSys (Safe and Secure Adaptive Collaborative Systems) project [31] in collaboration with the Volvo Group Collaborative Robot Systems Laboratory [16]. Within this project, design and runtime compliance of collaborative systems is in focus. Given the disruptive nature of the type of systems proposed by the case study owners (i.e., 1) self-driving vehicles in quarry sites cooperating with manned vehicles and 2) robots performing collaborative tasks within a manufacturing factory), and given the ongoing regulatory disruption within the automotive and robotics domain, we intend to analyse the ongoing changes within the standards that are under discussion in order to design our barbells. Specifically, we need to identify which parts of the standards represent the maximal certainty (known knowns)/low risk and which parts represent the maximal uncertainty (unknown unknowns). Once this is identified, in cooperation with the industrial partners and regulators in the automotive and robotics domains, we will design a series of case studies to explore the possible benefits of having a bi-modal strategy-oriented for the regulatory frameworks on one side and a bi-model strategy-oriented compliance management on the other side. The goal would be to show that it could turn out to be an optimal solution to combine traditional and binding regulatory frameworks with experimental and sunset-based regulatory frameworks on one side and compliance by the book (via compliance checking) and experimental demonstration (via argumentation) on the other side. Our research will also target the elaboration of heuristics for distributing the weights on the two extremes. Initially, the focus will be on process-related requirements and then it will be extended to product-related requirements.

7 Conclusion

In this paper, we have presented our vision for an optimal solution for handling disruptive technological solutions impacting regulatory frameworks and compliance management. Specifically, after having introduced a broader perspective on the impact and state of the art practices in handling corresponding disruptive consequence within regulatory frameworks and compliance management, we have proposed our Barbell strategy-oriented vision and our research roadmap towards the concretisation of our vision.

References

1. Bloomfield, R., Khlaaf, H., Ryan Conmy, P., Fletcher, G.: Disruptive innovations and disruptive assurance: assuring machine learning and autonomy. Computer **52**(9), 82–89 (2019)

2. Castellanos Ardila, J.P., Gallina, B.: Towards increased efficiency and confidence in process compliance. In: Stolfa, J., Stolfa, S., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2017. CCIS, vol. 748, pp. 162–174. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64218-5_13
3. Castellanos Ardila, J. P., Gallina, B.: A personal opinion survey on process-based compliance checking in the safety context. In: 13th International Conference on the Quality of Information and Communications Technology (QUATIC) (2020)
4. Castellanos Ardila, J.P., Gallina, B.: Separation of concerns in process compliance checking: divide-and-conquer. In: Systems, Software and Services Process Improvement (EuroAsiaSPI). Springer, Düsseldorf (2020)
5. Castellanos Ardila, J.P., Gallina, B.: Formal contract logic based patterns for facilitating compliance checking against ISO 26262. In: 1st Workshop on Technologies for Regulatory Compliance (TeReCom), pp. 65–72. CEUR-Workshop Proceedings (2017)
6. Castellanos Ardila, J.P., Gallina, B., Governatori, G.: Lessons learned while formalizing ISO 26262 for compliance checking. In: 2nd Workshop on Technologies for Regulatory Compliance (TeReCom), pp. 1–12. CEUR-Workshop Proceedings (2018)
7. Castellanos Ardila, J.P., Gallina, B., Ul Muram, F.: Facilitating automated compliance checking of processes in the safety-critical context. Electron. Commun. EASST **078**, 1–20 (2019)
8. Coglianese, C.: Listening, learning, leading: a framework for regulatory excellence. J. Nurs. Regul. **8**(4), 64 (2018)
9. Cortez, N.: Regulating disruptive innovation. Berkeley Technol. Law J. **29**(137) (2014). . SMU Dedman School of Law Legal Studies Research
10. Emmerich, W., Finkelstein, A., Montangero, C., Antonelli, S., Armitage, S., Stevens, R.: Managing standards compliance. IEEE Trans. Software Eng. **25**(6), 836–851 (1999)
11. Gallina, B.: A model-driven safety certification method for process compliance. In: 2nd International Workshop on Assurance Cases for Software-Intensive Systems, Joint Event of ISSRE, Naples, Italy, 3–6 November 2014, pp. 204–209. IEEE (2014)
12. Gallina, B., Gómez-Martínez, E., Earle, C.B.: Deriving safety case fragments for assessing MBA safe's compliance with EN 50128. In: Clarke, P.M., O'Connor, R.V., Rout, T., Dorling, A. (eds.) SPICE 2016. CCIS, vol. 609, pp. 3–16. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-38980-6_1
13. Gallina, B., Gómez-Martínez, E., Benac Earle, C.: Promoting MBA in the rail sector by deriving process-related evidence via MDSafeCer. Comput. Stand. Inter. **54**, 119–128 (2017)
14. Gallina, B., Ul Muram, F., Castellanos Ardila, J.P.: Compliance of agilized (software) development processes with safety standards: a vision. In: 4th international workshop on Agile Development of Safety-Critical Software (ASCS) (2018)
15. Geman, D., Geman, H., Taleb, N.N.: Tail risk constraints and maximum entropy. Entropy **17**, 1–14 (2015)
16. Götvall, P.-L.: Volvo group collaborative robot systems laboratory: a collaborative way for academia and industry to be at the forefront of artificial intelligence. KI - Künstliche Intelligenz **33**(4), 417–421 (2019). https://doi.org/10.1007/s13218-019-00622-0
17. Governatori, G., Shek, S.: Regorous: a business process compliance checker. In: 14th International Conference on Artificial Intelligence and Law (ICAIL), pp. 245–246. ACM (2013)

18. Governatori, G.: The regorous approach to process compliance. In: IEEE 19th International Enterprise Distributed Object Computing Workshop (EDOCW), pp. 33–40. IEEE (2015)
19. Hesselberg, J.: Unlocking Agility: An Insider's Guide to Agile Enterprise Transformation. Addison- Wesley Signature Series (Cohn), Boston (2018)
20. Institute for Occupational Safety and Health of the German Social Accident Insurance (IFA): Collaborative robots (COBOTS). Safe co-operation between human beings and robots. https://www.dguv.de/ifa/fachinfos/kollaborierende-roboter/index-2.jsp
21. ISO/TC 22/SC 32: ISO 26262-1:2018 Road Vehicles-Functional Safety - Part 1: Vocabulary (2018)
22. ISO/TC 299 Robotics: ISO 10218-1 Robots and robotic devices - Safety requirements for industrial robots - Part 1: Robots (2011)
23. ISO/TC 299 Robotics: ISO/TS 15066:2016 Robots and robotic devices - Collaborative robots (2016)
24. ISO/TC 299 Robotics: ISO/CD 10218-2 Robotics - Safety requirements for robotics in an industrial environment - Part 2: Robot systems and integration (under development)
25. ISO/TC 299 Robotics: ISO/DIS 10218-1 Robotics - Safety requirements for robot systems in an industrial environment - Part 1: Robots (under development)
26. Joshida, J.: AV Safety Standards Set to Arrive in 2020. https://www.eetasia.com/news/article/AV-Safety-Standards-Set-to-Arrive-in-2020
27. Lau, A.: Nurturing the Anti-Fragile Student - The Barbell Strategy and Academic Excellence sans the Strain. In: EduTECH Philippines (2020)
28. Organisation for Economic Cooperation and Development (OECD): Reducing the Risk of Policy Failure: Challenges for Regulatory Compliance. http://www.oecd.org/regreform/regulatory-policy/1910833.pdf
29. Pries-Heje, J., Johansen, J.: The SPI Manifesto (2009)
30. Russo, D., Ciancarini, P.: A proposal for an antifragile software manifesto. Procedia Comput. Sci. **83**, 982–987 (2016)
31. SACSys Team: SACSys project. http://www.es.mdh.se/projects/540-Safe_and_Secure_Adaptive_Collaborative_Systems
32. Sadiq, S., Governatori, G., Namiri, K.: Modeling control objectives for business process compliance. In: Alonso, G., Dadam, P., Rosemann, M. (eds.) BPM 2007. LNCS, vol. 4714, pp. 149–164. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-75183-0_12
33. Taleb, N.N.: The Black Swan. Random House, New York (2007)
34. Taleb, N.N.: Antifragile. Things that gain from disorder. Random House, New York (2012)
35. Ul Muram, F., Gallina, B., Gómez-Martínez, L.: Preventing omission of key evidence fallacy in process-based argumentations. In: 11th International Conference on the Quality of Information and Communications Technology (QUATIC), pp. 65–73 (2018)
36. Underwriters Laboratories Inc. (UL): UL 4600 - Standard for Safety for the Evaluation of Autonomous Products (2020)

An Automated Pipeline for the Generation of Quality Reports

Ricardo Eito-Brun(✉)

Universidad Carlos III de Madrid, 28903 Getafe, Spain
reito@bib.uc3m.es

Abstract. This document describes the implementation of a document automation process to automate the creation of quality reports according to the standards used in the Aerospace software industry. Companies developing software in this industry must produce and deliver quality reports to provide evidences of the verification activities and the quality characteristics of software products. From a document engineering perspective, these reports must incorporate data and content concerning different software development processes and activities, data that are usually stored and managed with different software applications and tools.

The generation of this type of reports represents a complex data aggregation and integration process. This paper describes a solution based on the use of the XML language for data integration and a set of connectors that makes possible the collection of data managed in the external applications that support the software engineering activities (requirements management, architectural design, testing, etc.).

Keywords: Document automation · Aerospace industry · XML-based data aggregation · Software engineering · Software development · Software quality

1 Introduction

Software engineering discipline has developed a useful set of skills and competences to streamline the creation of complex documents based on the single-source content management paradigm and technical standards like DITA (Day 2003; Hackos 2016). Although technical communication is still understood – by a high number of people – as an activity focused on the elaboration of instructional materials and user guides -, the skills and competences of our discipline can report substantial benefits to other engineering tasks that require the elaboration and dissemination of technical data, information and knowledge about complex software-based systems.

The development of software projects involves the creation of different documents that communicate information about the system specifications and design, its behaviour and the decisions taken during their implementation. A distinction can be made between those documents that provide guidance on the use and operation of the software, and those documents that describe the system internal characteristics to support its maintenance and future evolution. Methodologies and standards like ISO

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 706–714, 2020.
https://doi.org/10.1007/978-3-030-56441-4_53

29110, ISO/IEC/IEEE 15289 or the Rational Unified Process (RUP) dictate the documents to be created, their structure and expected content.

Software documentation also includes the quality reports generated at intermediate stages of the project to provide evidences of the achieved tasks and product quality. These reports are extremely important in regulated industries like aerospace, automotive or medical software, as they are the basis for product qualification. These reports need to include the qualitative, expert assessment of engineers, as well as the quantitative data on which their assessments are based. The need of reusing and merging in a single document the data generated by different tools in heterogeneous formats (text files, CSV or XML, to name typical cases), makes the creation of these reports complex. We can apply the experience and the lessons learnt on topics like XML data management and single sourcing strategies to streamline this complex process.

The proposed case provides a practical answer to the need of introducing improvements in the software development process as a whole, and in particular in those aspects with a significant impact on teams' productivity. This clearly affects the development and maintenance of product documentation and quality data, an aspect that is typically overlooked by practitioners. The automation of document-production and maintenance activities is a good example of the improvements that can be reached according to the SPI Manifesto, whose Principle 3 "Base improvement on experience and measurements" indicates the need of optimising how people do the processes in their day-to-day business. Principle 10 "Do not lose focus" highlights the need of improvement making a contribution to better fulfil business goals.

2 Document-Oriented and Data-Oriented Content

Software developed for the European Space Agency (ESA) must be compliant with the ECSS standards. These standards provide requirements and guidelines that govern engineering activities and specify the documents to be created and delivered at the different project milestones. ECSS standards establish two quality-related reports: the Software Verification Report (SVR) and the Software Product Assurance Milestone Report (SPAMR). The SVR has the purpose of "presenting gathered results of all the software verification activities that have to be executed along the software development life cycle". The purpose of the SPAMR is "to collect and present at project milestones the reporting on the software product assurance activities performed during the past project phases". Both documents need to include data that are handled with other tools, like traceability between requirements and test cases, the results of tests and inspections, adherence to coding rules, software metrics or the status of software problems.

ECSS standards do not impose the use of mark-up languages. But it is clear that XML may be applied in the document generation process to reduce costs and improve time-to-delivery. In this case, an XML-based solution has been designed to automate the creation of quality reports. The solution must support authors in different tasks: a) collection of the input data from different tools and their aggregation in the target document; b) content editing, to incorporate additional textual information and c) publishing of the final deliverable in formats like PDF, Word or HTML.

Quality reports are good representatives of the fuzziness that characterizes some documents that contain both unstructured information and structured data. Authors in charge of writing these reports must collect the data from all those tools, and spend time formatting the data before integrating them in the final document. This is a time consuming, error-prone process. Automating all these tasks can save time to the authors. If the manual processing of data is reduced, authors can focus their effort on the assessment of the data and the elaboration of written conclusions.

Another factor that leads to reconsider the current approaches is the increasing popularity of agile methodologies. Agile methods recognise the importance of putting relevant information at the disposal of the interested parties, but the short periods allocated to the development of increments is seen as an obstacle for the elaboration of all the documents that are requested in regulated software industries (Heeager 2012). The automation of the document generation tasks can be a significant contribution to demonstrate the feasibility of using agile practices in regulated industries.

3 Data Integration and Document Generation

The document generation process is split into four main tasks:

a) Data generated by software engineering tools is collected and moved to a working folder. These data may be stored in different format, depending on the tools' capabilities (CSV, plain text files or XML).
b) Those data are converted to XML by means of XSLT stylesheets. This results in a set of intermediate XML files. These XML files can be combined automatically to create a draft version of the document.
c) Authors use an XML editor (Altova Authentic) to complete the creation of the report. Using the editor, they can edit and enhance the draft version with additional information and content.
d) Once the document is completed, the final deliverables in PDF, HTML and Word format can be created by running a stylesheet.

Software Quality Reports (SVR and SPAMR) need to combine data coming from heterogeneous sources and textual content entered by the document authors with the analysis of those data. To support the aggregation data from different tools, the use of the XML language was considered as one pillar for the target solution. Document analysis led to the definition of different XSD schemas (Fig. 1):

- One document schema was designed to model the target reports structure (SVR and SPAMR). This document reused part of the elements defined in the DITA standard, popular in the context of end-user documentation.
- Additional schemas were designed for the different data that were required as part of the report content, in particular:
  - Traceability data, between requirements, element design, and test cases.
  - Results of coding standards verification.
  - Software Metrics.
  - Problem data and status.
  - Verification tasks.

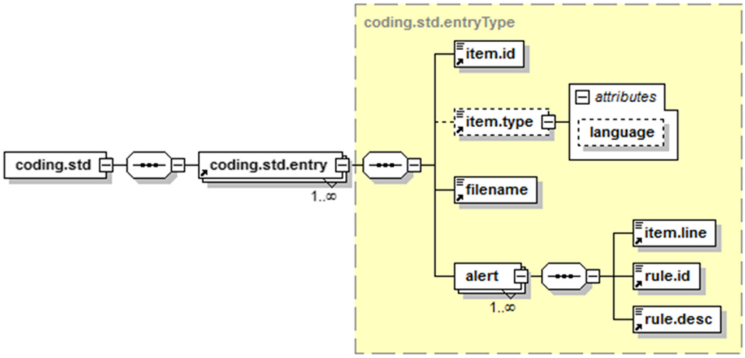


Fig. 1. Fragment of the XSD schema

One characteristic of software development projects is that all the data above are usually managed with different, specialized tools. Information about requirements and traceability needed to be taken from tools like IBM DOORS and Enterprise Architect. The verification of coding standards was done with specialized tools like CheckStyle, PMD and Findbugs (for Java projects) and one legacy tool, Foresys, for Fortran code. Metrics were obtained with another tool, SciTools Understand®. Information about software issues and tasks were stored in one collaborative, web-based platform, Redmine. The tools listed above are widely used to manage the different data generated in the software development projects. Of course there are additional tools and alternatives. It does worth to remark that the design of XSD schemas to serialize data about traceability, problems, metrics, etc., is independent of the use of one tool or another.

Figure 2 shows a schema of the data conversion and exchange approach:

The proposed approach consists on the use of custom data connectors to:

a) Automatically get the data needed for the generation of the quality reports from these tools,
b) Convert these data into an intermediate XML file that makes possible its processing and aggregation and
c) The authors can then incorporate the XML data into the final document using a standard XML editor.

The process automates the harvesting of the data from the applications mastering the master data, and its conversion into the intermediate XML format and their integration into the report draft that will be later edited by the authors. The use of XML gives also the possibility of generating multiple representations of the documents in the formats used for its final delivery: Microsoft Word and PDF, using standard XSLT stylesheets.

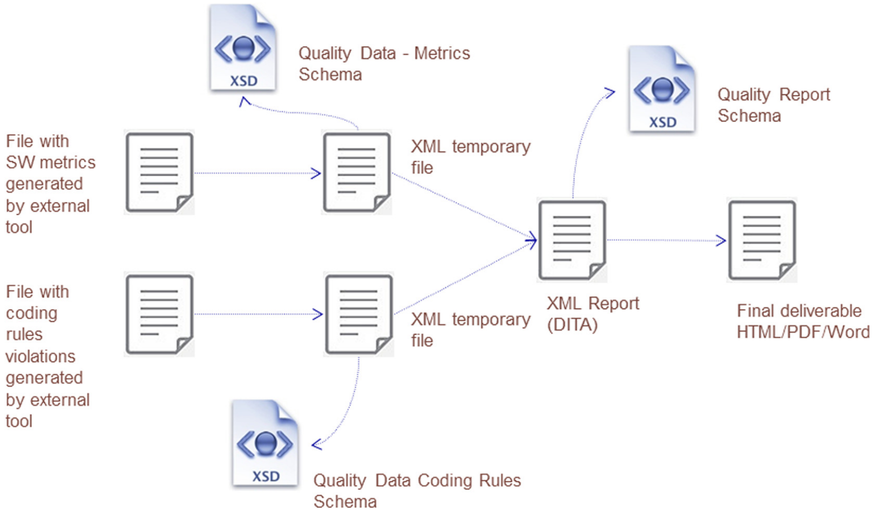


Fig. 2. Overview of the document processing pipeline

4 Technical Implementation

The automation of the process is mainly based on the use of XSLT transformations, designed with the Altova MapForce and StyleVision tools. The module provided by Altove MapForce to convert CSV and plain text files with no markup into XML was extremely useful for managing the conversion of data created by legacy tools that do not have the capability of exporting to XML. Data extraction from the software engineering tools was done using different approaches. In some cases, the tools offer a command-line interface that makes possible the scheduled execution of commands; in other cases, more complex approaches were needed (for example, to collect data from the DOORS requirements and traceability tool)

In addition to the technical aspects, the implementation of this solution consisted of these subtasks:

- Data requirements analysis: an analysis of the ECSS standards was done to identify the data to be incorporated in the quality reports. Data transformation requirements were studied in this preliminary phase.
- Identification of data providers: these are the software tools used by the company to generate the data identified in the previous step.
- Data extraction, conversion and merging: this task included:
  - the development of software-based connectors to automate the extraction of source data and their conversion to XML using XSLT.
  - The development of a small VB.NET program to merge XML documents and create the draft version of the document that can be later edited using the XML editor.
- The design of XSLT/XSL-FO stylesheets for publishing the document in PDF, RTF and HTML.

4.1 Data Requirements Analysis

The analysis of the ECSS standards resulted in a list of data requirements derived from the description provided in the standard for the two reports under study (SVRs and SPAMRs).The analysis considered the requirements that must be satisfied by these documents. As an example, Annex M of ECSS-E-ST-40C requests to complete its chapter 4.5 of the SVR with "evidences for the clause 5.7.3.5 of ECSS-E-ST-40C that requests the traceability between acceptance tests and the requirements baseline". In the case of the SPAMR, annex C of ECSS-Q-ST-80C states that the chapter 6 must report the "adherence to design and coding standards", and chapter 7 must "report on the collected product and process metrics, the relevant analyses performed, the corrective actions undertaken and the status of these actions".

Data requirements analysis led to the identification of a set of entities and data items that these reports should aggregate in raw or consolidated form.

4.2 Identification of Data Providers

Data identified by the previous analysis are usually created and managed with the help of specialized tools, and stored in different formats. Although companies in the European software industry may rely on different tools to support engineering processes, a certain degree on consistency is observed. For example, IBM® DOORS™ is widely adopted for requirements and traceability management; in the case of source code verification, SonarQube has become popular in projects based on the Java programming language.

In the development of this case study, the list of data providers included: SciTools Understand, Logiscope, SonarQube and a Fortran legacy tool for software metrics and coding rules verification, TestLink for test management, Redmine, used to handle information about task, problems and change requests (ECRs) and DOORS.

4.3 Data Extraction, Conversion and Merging

As indicated in a previous paragraph, data extraction from the provider tools and their conversion to XML was implemented using different strategies, depending on the capabilities offered by the different tools. In some cases this was implemented by executing scheduled commands using the command-line interface. In other cases, for example with Redmine, the Rest API of the tool was used to gather XML data about software problems, tasks and change requests. Different strategies were defined depending on the provider tool, in order to:

a) Gather the data needed for the generation of the document,
b) Convert these data into intermediate XML files and
c) Merge the data into a preliminary, draft document.

XSLT functions for data filtering and grouping were used. For example, for some metrics the document must only report those values that are above or below a specific threshold. In this case, the XSLT stylesheet not only needs to convert the metrics data to XML; it is also convenient to filter the dataset to leave only those data that must be included in the report (Fig. 3).

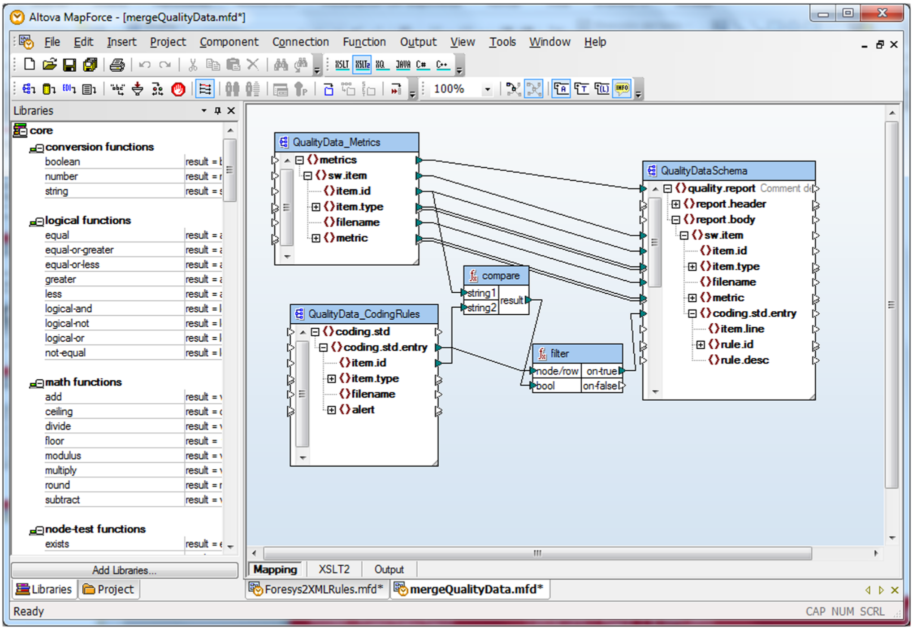


Fig. 3. Example of an intermediate XSLT stylesheet to merge data.

For the different types of data that must be aggregated in the quality and verification reports, separate XSD schemas were designed to serialize the data collected from the source tools:

- Traceability data between requirements and element design, and between requirements and test cases.
- Results of the verification of the adherence to coding standards.
- Software Metrics, containing metrics like module, class and method size, cyclomatic complexity, nesting level, percentage of comments, etc.
- Problems and their status and request for deviations and waivers.
- Verification tasks, their status and basic information like start and ending date, person in charge, etc.
- Test results and coverage.

4.4 Content Edition

Once the draft XML document is built, it can be manually edited to incorporate additional information. At this point, the users are provided with an XML editor with standard functions to edit and incorporate new content. In the case of the Altova Authentic editor, it is remarked that it may be extended with the help of an scripting language to incorporate additional functions, like searching for XML content in the file system or execute remote calls to other applications and tools.

Figure 4 shows the final quality report generated after the aggregation of the data collected in XML format from the different tools providing the results of the software verification and quality activities. The user-friedly version of the resulting XML document is the result of merging and aggregating the different data collected automatically after their processing and normalisation. In this case, the screenshot shows the view of the report with information about source code metrics.

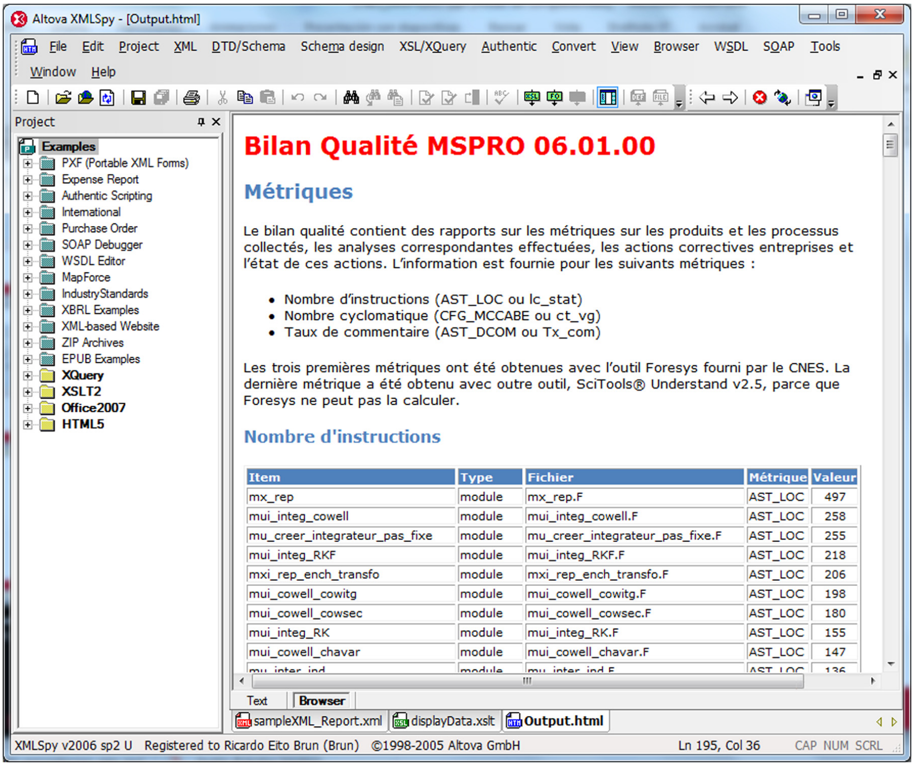


Fig. 4. HTML view of the final document.

5 Conclusions

The proposed solution was tested and validated with the help of staff working for a French aerospace company developing software according to the ECSS standards. The tool was presented to four authors with experience on the creation of this type of reports, who were asked to check the potential benefits of the tool and the feasibility of using it in a real scenario. Positive feedback was provided, and the tool was considered to be an effective and efficient solution to reduce the time needed to build these documents. In particular, avoiding the manual formatting of the data coming from the provider tools was the most-valued feature.

The conclusions can be summarized as follows:

- The proposed solution was recognized as a valuable tool to reduce the time needed to generate these documents.
- Keeping the input data as separate XML files was also positively valued: these data can be saved for future reuse as they are arranged according to a well-defined XML schema.
- The XML editor was considered user-friendly, although some users showed their reluctance to use a new document editing tool.
- A reduction of potential errors in the generation of these reports was observed and positivelly assessed by the involved staff.

As a general conclusion, we can conclude that the technical skills acquired with the use of documentation standards like DITA and other XML-based standards can be successfully applied in other areas of the company. XML-based content management processes and related tools and technologies (XSLT in particular) can be widely used in the creation of other documents, in particular those that require reformatting or repurposing of existing data and content. This includes not only quality reports, but also other system and software documents like those related to logistics and configuration management. This is probably the most important lesson learnt as a result of this activity: significant saving and improvements may be obtained by applying automated document generation techniques and tools at the different activities of the SW development process.

References

Day, D., Hennum, E., Hunt, J., Priestley, M., Schell, D.: An XML architecture for technical documentation: the darwin information typing architecture. In: Annual Conference-Society For Technical Communication, vol. 50, pp. 248–254 (2003)

Hackos, J.T.: International standards for information development and content management. IEEE Trans. Prof. Commun. **59**(1), 24–36 (2016)

Pries-Heje, J., Johansen, J. (eds.): SPI Manifesto. Version A.1.2.2010. https://2020.eurospi.net/images/eurospi/spi_manifesto.pdf

Heeager, L.T.: Introducing agile practices in a documentation-driven software development practice: a case study. J. Inf. Technol. Case Appl. Res. **14**(1), 3–24 (2012)

European Space Agency: ECSS-E-ST-40C. Space Engineering – Software Engineering, ESA Std. (2009)

European Space Agency: ECSS-Q-ST-80C. Space Engineering – Software Quality, ESA Std. (2009)

Team Skills and Diversity Strategies

Social and Human Factor Classification of Influence in Productivity in Software Development Teams

Liliana Machuca-Villegas[1,2](✉), Gloria Piedad Gasca-Hurtado[2](✉), Luz Marcela Restrepo Tamayo[2], and Solbey Morillo Puente[2]

[1] Universidad Del Valle, Calle 13 # 100-00, 760032 Cali, Valle del Cauca, Colombia
liliana.machuca@correounivalle.edu.co
[2] Universidad de Medellín, Carrera 87 No. 30-65, 50026 Medellín, Colombia
{gpgasca, lmrestrepo, smorillo}@udem.edu.co

Abstract. Software development organizations continuously seek to increase the productivity of their teams. However, productivity in the development of a software product is characterized by a high dependence on social and human factors. Therefore, studying these factors is essential to meet the performance needs of software development organizations. The above supports the motivation to consider the social and human factors that may influence the productivity of software development teams. This research has been conducted to classify such factors associated with the productivity of software development teams. For this purpose, a methodology was designed based on systematic literature review processes and evaluation processes with a psychology expert. The classification of social and human factors and their respective definitions are the partial results of this investigation. The proposed classification is an essential input for the establishment of improvement actions.

Keywords: Social and human factor · Productivity · Software development team

1 Introduction

The software development process is people-centered; the work team's integration and cohesion to meet the project goals are required [1]. Therefore, social and human factors (SHF) play an essential role in the productivity and success of the process [2]. Software development organizations are interested in this issue, which constantly seek to increase the productivity of their teams [3]. In other words, software development teams' productivity has a high level of dependency on social and human factors [4, 5]. Consequently, studying these factors is essential to meet the performance needs of software development organizations.

Studies related to factors of influence on software development productivity were found. Oliveira et al. [6] presented the results of a tertiary review about factors of influence on developer productivity. This review is based on secondary studies where are presented with a set of influencing factors. On the one hand, Wagner and Ruhe

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 717–729, 2020.
https://doi.org/10.1007/978-3-030-56441-4_54

presented a classification of technical and non-technical (soft) factors [7]; Trendowicz and Münch identified influence and context factors [8]; Dutra et al. classified factors into team emerging state factors, individual characteristic factors, and support task factors [9].

On the other hand, Yilmaz [10] studied social aspects and personality profiles that should be considered in the productivity of software development teams. He argues that productivity depends on the social aspects of productivity, which can be achieved by better social alignment. That means social factors influence software productivity [11]. Likewise, according to Cunha De Oliveira's results [4], the most outstanding factors that influence software developer productivity are human factors over organizational and technical factors.

The above supports the motivation to study the social and human factors that may influence the productivity of software development teams. In previous studies, 57 SHF were identified, grouped into factors at the individual and social levels [12]. As a continuation of this paper, factors were evaluated from the context of organizational psychology and software engineering. As partial results of this study, thirteen SHF and their respective definitions were established. SHF were analyzed from a global and integrated perspective of the social and the individual, understanding SHF within the individual-group interaction. They were also analyzed in a non-hierarchical way to focus on how they can favor a software development team's results. This set of factors seeks to facilitate the identification and definition of SHF to propose strategies for improving productivity.

This paper is organized as follows. Section 2 describes studies related to factors that influence software development productivity. Section 3 shows the methodology followed to evaluate the factors. Section 4 presents the SHF proposal and its definitions. Finally, in Sect. 5, the conclusions and future work are mentioned.

2 Background and Related Work

This section presents studies associated with the importance of studying social and human factors in software engineering. These studies are addressed towards software development productivity.

Gómez Pérez [13] studied software development teams, their characteristics, and their interactions to determine personality factors, task characteristics, team processes, and work climate that influence team effectiveness. The effectiveness was measured based on the product quality and the satisfaction of the team members. Results showed a relationship between software quality and team satisfaction with the work climate. Additionally, they indicated a positive and direct influence of certain personality factors such as extroversion on the software quality. Based on the results obtained, Gómez Pérez proposed a guide of recommendations for the creation of effective teams.

Sommerville [14] analyzed some factors associated with personnel management in software development projects and considers them as an important element in the success of a project since the project manager must motivate his team. Regarding the factors that may influence teamwork, Sommerville [15] reported the following: group composition, cohesion, communication, and organization. Similarly, Weimar et al. [16]

carried out a study of the quality factors of teamwork that influence team performance and software project success. These factors are measured in terms of product quality, efficiency, and effectiveness.

Further, Murphy-hill et al. [5] conducted a study with developers from three organizations on productivity. In the study, they identified the dominant factors are oriented towards non-technical factors such as enthusiasm at work, peer support for new ideas, and receiving useful comments on performance. Likewise, Dias Canedo y Almeida Santos [17] investigated the factors that affect software development projects' productivity and open-source projects. They recognized 37 factors that influence productivity, categorized into people, product, organization, and open-source projects.

Factors influencing productivity have been identified in the context of agile development. Fatema & Sakib [18] presented productivity factors in agile work teams using system dynamics. In their study, team effectiveness, team management, motivation, and customer satisfaction were the most prominent factors. Similarly, Iqbal et al. [19] empirically analyzed the factors of influence on agile teams. They determined factors such as inter-team relationships, team speed, team vision, and other factors related to the team member and lead roles, as the most common factors.

Researches have been carried out aimed at the study of emotions in software developers and their influence on productivity. Graziotin et al. researched the influence of happiness on developers [20], their results are interesting. The consequences generated by happiness influence productivity, code quality, and personal well-being.

Finally, there are studies discussed factors and emotions that influence the productivity of software development teams. These studies present different approaches to classifying factors, which SHF are not considered explicitly. Generally, studies focus on the analysis of factors influencing productivity and only a few propose recommendations for their improvement. On the other hand, the factors related to the work team are presented from a very general perspective. Therefore, a factor classification at a social and human level is necessary to generate improvement proposals for the work team.

3 Methodology

A methodology was used for the identification and selection of SHF that influence software development productivity [12]. It is based on systematic literature review processes and evaluation processes with experts in psychology and software development. A summary of the phases proposed in the methodology is shown in Fig. 1.

Phase 1 (Parameterization) defines the study purpose. In this phase, objectives, research questions are established, and secondary studies are selected. During phase 2 (Data Extraction), the data extraction strategy of secondary studies is specified. Forms are used to store general data of the articles reviewed and the factors identified. These data are the inputs for phase 3 (Classification of SHF). The data are analyzed and the SHF selection process is carried out. This phase is one of the essential phases of the process since, through its cycle and iteration scheme, it is possible to achieve the SHF classification. Finally, in phase 4 (Validation), the identified factors are evaluated. In this phase, two validation cycles are proposed. In cycle A, an evaluation with a

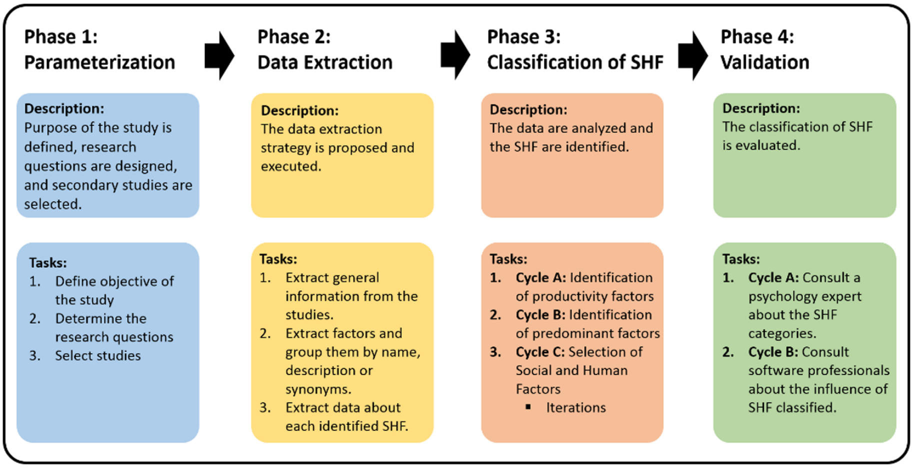


Fig. 1. SHF identification and selection methodology [12]

psychology expert to review the SHF from its definition and organizational psychology. In cycle B, an evaluation with software professionals to measure the perception of SHF through a survey.

The first three phases were applied in previous studies obtaining 57 preliminary SHF [12]. The evaluation of these factors (Phase 4 - Cycle A) is presented in this paper.

4 Social and Human Factor Classification Proposal

In this research, SHFs can be considered as human being characteristics that identify him based on his behaviors from a social and individual perspective [4]. In addition to these characteristics (person internal elements), there are stimuli (person external elements) that constitute the mental and influence the person from the social environment where he lives. Therefore, opportunities to execute certain behaviors such as collaboration or empathy, can be created [21].

This section presents the selected SHFs and its definition. Early results are exposed in [12], and the final description of the process is summarized below.

4.1 Factor Selection Process

This process began with the study of SHF that influence software development productivity. In [12], 57 preliminary factors related to aspects at a personal and social level were identified. To get a final classification of these factors, they should be assessed. To this end, factors were reviewed from a psychological and software engineering point of view. Their meaning, the relationship between them, and the complexity of their study within this research were analyzed. During a semester and in several frequent meetings, these factors were assessed by a psychologist along with the research team (two software engineers and two statisticians).

For this process, it was taken into account that the human being is an individual integrated and interrelated by biological (inheritance), psychological (individual) and social (culture) factors [22]. For this reason, the SHF analysis was carried out in an integrated way. We understand that human being complexity implies a whole that encompasses relationships he has with himself and with the social environment where he lives. Additionally, the influence (positive or negative) of social interactions on an individual's relationships is recognized.

We defined three selection criteria to facilitate the evaluation process with the psychology expert. These criteria show how we have carried out the analysis and revision of the factors in such a way that it allows us to delimit the scope of this research. In Table 1 is described these criteria.

Table 1. Selection criteria

ID	Criterion	Description
SC-1	Complexity	Factors whose study involves a whole knowledge subspecialty (sciences such as psychology or sociology), considering definitions from various authors and currents. Therefore, they are factors that are classified as complex and are excluded from the study because they demand a higher analysis that may be outside the scope of this investigation
SC-2	Group	Factors of human behavior that can be analyzed in an integrated way, either because they are factors that occur together or one of them, result from the other. Additionally, these factors are associated with their affinity and the difficulty that would imply delimiting them
SC-3	Exclusiveness	Factors considered mutually exclusive because its conceptual definition can be approached without intertwining with other factors, but it does not mean that it is not interrelated with the other factors Considering the factors as exclusive does not imply that they are more important than the others

It is necessary to clarify that we proposed the separation of SHF for analytical purposes and from a conceptual perspective. We are aware that, in real life, a human being is complex integrity that involves all the SHF mentioned here.

We also contemplate the study of those pertinent factors for the intervention of a software development team through gamification and those SHF that experts (software development and psychology) consider as pillars of the teamwork and can be studied for their relationship with other factors.

According to the methodology, the SHF evaluation process with a psychology expert belongs to Phase 4-Validation, Cycle A. The expert psychologist has a profile oriented towards a cognitive-behavioral approach, relying on neuroscience for the study of behavior and higher mental processes. During this cycle, we performed tasks to facilitate the evaluation process and definition of the selected factors. Firstly, we define each of the factors, and then we apply the selection criteria to choose the factors presented in this paper. As a result of applying the criteria, some SHF were removed, others were grouped, and others remained exclusive. An outline of the process is shown in Fig. 2.

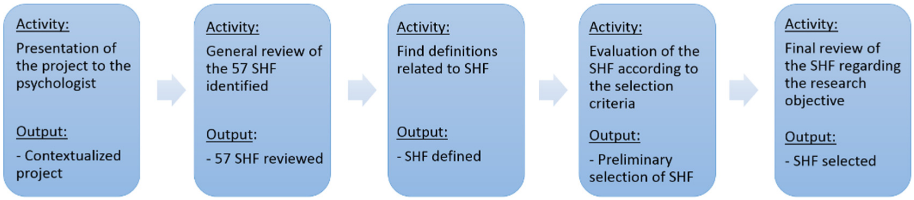


Fig. 2. SHF evaluation process

Removed SHF were excluded according to selection criteria SC-1. These factors are Group psychology, Diversity, Change resistance, Intelligence, Learning skills and styles, Rationality, Abstract/reflective thinking, Personality, Thinking styles, Experience, and Knowledge. Their study and their analysis that implies a whole knowledge subspecialty from psychology or sociology are complex. These factors constitute fundamental elements in the study of human social behavior. However, they are finally removed because their study's complexity would be outside the scope of the research objective.

Grouped SHF were integrated with other factors by applying selection criteria SC-2. These factors are Motivation, which groups Team motivation and commitment factors. Collaboration, which brings together the factors Cooperation, Sharing knowledge and information, Camaraderie, Trust. Third, Team Cohesion, which encompasses the Team Identity, Clear Goals, Teamwork capabilities factors. Empathy and Interpersonal Relationship is the fourth factor, which includes the factors Mutual respect, Group learning, Attitudes, Judgment, Bias. Fifth, Leadership, which contains the Sense of Eliteness factor. Autonomy groups the factors Independence and Self-efficacy. As a seventh factor, Capabilities and Experiences in the software development process, which brings together the factors Analyst Capability, Design experience, Programmer Capability, Development skills, Applications Experience, Platform Experience, Language and Tool Experience. Finally, Capabilities and Experiences in software project management include the factors Manager Capability, Manager Application Experience, Management quality and style, Estimation skills, Integrator experience, and skills.

Exclusive SHF were selected factors when applying selection criteria SC-3. The autonomy of these factors does not imply that they are disarticulated from the holistic study of human behavior. Instead, their conceptual definition can be approached without mixing with the other factors, constituting a single factor to consider. The selected factors are Communication, Commitment, Work satisfaction, Emotional Intelligence, and Innovation.

4.2 Selected Factors

After the evaluation process, 13 SHF were selected. We took into account the following topics in the identification and selection of the SHF definitions: i) work team, "a small number of people with complementary skills who are committed to a common purpose, set of performance goals, and approach for which they hold themselves mutually accountable" [23]. ii) Software development is a people-centered process that requires the integration and cohesion of the work team to meet the project goals [1], relies predominantly on teamwork [24]. iii) Gamification, it favors the development of social and human factors and constitutes itself as a potential tool to influence the commitment, communication, and motivation of work teams in software engineering [25–27].

Knowing the SHF of influence on software development productivity can help software organizations face the difficulties associated with them. For instance, reducing software project management failures, reducing development time, reducing the cost of the product, and finally improving the productivity of the work team. Likewise, our research support improvement strategies in software development processes and designing productivity measurement instruments at the level of teamwork.

In Table 2, the SHF selected are presented with the definition of each one. These definitions will ease the validations of these factors based on the design of a questionnaire. The purpose of the questionnaire in the future is to measure the perception of the influence of social and human factors in development software teams.

To better understand how influences the factor on software development productivity, we define the factor in terms of items. The items were derived from the theoretical definition of each SHF. An example of the collaboration factor items is presented in Table 3. As shown in the example, it was done with other factors.

Table 2. Human and social factors and definitions

Factor	Definition
Communication	Communication is the method that allows a human to link with another human [21]. It is the bridge that channels the transmission of thoughts, emotions, and information [28]. The valid forms of communications (verbal, written, gestural) are the ones that take into account the relevance of the messages, the correct use of the language, the adequate moment to transmit the information, the method used, and the interpretation of the perception of the interlocutor In the work teams, the forms of communication ease the adequate flux of information, promoting dynamics that influx the outcome of software development projects [7, 29]. That can be seen in the clarity of instructions and objectives, the preventions of reprocessing, improvement of tasks and activities, and assertive conflict solving
Collaboration	The collaboration is related to the sentiment of feeling supported by others, a tendency of working together in a solidarity environment. In teamwork, collaboration arises when various members work together for a mutual benefit [30]. The capacity to share knowledge, information, and know-how to achieve a common goal is also essential. Collaboration is related to trust, friend relationships, or cordiality between team members to create a proper environment and achieve goals Collaboration is opposed to competition even though the possibility of achieving personal or collective progress can be a source of conflicts, frustration, envy, and aggressiveness [21]

(*continued*)

Table 2. (*continued*)

Factor	Definition
Commitment	It is the level of responsibility a team member is willing to assume in his/her tasks inside the work team. In the same way, the team is responsible for the goals of the project. A responsible person is consequent with his/her actions, achieves his/her tasks, and can assume his/her mistakes [21]
Motivation	Motivation is the factor that moves the team member into action [21]. The motivation is inherent to every individual, has an intensity, force, and duration. It varies according to the objective and determines part of human behavior [31]. The motivations can be uncountable and, according to Herzberg [32], cited by [33], are classified in: Intrinsic motivations: they surge of the pleasure of doing the job. Imply desire to do the task because they matter, like are exciting or are part of something valuable (autonomy, mastery, and purpose) [34]. They are related to the position and the tasks of the position. It produces a long-lasting satisfaction effect, an increase of productivity reaching excellence, way better than normal levels. It also includes feelings of realization, growth, and professional recognition. These are evidenced in the execution of tasks and challenging activities that are important for the job. When the motivation factors are optimal, the satisfaction is substantially improved; when not, satisfaction is lost [33] Extrinsic motivations are related to the job's external factors, like the physical conditions that surround the individual when he/she is working. Comprehend the physical and environmental conditions of the job, the salary, the social benefits, the corporate policies, the type of supervision, the status of the relationships between directors and employees, the rules of procedure, the opportunities, etc. These aspects influxes in a limited way, the workers' behavior because, when optimal, they only prevent dissatisfaction [33]
Work satisfaction	Work satisfaction is determined by the discrepancy between what he/she wants and what he/she has in the job [35]. It relates to the content or the challenging and stimulant activities of the position. The expectations, personal and professional growth, recognition, opportunities, salary, work environment, and others are the aspects that measure work satisfaction [21]
Leadership	Leadership is the capacity that some people have to influence others inside the team focused on achieving goals and objectives [36]. The leader exalts in the team because he/she has a group of qualities and characteristics: Important security in himself/herself and his/her convictions and ability to take decisions and risks. He/she also stands out in three dimensions: His/her activity, the capacity to perform tasks and objectives, and the affection he/she receives. Teamwork requires a collaboration attitude and a high integration capacity, which depends on a good leader and his/her capacity to coordinate the team members' functions. This way, high efficacy levels can be achieved in comparison with individualized work [21]
Innovation (creativity)	The innovation manifest in an individual or teamwork with creative thinking as the capacity to establish a new relationship between facts or to integrate them originally or innovatively. The innovation is related to the creativity needed to elaborate new, different, valuable trough experiences and know-how [21]
Emotional intelligence	The ability that an individual has to identify and process adequately his/her emotions. That implies not getting dominated by the emotions, instead of control his/her behavior. Emotional intelligence relies on the quality of the interpersonal relationships with the team members. It will let him/her recognize the emotional state of his/her teammates and act empathetically, forging ties of friendship and confidence. Although it is an individual factor, the adequate management of emotions in complex group situations, like conflict solution, has a significative influx (positive or negative) in the work environment [37, 38]

(*continued*)

Table 2. (*continued*)

Factor	Definition
Autonomy	It is defined as the capacity to take workable decisions independently from the management. It directly relates to the liberty the employee and the team have to take decisions related to the project and the tasks in it [9]. Could include the decisions: what should be done first, which projects to investigate, which ideas to develop, how to solve problems, and how to adapt the job without external interference Autonomy impulses creativity and improves intrinsic motivation. Additionally, it is related to auto efficacy and how individuals trust his/her abilities and capacities [39]
Empathy and interpersonal relationship	Empathy is the mental state in which an individual relates to another individual or a group sharing the same mood. It needs a disposition of openness to the other, so the feelings that they are feeling can be understood. In social interaction, values like respect are necessary. This implies a consensus in the group rules, establishing limits, and a margin that everybody gives and receives [21]. This way, good personal relationships are promoted as the social nature of the human being inside the socialization dynamics with his/her teamwork Empathy used to be a bridge to build social relationships. Nevertheless, empathy does not happen without processes like admiration or proximity with the other. This allows a step-by-step sociocultural knowledge in which every person appropriates the stimuli he/she receives from others and the environment. With the stimuli, the behavior is molded according to the needs [21]. Inside these processes, aspects like attitudes and partiality exist. These aspects act upon the decisions an individual takes based on value judgments according to the received stimuli allowing him/her to do a quick (and unconscious) global evaluation relatively stable that determines his/her conduct
Team cohesion	The identity of the team is an essential aspect of cohesion. To this end, the individual should feel parts of the team acting motivated and autonomically. Additionally, the team's objectives must be consistent with the objective of the individual and his/her goals. The cohesion of a work team refers to the level of integration of its members so that all efforts are focused on the same common goal. This integration eases teamwork due to the ability to distribute responsibilities between team members for decision-making and to have different points of view to avoid biases and wrongdoings in the work process [21]. In teamwork, a social characteristic where the dynamics that meddle social relationships, empathy, and collaboration prevails. This makes an individual feel an active and positive part of a team and makes him/her a potentiator of the team's performance. Likewise, the definition of clear objectives allows team members to focus on the achievement of the tasks planted by the group
Capabilities and experiences in software development process	It is the knowledge and experience in the analysis, design, and development of a software product according to every member of the team's goal. Demands familiarity with the dominance of the application, the hardware and software platform and the tools and languages of programming
Capabilities and experiences in software project management	It is the knowledge application, abilities, tools, and techniques directed at the project activities to achieve the requisites of it [40]. These capacities and experiences are directed to functions like planning, dynamic of resources organization (human, economic and materials), decision making, procedures elaboration and achieving that team members do the tasks they are responsible for [21]

Table 3. Example factor items

SHF	Items
Collaboration	• To improve productivity in software development processes, team members should collaboratively work to achieve the project's goals • There should be trust among team members for the performance of their duties and protection of common interests, which contributes to improving productivity in software development processes • To improve productivity in software development processes, team members should be willing to help, endorse, and support their colleagues • To improve productivity in software development processes, team members should share their knowledge, information, and experience with their colleagues

5 Conclusions and Future Work

This study presents a set of social and human factors that influence software development productivity. These factors have been identified and selected using a methodology based on systematic literature review processes and an evaluation process with an expert in psychology.

As a result of this research, 13 social and human factors are presented with their definitions. A rigorous and systematic process was performed. It began with identifying the factors from literature and ended with the evaluation of the factors by a psychologist.

SHF were analyzed from a global and integrated perspective of the social and the individual, understanding the SHF within the individual-group interaction. They were also analyzed in a non-hierarchical way to focus on how they can favor the results of a software development team.

During the identification and selection of the SHF definitions, we tried that the definitions were associated with the concepts: work team, software development, and gamification. We used these three concepts as a foundation to guide the description associated with SHF and to maintain the scope of this research. However, this does not mean that each SHF explicitly shows the relationship with these three concepts.

It is important to take into account criteria from psychological and software engineering point of view for the selection of SHF, so that an analysis of them from the human component and the work context is allowed.

Currently, we have identified a future line of work related to the level of influence of the factors. To achieve this influence, we propose to use a system dynamics based approach to establish a nonlinear relationship between attributes. A Cause-effect relationship can support clarify and quantify the influence of the SHF.

Additionally, we are designing a survey-based study to measure the perception of the influence of social and human factors on software development teams. We are identifying the predominant SHF considered by the practitioners in the software industry consulted through this survey. We plan to analyze and determine the relationship between SHF using factorial analysis.

Acknowledgements. This study is part of the "Classification and Influence of Social and Human Factors on Software Development Team Productivity in Small and Medium IT Businesses" project, financed by La Universidad de Medellín (Colombia), El Centro de Investigaciones en Matemáticas (Zacatecas-México) and La Universidad de Guadalajara (México).

We thank Kimberly Londoño Ruiz, psychologist, for her support in the review, selection, and definition of the social and human factors presented in this paper.

References

1. Hernández, L., Muñoz, M., Mejia, J., Peña, A.: Gamificación en equipos de trabajo en la ingeniería de software: Una revisión sistemática de la literatura Gamification in software engineering teamworks: a systematic literature review (2016)
2. Muñoz, M., et al.: A model to integrate highly effective teams for software development. In: Stolfa, J., Stolfa, S., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2017. CCIS, vol. 748, pp. 613–626. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64218-5_51
3. de Barros Sampaio, S.C., Barros, E.A., De Aquino, G.S., et al.: A review of productivity factors and strategies on software development. In: Proceedings of 5th International Conference on Software Engineering Advances, ICSEA 2010, pp. 196–204 (2010). https://doi.org/10.1109/icsea.2010.37
4. De Oliveira, E.C.C.: Fatores de influência na produtividade dos desenvolvedores de organizaciones de software. Universidade Federal Do Amazonas (2017)
5. Murphy-hill, E., Jaspan, C., Sadowski, C., et al.: What predicts software developers' productivity? IEEE Trans. Softw. Eng. 1–13 (2019)
6. Oliveira, E., Conte, T., Cristo, M., Valentim, N.: Influence factors in software productivity - a tertiary literature review. In: Proceedings of 30th International Conference on Software Engineering Knowl Engineering, pp. 68–103 (2018). https://doi.org/10.18293/seke2018-149
7. Wagner, S., Ruhe, M.: A systematic review of productivity factors in software development. In: Proceedings of the 2nd International Software Productivity Analysis and Cost Estimation (SPACE 2008), pp. 1–6 (2008)
8. Trendowicz, A., Münch, J.: Factors influencing software development productivity-state-of-the-art and industrial experiences. Adv. Comput. **77**, 185–241 (2009). https://doi.org/10.1016/S0065-2458(09)01206-6
9. Dutra, A.C.S., Prikladnicki, R., França, C.: What do we know about high performance teams in software engineering? Results from a systematic literature review. In: 41st Euromicro Conference on Software Engineering and Advanced Applications (2015)
10. Yilmaz, M.: A software process engineering approach to understanding software productivity and team personality characteristics: an empirical investigation. Dublin City University (2013)
11. Yilmaz, M., O'Connor, R.V., Clarke, P.: Effective social productivity measurements during software development—an empirical study. Int. J. Softw. Eng. Knowl. Eng. **26**, 457–490 (2016). https://doi.org/10.1142/S0218194016500194
12. Machuca-Villegas, L., Gasca-Hurtado, G.P.: Towards a social and human factor classification related to productivity in software development teams. In: Mejia, J., Muñoz, M., Rocha, Á., A. Calvo-Manzano, J. (eds.) CIMPS 2019. AISC, vol. 1071, pp. 36–50. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-33547-2_4
13. Gómez Pérez, M.N.: Estudios experimentales sobre la influencia de la personalidad y el clima en el desarrollo de software. Guías para gestión de equipos en proyectos de ingeniería del software. Universidad Autónoma de Madrid (2010)

14. Sommerville, I.: Software Engineering, 10th edn. Pearson, London (2016)
15. Sommerville, I.: Ingenieria del Software, 10th edn. Pearson, London (2005)
16. Weimar, E., Nugroho, A., Visser, J., Plaat, A.: Towards high performance software teamwork. In: Proceedings of 17th International Conference on Evaluation and Assessment in Software Engineering – EASE 2013, pp. 212–215 (2013). https://doi.org/10.1145/2460999.2461030
17. Canedo, E.D., Santos, G.A.: Factors affecting software development productivity: an empirical study. In: ACM International Conference on Proceeding Series, pp. 307–316 (2019). https://doi.org/10.1145/3350768.3352491
18. Fatema, I., Sakib, K.: Using qualitative system dynamics in the development of an agile teamwork productivity model using qualitative system dynamics in the development of an agile teamwork productivity Model. Int. J. Adv. Softw. **11**, 170–185 (2018)
19. Iqbal, J., Omar, M., Yasin, A.: An empirical analysis of the effect of agile teams on software productivity. In: 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), pp. 1–8. IEEE (2019)
20. Graziotin, D., Fagerholm, F., Wang, X., Abrahamsson, P.: What happens when software developers are (un)happy. J. Syst. Softw. **140**, 32–47 (2018). https://doi.org/10.1016/j.jss.2018.02.041
21. Vallejo-Nágera, J.: Guía Práctica de Psicología. Ediciones Temas de Hoy (2002)
22. Belloch Fuster, A., et al.: El modelo bio-psico-social: un marco de referencia necesario para el psicólogo clínico. Clin. Health **4**, 181–190 (1993)
23. Katzenbach, J.R., Smith, D.K.: The Discipline of Teams. Harvard Business Press, Boston, Massachusetts (2008)
24. Tang, X., Kishore, R.: The antecedents and consequences of agile practices: a multi-period empirical study of software teams in time-bound projects. In: Proceedings of the International Conference on Information Systems (ICIS) and International Research Workshop on IT Project Management. Saint Louis, Missouri, USA, pp. 142–157 (2010)
25. Muñoz, M., Hernández, L., Mejia, J., Gasca-Hurtado, G.P., Gómez-Alvarez, M.C.: State of the use of gamification elements in software development teams. In: Stolfa, J., Stolfa, S., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2017. CCIS, vol. 748, pp. 249–258. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64218-5_20
26. Gasca-Hurtado, G.P., Gómez-Álvarez, M.C., Manrique-Losada, B.: Using gamification in software engineering teaching: study case for software design. In: Rocha, Á., Adeli, H., Reis, L.P., Costanzo, S. (eds.) WorldCIST'19 2019. AISC, vol. 932, pp. 244–255. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-16187-3_24
27. de Paula Porto, D., Ferrari, F.C., Fabbri, S.C.P.F.: Improving project manager decision with gamification. In: ACM International Conference on Proceeding Series (2019). https://doi.org/10.1145/3364641.3364675
28. Consuegra, N.: Diccionario de psicologia-Ecoe Ediciones (2010)
29. Paiva, E., Barbosa, D., Lima Jr., R., Albuquerque, A.: Factors that Influence the productivity of software developers in a developer view. Innov. Comput. Sci. Softw. Eng. (2010). https://doi.org/10.1007/978-90-481-9112-3
30. Tomasello, M.: Por qué cooperamos. Primera ed. Katz Editores (2010)
31. Hernández López, A.: Medidas de productividad en los proyectos de desarrollo de software: una aproximación por puestos de trabajo (2014)
32. Herzberg, F., Mausner, B., Bloch Snyderman, B.: Motivation to Work. Wiley, New York (1959)
33. Chiavenato, I.: Administración de recursos humanos. Quinta Edición Santa Fé de Bogota. Editorial McGraw-Hill (2000)
34. Pink, D.: La sorprendente ciencia de la motivación. TEDGlobal (2009)

35. Judge, T.A., Piccolo, R.F., Podsakoff, N.P., et al.: The relationship between pay and job satisfaction: a meta-analysis of the literature. J. Vocat. Behav. **77**, 157–167 (2010)
36. Vivas, P., Rojas, J., Torras, M.: Dinámica de Grupos. Fundación para la Universidad Oberta de Catalunya FUOC (2009)
37. Feldman, R.: Psicología Con Aplicaciones A Países De Habla Hispana. Décima Edi. McGraw Hill (2014)
38. López Jordán, M.E., González Medina, M.F.: Inteligencia Emocional. Ediciones Gamma (2009)
39. Riso, W.: Terapia cognitiva. Fundamentos teóricos y conceptualización del caso clínico. Editorial Norma (2006)
40. Project Management Institute, IEEE Computer Society: Software Extension to the PMBOK® Guide Fifth Edition, Fifth. Project Management Institute, Inc. (2013)

What Motivates VSEs to Adopt an International Standard Such as ISO/IEC 29110? An Exploratory Analysis

Mirna Muñoz[1(✉)], Jezreel Mejia[1], Adriana Peña[2], Claude Laporte[3], and Gloria Piedad Gasca-Hurtado[4]

[1] Centro de Investigación en Matemáticas, Parque Quauntum, Ciudad del Conocimiento, Avenida Lassec, Andador Galileo Galilei, Manzana, 3 Lote 7, 98160 Zacatecas, Zac., Mexico
{mirna.munoz,jmejia}@cimat.mx

[2] Departamento de Ciencias de la Computación, CUCEI de la Universidad de Guadalajara, Av. Revolución No. 1500, 44430 Guadalajara, Jal., Mexico
adriana.pena@cucei.udg.mx

[3] Department of Software and IT Engineering, École de technologie supérieure, 1100, Notre-Dame Street West, Montréal, Québec H3C 1K3, Canada
Claude.Laporte@etsmtl.ca

[4] Ingeniería en Sistemas, Facultad de Ingenierías, Universidad de Medellín, Carrera 87 No. 30-65, Medellín, Colombia
gpgasca@udem.edu.co

Abstract. Very Small Entities (VSEs) of México are important to the economy because of their contribution to the growing demand of software. The ISO/IEC 29110 series has become a highly adopted standard in México because it is one of the Quality Standards recognized by the government, the industry and academia. This is confirmed by the increasing number of VSEs certified to this standard since 2013. Over 46 of the 53 VSEs certified in Latin America are from México. The objective of this research was to perform an exploratory analysis of the factors that motivate and affect the motivation of VSEs in the implementation of this standard. To perform this research, a case study was conducted and a questionnaire to collect data was developed. The sample of our research is a set of 15 VSEs that have adopted the software Basic profile of ISO/IEC 29110 to improve their software development process. This research provides the factors that motivates positively and negatively the adoption of the ISO/IEC 29110 standard by VSEs.

Keywords: ISO/IEC 29110 · Software Process Improvement (SPI) · Very Small Entities (VSEs) · Factors motivating implementation · Factors affecting implementation

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 730–741, 2020.
https://doi.org/10.1007/978-3-030-56441-4_55

1 Introduction

Software development organizations have shown significant economic activity in recent years around the world [1, 2] because the growth in the importance of software to generate software products and services capable of providing high-performance solutions to problems of different domains (industrial, agricultural, aeronautics, Information and Communication Technologies, among others).

Around the world, the importance of Very Small Entities (VSEs), representing over 92% in many countries [1, 3], is recognized. VSEs are becoming the main producer in a supply chain of software products because they develop software for many medium and large companies and organizations [3, 4].

This fact provides the opportunity for software development organizations to produce high-quality software products and services to satisfy market needs. But also, it highlights the increasing need for improving their development processes to steadily grow in the software market.

A VSE could be an enterprise, an organization (e.g., public or non-profit organization), a project or a department having up to 25 people [1–6], that often must work hard to survive, having the minimal time and effort to invest for improving their operation and processes.

In this context, many authors have addressed the challenges to implementing Software Process Improvements (SPI) in this type of organizations [2–7] mainly because of their specific characteristics, e.g., lack of knowledge and practical experience regarding the implementation of SPI models and standards, they must work hard to survive and, they cannot get the all the employees and expertise needed, cost and time for activities related to SPI.

To help VSEs, the ISO/IEC 29110 has been developed. The ISO/IEC 29110 series provides standards and guides that are based on subsets of international engineering standards called profiles for VSEs [2, 3, 5].

This standard has been recognized by the Mexican government and the industry as one quality standard. This is why it has become a highly adopted standard by VSEs of México.

Nowadays, México is the country in Latin America that has about 86% of VSEs certified to this standard. The goal of this paper is to perform an exploratory analysis of what motivates VSEs to adopt an international standard, specifically the ISO/IEC 29110, to identify the factors that motivate and affect their implementation.

After the introduction, the rest of the paper is organized as follows: Sect. 2 provides the background of this research including the strategy of México to improve the competitiveness of VSEs; the related works and an overview of why ISO/IEC 29110 is interesting for the Mexican software industry; Sect. 3 presents the characterization of the specific environments of VSEs and their problems toward the implementation of SPI; Sect. 4 provides the exploratory analysis performed to identify the factors that motivate and affect the implementation of ISO/IEC 29110; Sect. 5 presents the results and, finally Sect. 6 provides the conclusion and future work.

2 Background

2.1 Mexican Government Strategy

According to a report published by the National Survey on Productivity and Competitiveness of Micro, Small and Medium Enterprises (ENAPROCE by its Spanish acronym) in 2018, there are about four million enterprises in México and about 97.3% of them are very small enterprises [8].

To help the software industry (VSEs, SMEs, and large enterprises) México created the PROSOFT 3.0 program [9]. This program aimed to provide economical support to software development organizations, so that they can become competitive internationally in the IT services sector and it ensured their long-term growth.

Two strategies are highlighted of this program because they are focusing on reinforcing the Mexican organizations to make them competitive internationally [9]:

- *Increasing the quality and quantity of talent in the software development and its services*: it refers to the availability of more and better human capital to take advantage of the great growth of the sector both at the local level and to access to international markets.
- *Achieving an international level in process capacity*: it refers to increase the capacity of the IT services sector adopting models and international standards that could increase the productivity and quality of software organizations so that the Mexican industry can be competitive internationally.

In this context, based on the PROSOFT program, the government of Zacatecas created a program in 2017 that aims to help VSEs to improve their competitiveness by implementing the ISO/IEC 29110.

2.2 Related Works

Model and standards for the software industry are elaborated to contribute to the development of quality products within budget and schedule, by optimizing efforts and resources. However, their implementation in software development organizations represents a great challenge. This fact has been highlighted by researchers who have published research works focused on identifying those challenges [1–7, 10]. However, we have not identified similar studies in México.

Among the studies, we have selected three of them because their goal is very closed to the one established in this paper. Next, the related works are briefly described.

Davila and Pessoa in [10] perform a case study at Perú, in a company to identify the factors, both positive and negative, that drive the implementation of the software Basic profile of the ISO/IEC 29110. They implemented a case study and a questionnaire to perform the analysis and targeted three projects within the company. They have identified the experience in IT and top management support as factors that positively influence the implementation of ISO/IEC 29110. They also identified competitive pressure, perceived usefulness, perceived ease of use, and user training as factors having a negative influence on the implementation of ISO/IEC 29110.

Wongsai, Siddoo, and Wetprasit in [11] performed a study in Thailand to identify the factors, having both negative and positive impacts that influence the implementation of the software Basic profile of ISO/IEC 29110 in VSEs. They analyzed 15 papers, from a Systematic Literature Review, then they evaluated them with ISO standard's experts to analyze its potential. They identified six categories of factors: finance, customer, internal business processes, learning and growth, organization, and SPI project program. Each category has a set of negative and positive factors. However, they highlighted the importance of effective management of available resources, staff turnover, and individual skills and experience.

Larrucea, O'Connor, Colomo-Palacios, and Laporte in [4] provided a study of hurdles that VSEs must face in the implementation of Software Process Improvements as well as the opportunities that the implementation of software process improvements offers to them. The results of this paper are based on the software industry experience of authors in multiple countries. The results provided a set of 10 barriers and 8 opportunities classified into 4 categories: financial, skills, culture, and reference models.

2.3 ISO/IEC 29110 Series and Mexican Software Industry

Since 2005, Working Group 24, of the ISO/IEC JTC1 SC7, developed the ISO/IEC 29110 Systems and Software Engineering series as a solution to help VSEs to face challenges such as improving productivity and quality with minimum costs [12].

In México, this standard series has been recognized as one quality standard equated for Mexican software industry because its features such as [12]: (1) it provides a roadmap of 4 profiles that could be selected by a VSE according to its specific needs; (2) it has two processes as the foundation for project management and the software implementation; (3) it works in VSEs using any development approach, methodology or tool; and (4) it provides a step-by-step guide based on a set of process elements such as objective, activities, task, roles and work products.

It is important to highlight that nowadays the ISO/IEC 29110 has one certifiable profile, the Basic profile. This profile is targeted for VSEs developing a single product by a single work team and has two fundamental processes (i.e., Project Management and Software Implementation) [3, 12]. These features make it very interesting for Mexican VSEs.

Table 1 provides an overview of the 2 processes and the activities of the software Basic profile of ISO/IEC 29110.

3 Very Small Entities

As mentioned in the introduction, the term 'Very Small Entities' covers organizations (e.g., public or non-profit organization), project, or departments having up to 25 people. These type organizations have specific features that become a real challenge in the implementation of a model or a standard. Table 2 provides an analysis of the VSEs' environment characterization and problems. To perform the analysis, we took some of the categories proposed in [4, 10, 11] as a base and named as factors.

Table 1. ISO/IEC 29110 processes of the software Basic profile [13]

Process	Purpose	Activities
Project management	Establishing and carrying out the tasks related to project management systematically, so that the project's objectives are completed with the expected quality, time, and costs	Project planning Project plan execution Project assessment and control project closure
Software Implementation	Performances, in a systematic way, of the activities related to the analysis, design, construction, integration, and test, according to the requirements specified of new or modified software products	Initiation Analysis Design Construction integration and tests Delivery

4 Exploratory Analysis of Factors that Motivates the Implementation of ISO/IEC 29110

This section presents the results of an exploratory analysis performed, in México, to identify, in a qualitative way, the factors having more influence in the adoption of the ISO/IEC 29110 standard, more specifically in the software Basic profile. To achieve it, we took the 4-step schema proposed in [10] to perform a case study: (1) to design the case study; (2) to prepare the data collection; (3) to analyze the evidence and (4) the report results. Next, a brief description of each step as well as the development of each step is provided.

4.1 Design the Case Study

This step aims to set up the case study so that it should be established the objective, what will be studied, the base theory, the research question, and the strategy to collect data [10].

- *Objective*: to perform an exploratory analysis to identify qualitatively the factors that have more impact to motivate VSEs in the implementation of ISO/IEC 29110.
- *The case*: to get data from VSEs after the implementation of ISO/IEC 29110 to identify the factors that motivate the implementation of ISO/IEC 29110 in their organizations, and that keep them using it.
- *Theory*: VSEs have shown significant economic activity in recent years [1], and in México are not the exception. The Mexican government has established a program to increase the competitiveness of the software industry [9], where ISO/IEC 29110 was recognized as one quality standard. Therefore, it has become a highly adopted standard by VSEs of México, reporting 46 of the 53 certified VSEs of Latin America [14].

Table 2. Characterization of VSEs environments and problems

Factors	VSE environment	Related problems
Financial [4, 10, 11]	It depends on project profits; with no or limited budget to invest in resources (human, time, and infrastructure)	• Lack of budget to invest in the improvement of the software process
Customer [11]	It has a high dependence on customers	• The pressure to deliver customer products.
Internal business [10, 11]	It is under increasing pressure to improve their productivity and quality while keeping costs to a minimum	• Lack of budget to invest in the improvement of the software process
Processes [4]	Most of them do not have defined processes, and software development is built as a craft	• Lack of experience in defining and using processes or software engineering practices
Learning and Growth/Skills [4, 10, 11]	It depends on the experience and skills of current personnel	• Lack of personnel so that one employee has several tasks assigned. Therefore, it is complicated that he/she performs only activities related to SPI • Personnel is often hired without having experience
Organization/Culture /HypSlash> [4, 10, 11]	It must work harder to survive	• Lack of knowledge about the importance of the software development process in product quality • Lack of process culture. • Lack of knowledge to be able to select an appropriate model or standard
SPI project/reference models [4, 11]	They prefer to adopt practices instead of a complete model or standard because they perceived models and standards as implementing bureaucracy in the organization without any profit	• The cost of implementing a model or standard is very expensive • Lack of or few experiences in the implementation of models and standards • The resistance to implementing a complete standard

- *Research question*: What motivates VSEs to adopt an international standard such as ISO/IEC 29110?
- *Data collection strategy*: (1) to collect data, a questionnaire was developed in Google forms; (2) to share the questionnaire to VSEs that have obtained the ISO/IEC 29110 certification; (3) to analyze the answers of VSEs, and (4) to identify the motivators and establish conclusions.

4.2 Prepare the Data Collection

This step aims to define how the collection will be performed [10]. As mentioned in the data collection strategy, a questionnaire was developed to get data from VSEs certified in the software Basic profile of the ISO/IEC 29110. The sample was composed of VSEs of the Zacatecas region who participated in a project to increase their competitiveness by implementing international standards such as the ISO/IEC 29110. The VSEs that participated are stand-alone enterprises (SDE) and Software Development Centers (SDC) of academic institutions, that implemented the ISO/IEC 29110 in courses and projects, of the region of Zacatecas.

This project started in 2017 [3] and has continued until 2020, having achieved 19 VSEs certified [14]. Table 3 shows the VSEs certified between 2017 and 2020.

Table 3. VSEs certified by year in the Zacatecas Region

Year	Number of VSEs participating in a certification process	VSEs participating in a monitoring audit
2017	8 VSEs (4 SDE and 4 SDC)	
2018	7 VSEs (2 SDE and 5 SDC)	7 VSEs (3 SDE and 4 SDC)
2019	6 VSEs (4 SDE and 2 SDC)	14 VSEs (5 SDE and 9 SDC)
2020	2 VSEs (1 SDE and 1 SDC)	19 VSEs (9 SDE and 10 SDC)

As Table 4 shows, the developed questionnaire has 7 questions focused on getting qualitative data from VSEs regarding their experience after getting the certification. Besides, it was supplied in two ways, by email and using a Google form.

Table 4. Questionnaire to collect data

#	Question
1	Why implement ISO/IEC 29110 in your organization? (What is your business motivation?)
2	Why did you work to get certified?
3	What difficulties did you have in the implementation of ISO/IEC 29110?
4	What difficulties did you have to obtain the ISO/IEC 29110 certification?
5	What benefits did you have in the implementation of ISO/IEC 29110?
6	What benefits did you have to obtain the certification of ISO/IEC 29110?
7	What type of support do you expect from the government to continue within the project?

4.3 Analyze Evidence

This step aims to collect data and the analysis and interpretation of the data [10]. It is important to highlight that the questionnaire was sent to 19 VSEs. However, we received 15 answers that mean 79% of the sample. Besides, 8 answers were received by Google forms and 7 answers were received by e-mail.

The analysis of the data collected is presented focusing on questions defined in Table 4 and using the same factors of Table 2.

Table 5 shows the analysis of the VSEs answers, it is structured as follows described: the first column on the left side lists the categories and, from columns 2 to 8 are listed the questions.

4.4 Report Results

This step aims to provide the findings of the analysis performed on the collected data [10].

Therefore, using the data presented in Table 5, findings regarding the motivators that impacted VSEs in the implementation of the software Basic profile of the ISO/IEC 29110 are next listed. It is important to mention that to consider a factor, it should be mentioned by at least three of the VSEs:

(Q1) Related to the reason for implementing the ISO/IEC 29110, they highlighted two factors. The most highlighted factor is *customer*, VSEs mentioned that with the implementation of ISO/IEC 29110, they will be able to get new customers or markets. The other factor was *learning and growth/skills*, VSEs mentioned that they will update their processes, products, or services based on the knowledge acquired.

(Q2) Related to work for getting the ISO/IEC 29110 certification, they highlighted three factors, *processes, customer, and SPI projects/reference models*. About *processes*, VSEs mentioned that they will achieve the implementation of processes and, therefore they will be more effective in software development. About *customers*, VSEs mentioned they will be able to obtain recognition based on an international standard. Finally, about *SPI project/reference models*, VSEs mentioned that they will ensure the quality of software projects because they are working under the ISO/IEC 29110 standard.

(Q3) Related to the difficulties detected in the implementation of ISO/IEC 29110, VSEs highlighted three factors *processes, learning and growth/skills, and organization/culture*. About *processes,* VSEs mentioned they had never developed software using documented processes. About*, learning and growth/skills,* VSEs mentioned the lack of collaboration of the teammates, as there is no knowledge of processes and there is also no structured way of working. Finally, about *organization/culture,* VSEs mentioned the resistance to change to work under a formal or systematic way to develop software.

(Q4) Related to the difficulties detected in getting the ISO/IEC 29110 certification, VSEs highlighted two factors *learning and growth/skills,* and *SPI project/reference models*. Regarding *learning and growth/skills,* VSEs mentioned the lack of experience in performing some activities related to quality such as verification, test definition and execution, architecture, and design. About *SPI project/reference models,* VSEs mentioned it was the first time in applying a standard, having to get a comprehension, and an understanding of it.

(Q5) Related to benefits obtained with the implementation of ISO/IEC 29110, they highlighted two factors *processes* and *customers*. About *processes*, VSEs mentioned they achieved a definition, documentation, and implementation of processes for developing their software projects. About *customers,* VSEs mentioned an increase in awareness of the need to work under a quality standard.

Table 5. Analysis of the collected answers

Factors	Questions						
	Q1	Q2	Q3	Q4	Q5	Q6	Q7
Financial [4, 10, 11]	E4, E9				E5	E1, E2, E4, E6, E7, E8, E9, E15	E4, E6, E7, E8, E9, E10, E12, E15
Customer [11]	E1, E2, E4, E8, E9, E11, E12, E14, E15	E2, E6, E7, E10	E3		E2, E4, E6, E7, E10, E11	E11	E6, E7, E9, E11
Internal business [10, 11]	E3, E5, E9	E4, E5, E8		E3, E10	E14, E15	E11, E14	
Processes [4]	E8, E12	E1, E2, E12, E14	E3, E4, E5, E9, E10, E12	E8, E11	E3, E8, E9, E10, E11, E12, E13	E10, E11	E9
Learning and Growth/Skills [4, 10, 11]	E6,E7, E8,E10	E3,E11	E1,E3; E6, E7	E4,E5, E6, E7, E13	E9,E10	E3,E6, E7, E8, E9, E15	E2,E3, E9,E13, E14
Organization/Culture [4, 10, 11]	E3	E3, E15	E1, E2, E3, E14	E9	E3, E9, E10	E5, E11, E13	
SPI project/reference models [4, 11]		E3, E8, E9, E14	E3, E8, E15	E1, E3, E9, E11, E12	E1, E9		E5, E8
Other	E13			E2, E14		E14	E1, E15

(Q6) Related to benefits obtained from achieving the certification of ISO/IEC 29110, they highlighted two factors *financial,* and *learning and growth/skills*. About *financial*, VSEs mentioned the government support in sponsoring a percentage of the cost of the Certification. About *learning and growth/skills,* VSEs mentioned the support obtained from a research center in training, reviews, and recommendations that help them to have a better understanding of the standard and how to implement it within their organizations based on their particular way of work.

(Q7) Related to the type of support expected from the government to continue keeping the certification, they highlighted three factors, *financial, customer,* and *learning and growth/skills.* The most highlighted factor was *financial*, VSEs mentioned they need the government to keep the financial support provided. About *customers*, VSEs mentioned they expect that the government promotes them and link them with potential customers. Finally, about *learning and growth/skills*, VSEs mentioned they expect the government to continue providing training toward achieving continuous improvement culture.

4.5 Discussion

Having the results of the exploratory analysis, we compare our results with the results of the selected related works [4, 10, 11].

Regarding the motivators that affect positively the implementation of ISO/IEC 29110:

(1) *Recognition of customers:* none of the related works mentioned this factor. However, in our results for both SDEs and SDCs, it represents a positive motivator. On the one hand, for SDEs, it allows the widening of its market. On the other hand, for SDCs, it allows getting the attention of new customers regarding the capacity of producing quality software and the software engineer's talent.

(2) *Implementation of a structured way of developing software:* this motivator is mentioned in the three related works [4, 10, 11] as a motivator having a relevant impact. However, in [10] authors mentioned that it has a neutral impact. In our results, for both SDEs and SDCs, ISO/IEC 29110 is a standard that provides the elements for developing software in a structured way.

(3) *New knowledge:* this motivator is mentioned in the three related works [4, 10, 11] as a motivator having an impact. In our results, for SDEs imply to train the personnel to be able to update the way they work toward using recent technology. For SDCs, on the one way for teachers, it allows them to be updated in recent technology, and therefore they can teach it to their students.

(4) *Economical support:* this motivator is mentioned in the three related works [4, 10, 11] as a barrier. However, in our results, it is recognized as one of the main motivators to push them to keep within the program of continuous improvement.

(5) *Technical support:* this motivator is mentioned in the three related works, [4] recognized it as a barrier, [10] recognize it with negative impact, due to the lack of time and money to training employees. In [11] it is considered as a success factor. In our results it is recognized as one of the main motivators because this technical support facilitates, both SDEs and SDCs, to understand the standard and its implementation.

Regarding the motivators that affect negatively the implementation of ISO/IEC 29110:

(1) *Lack of experience using international standards:* this motivator is mentioned in the three related works [4, 10, 11] as a barrier. In our results, it is confirmed that even when ISO/IEC 29110 has a minimal set of required work products, the lack of experience in both SDEs and SDCs was identified as a motivator with a negative impact.

(2) *The resistance to change of the organizational culture:* this motivator is mentioned in two related works, in [4] and [11] as a barrier. In our results, it is a factor with negative impact identified in both the SDEs and the SDCs and it is very related to the previous one, most of the time due to the lack of knowledge people feel afraid of participating in new improvements initiatives or adopting new proven practices.

5 Conclusions and Future Work

This paper presented an exploratory study to identify qualitatively the factors that affect positively and negatively the implementation of the software Basic profile of the ISO/IEC 29110 in VSEs.

As illustrated by the results, on the one hand, questions Q1, Q2, and Q5 allowed the identification of motivators that affect positively the implementation of ISO/IEC 29110 highlighting the recognition of customers, the implementation of a structured way of developing software, and their new knowledge.

On the other hand, questions Q3 and Q4 identified motivators that affected negatively the implementation of ISO/IEC 29110 highlighting the lack of experience about using an international standard that brings a challenge in the understanding and comprehension of the standard as well as the resistance to change of the organizational culture.

Finally, Q6 and Q7 highlighted two very important motivators that allowed VSEs to keep in continuously improving their development process using the international standard ISO/IEC 29110: the economic support provided by the government of Zacatecas and the technical support provided for the training about continuous software process improvement. As future work, the authors are collecting data to perform a quantitative analysis of the benefits obtained and the factors motivating the implementation of ISO/IEC 29110 in VSEs.

References

1. Laporte C., Muñoz M., Gerançon B.: The Education of students about ISO/IEC 29110 software engineering standards and their implementations in very small entities. In: IEEE Canada International in Humanitarian Technology Conference (IHTC), Toronto, Canada, pp. 94–98 (2017)
2. Laporte, C.Y., O'Connor, R.V.: Software process improvement standards and guides for very small organizations. An overview of eight implementations. CrossTalk **30**, 23–27 (2017)
3. Larrucea, X., O'Connor, R.V., Colomo-Palacios, R., Laporte, C.Y.: Software process improvement in very small organizations. IEEE Softw. **33**(2), 85–89 (2016). https://doi.org/10.1109/MS.2016.42
4. Larrucea, X., O'Connor, R.V., Colomo-Palacios, R., Laporte, C.Y.: Software process improvement in small organizations. IEEE Softw. 85–89 (2016)
5. Laporte, C., O'Connor, R.: Systems and software engineering standards for very small entities: accomplishments and overview. Computer **49**(8), 84–87 (2016)
6. Sanchez-Gordon, M.-L., de Amescua, A., O'Connor, R.V., Larrueca, X.: A standard-based framework to integrate software work in small settings. Comput. Stand. Interfaces **54**, 162–175 (2017). Part 3

7. O'Connor, R., Laporte, C.Y.: The evolution of the ISO/IEC 29110 set of standards and guides. Int. J. Inf. Technol. Syst. Approach **10**(1), 1–21 (2017)
8. INEGI: Encuesta Nacional sobre Productividad y Competitividad de las Micro, Pequeñas y Medianas Empresas (ENAPROCE) (2018). https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2019/especiales/ENAPROCE2018.pdf
9. Secretaría de Economía: PROSOFT 3.0 Program to develop the Information Technology Services Sector (2019). https://www.gob.mx/cms/uploads/attachment/file/440337/ro2019prosoft_1_pdf
10. Dávila, A., Pessoa, M.: Factors driving the adoption of ISO/IEC 29110: a case study of a small software enterprise. In: 2015 Latin American Computing Conference (CLEI), Arequipa, pp. 1–8 (2015)
11. Wongsai, N., Siddoo, V., Wetprasit, R.: Factors of influence in software process improvement: an ISO/IEC 29110 for very small entities. In: 2015 7th International Conference on Information Technology and Electrical Engineering (ICITEE), Chiang Mai, pp. 12–17 (2015)
12. ISO/IEC: Software engineering - Lifecycle profiles for Very Small Entities (VSEs) - Part 5-1-2: Management and engineering guide: Generic profile group: Basic profile. ISO/IEC TR 29110-5-1-2:2011 (2011). http://standards.iso.org/ittf/PubliclyAvailableStandards/index.html
13. Muñoz, M., Peña, A., Mejia, J., Gasca-Hurtado, G.P., Gómez-Alvarez, M.C., Laporte, C.: A comparative analysis of the implementation of the software basic profile of iso/iec 29110 in thirteen teams that used predictive versus adaptive life cycles. In: Walker, A., O'Connor, R. V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 179–191. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_14
14. NYCE: Companies certified to ISO/IEC 29110–4-1:2011 standard (2020). https://www.nyce.org.mx/wp-content/uploads/2020/01/PADRON-DE-EMPRESAS-CERTIFICADAS-EN-LA-NORMA-ISO-IEC-29110-4-1-16-01-2020.pdf

Educating Future Engineers – Meeting Challenges at Secondary School and University

Claudia Fussenecker(✉), Jörg Niemann, Alexander Paul, Martin Schlösser, and Marius Schöning

Düsseldorf University of Applied Sciences, Düsseldorf, Germany
Claudia.fussenecker@hs-duesseldorf.de

Abstract. Serious shortages of qualified professionals and technical job-specific skills are hampering Europe's sustainable growth. Even more so, relatively newer technologies, such as digitalization and Industry 4.0, require a new set of qualifications for future engineers. Deficiencies in these areas can have a negative impact on innovation and, therefore, also for the well-being of the industry and economy. New skills in education as well as new training methods are required in order to train successful engineers that meet the requirements set by industry and society. How does a future engineering education look like? To get a qualified picture of the necessary requirements, as well as possible limitations, different stakeholders, such as teachers, students of engineering studies and pupils in secondary schools or vocational schools have been questioned. This survey was combined with a literature review. Furthermore, a recently finished EU-research project regarding this subject has been taken into consideration and been evaluated in order to come up with future solutions.

Keywords: Engineering education · Skills · Teaching methods · STEM

1 Introduction

Many of today's societal, industrial and engineering challenges cannot be viewed from an old-fashioned angle of scientific knowledge, soft skills and technology. Instead they will have to be looked at from an interdisciplinary point of view that connects these branches and includes socio-economic capabilities [1].

Bearing this in mind the teaching strategies of our educational system will have to be redesigned to match these new standards, especially with regard to topics such as digitalization and Industry 4.0, where interdisciplinary knowledge is crucial. The main content of Industry 4.0 are information and communication technologies, such as Cyber-Physical Systems, Internet of Things, Physical Internet and Internet of Services. Students need to be confronted with these topics quite early in their education so they can make a well thought out decision in which industrial branch they want to pursue a career, or if they want an industrial career at all [1, 2]. In this paper not only the engineering education at university level will be taken into consideration but also the educational phase in school, which influences the choice of the study program.

Having a look at the already mentioned point of interdisciplinary knowledge one "of the main issues raised by industry are that many university graduates lack skills

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 742–753, 2020.
https://doi.org/10.1007/978-3-030-56441-4_56

which complement their technical knowledge, such as problem solving, team work and the ability to adapt to change" [3]. As one possible solution for this lack of skills, problem-based learning (PBL) in Science, Technology, Engineering and Mathematics (STEM) related subjects was introduced. PBL means that pupils will be given an understanding of the topics by solving "real-world, open-ended problems, mostly in groups", to give them a feeling for later work experience in the industry. It is based on four modern insights into learning: constructive, self-directed, collaborative and contextual [3, 4]. Teaching methods like these are referred to as inductive, which is the "natural human learning style", as every problem that needs to be solved comes with a real situation. The opposite of that is deductive, where students merely listen to what their professor has to say and try to memorize this information without having the ability to apply it to real world problems [5]. To match industry's requirements there needs to be a "relationship between academic institutions and industrial expectations" to improve students' knowledge as well as their transition into the industry, which cannot be fulfilled by simply choosing the deductive learning style for students [5, 6].

In addition to the above-mentioned points of the industry's expectations, Mahanija Md Kamal points out in the paper "Evaluation of an engineering program: A survey to assess fulfilment of industry requirements" from 2009 that "new engineering graduates are expected not only to be knowledgeable and technically competent but also good at communication, have leadership skills, be able to work in teams and have lifelong learning attributes" [7]. Even though the evaluation is already 10 years old, the importance of the mentioned skills increased. These requirements mainly arise due to the changing industrial environment influenced digitalization and Industry 4.0 [8]. Another crucial factor is of course the teachers' perspective on the different teaching approaches. When sticking to PBL as an example for practice-oriented teaching Madeleine Abrandt Dahlgren's paper "PBL from the teachers' perspective", published as long ago as 1998, provides us with valuable insights. Here she refers to the teachers' perspective on PBL and points out that "the teachers were in general positive to continue to work with the PBL program" and that they appreciated the closer and more effective contact with the students but that there are also difficulties, especially on how it is interpreted by the teachers. They looked at it either from their perspective as teaching staff or from the student's perspective as the receiving part and not as a whole correlated concept [9].

This shows that teachers also need trainings to be able to apply inductive teaching methods and to give their students an understanding of practice-oriented topics, whether it is through PBL or any other teaching style with practical relevance [10]. The challenge of today's educational system is to give pupils a realistic view of what to expect by choosing studies in an engineering subject.

2 ELIC – Engineering Literacy – An Erasmus Proposal

One funded EU-project, which deals with the challenges of educating future engineers in the right way, recently ended in September 2019. This research project triggered the competence of a teacher's teaching methods and technical didactics, especially those teaching secondary-level STEM-related subjects in terms of their ability to incorporate engineering topics in natural sciences classes, motivating more pupils to choose an engineering career. Teachers are seen here as a medium of change. Together with several universities and schools from Italy, Austria, Czech Republic and Germany, an E-Learning provider and the ECQA (European Certification and Qualification Association) a compliant skill set and exam for teachers, as well as a Massive Open Online Course (MOOC) platform to train teachers was developed. Previously, existing structures in various European countries were analyzed and compared with the requirements of industry and science. Based on the gaps, discovered in the analysis, a skill card was developed and a curriculum established. Additionally an online based training for teachers of STEM-subjects, to give them tools to improve their way of teaching, leading to a new generation of qualified engineers, was set up [11].

Funded by Erasmus+ this project falls under the Lifelong Learning Program, which aims at enabling people to pursue further learning experiences at any stage of their life. This leads to the lucky coincidence that the engineers of today also find themselves in a lifelong learning process. The project was funded for a duration of 24 months. The starting date was October 1st, 2017 [12].

Due to the requirements set up by the European Union and in order to receive the funding, measurable goals, such as an Intellectual Output (IO), had to be defined. For the ELIC-research project seven IOs, were determined, structuring the tasks of the two-year research work. After starting with the research on the training needs, a GAP-Analysis was conducted that took into account what kind of education the learning plan currently specifies, and compared it to the needs of society and industry. Their requirements were then also implemented into a Skill Card that named the engineering topics and their linkage to the STEM-subjects. This all culminated in the development of the training curriculum in which the detailed teaching content combined with the teaching tools, learning out-comes and time management was defined. Then a MOOC-Platform was set up in order to have a suitable platform to disseminate the content. This platform is still running [11].

The structure of the Intellectual Outputs can be used as a guideline for future research proposals, keeping in mind what was successful in this project. After that, stakeholders, such as school teachers, industrial partners as well as associations dealing with education matters were questioned in focus group interviews, in order to receive an honest feedback and ideas for future projects and research areas.

2.1 Results

The feedback given during a round table with teachers in Germany in May 2019 lead to the following conclusion and ideas for future research. The idea of coming up with already finalized teaching material, which includes state of the art practical examples, fostering an interdisciplinary approach of the STEM-subjects taught was evaluated as helpful. Even though the project language is English, teachers in Germany need all the material in German in order to avoid extra work and also possible user boundaries. This problem will also apply to the other participating countries.

Another Feedback, which unfortunately cannot be solved by an EU research proposal, is the fact, that in Germany each federal state has its own education policy, often leaving little or no space for experiments. Furthermore, an interdisciplinary approach is highly appreciated. The material needs to be directly linked to the teaching curriculum, including a standardized setting of level of difficulty, requirements, such as the needed level of theory etc.

Moreover, the teachers also expressed an interest in a "Rent an Expert"-model, where several times a year an expert from university or industry takes over the lesson. Ideally the expert brings in a toolbox of experiments to support their practical example and combines it with the current teaching content of STEM-subjects.

It can be noted that the teachers are in general very open minded regarding an interaction of different stakeholders, such as industry and universities, with their teaching and they also appreciate a constant exchange of knowledge. However, requirements have to be met so that the integration of new elements into the STEM lessons can succeed [13].

3 Methods

A survey was designed to obtain more information on how students currently studying an engineering subject and pupils in secondary school or vocational training view the current state of their engineering education. The survey was conducted in German and the results and graphs were then translated into English for this publication.

To receive representative results, around 100 participants each from secondary schools as well as universities were questioned. For each group of participants, pupils and students, a separate survey was conducted, however, the main focus remained the same and generally dealt with the following questions:

- The influence of STEM-subjects on study choice.
- The relatability of the content of STEM-subjects to engineering tasks.
- The influence of STEM-subjects on study success.

Both surveys were most interested in students' and pupils' overall satisfaction with how topics in school or university are taught. A specific focus was put on how their practical relevance could be emphasized more to reach a higher level of awareness of why certain issues are being taught, especially with regards to STEM-related subjects. An example for potential subjects for emphasizing practical issues is physics, which could illustrate wave theory by using autonomous driving cars and how they check their surroundings for obstacles. In Mathematics the use of derivations and integral calculus could be explained by including economical topics such as the decision of purchasing something for an investment [14, 15]. Here the survey also goes into detail on the pupils' suggestions on where and how practical teaching methods could be implemented. The purpose of the survey was to find useful information by analyzing two different perspectives in order to receive a clear picture of what can be improved at which step to reach the common goal of educating more qualified engineers in the future.

3.1 Student Survey

For students already studying engineering or an engineering related subject, the survey touches upon the problem whether their expectation of the studied subject was fulfilled and whether they felt well prepared for their studied subject by their previous high school/secondary school education.

In order to achieve a representative result with the survey, the aim was to have an even number of female and male students take part. The survey took place at the Duesseldorf University of Applied Sciences in June 2019 and 109 students participated. However, due to the fact that only 16% of the students in the department of mechanical and process engineering are female, only 17% of the participants of the survey were women [16]. But despite this uneven distribution of gender, it is still a representative survey, because this is a normal gender distribution in an engineering study program in Germany [17].

- The age of the participating students reaches from 18 to 32.
- Nearly 60% of the participants have got the general university entrance qualification. Only 40% received their entrance qualification through other accepted paths. As you can see most of the students were educated at a Gymnasium (the highest general education in Germany).

The participants say that in order to receive a successful university degree in an engineering subject the three most important subjects in school are Mathematics, Physics and English. In these three subjects they did not feel well prepared for an engineering study program.

As you can see in Fig. 1, 60.7% of the participants did not feel well prepared for studying engineering when they first started to study. This is a huge lapse in educating future engineers, which should already be taken care of during school education.

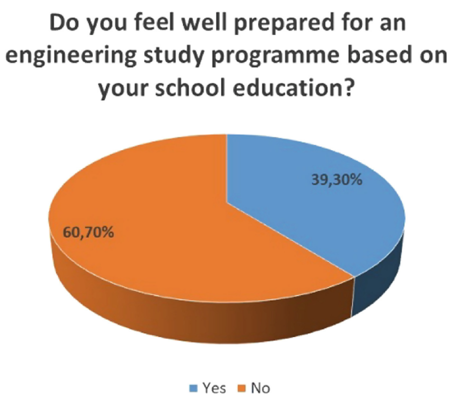


Fig. 1. Subjective preparation for studying engineering (students).

In Fig. 2 you can see that approximately 41% of the participants said that they do not know what the practical approach of the taught content of the STEM subjects could be. They know why the topics are relevant for their education, but they unfortunately do not know what to use them for. Only 31% of all participants knew why the topics are important and what they could use them for in the future.

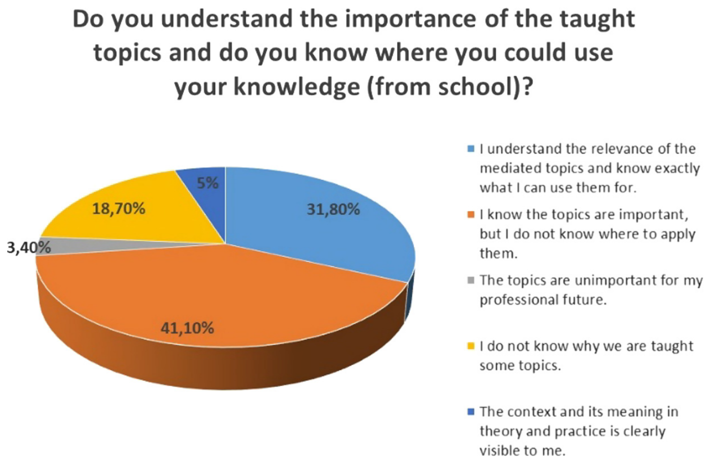


Fig. 2. Connection between theory and practice (students).

In general, over 90% of the participants asked for more classes focused on practical issues in general education and at the university. They said that it would help to get a more detailed overview on their possible future job tasks. Moreover, you can see in Fig. 3 that a larger number of classes focusing on practical problems would give the students a higher motivation, to deal with the relevant topics.

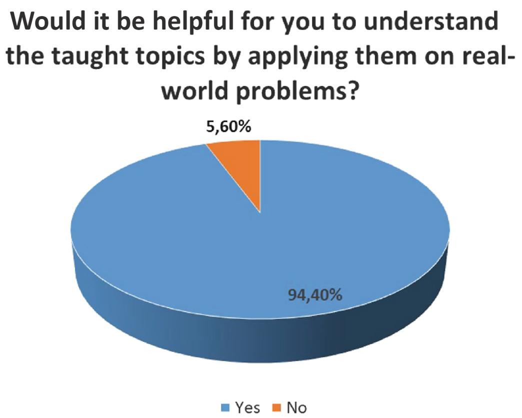


Fig. 3. Keep themselves busy with acquired topics (students).

3.2 Pupil Survey

At the beginning of the survey, pupils were explicitly asked, whether they were willing to pursue an engineering study program in the future.

- 39% of the 73 participants answered this question with yes and consequently took part in the questionnaire.
- 86% were from the German Gymnasium and 14% from vocational training school.
- 46% of the asked participants were 17 years old. All in all, the participants were between 16 and 19 years old. This is a normal distribution, because pupils in several school models were asked.
- 14% of them did an apprenticeship prior to their advanced technical college entrance qualification (Fachhochschulreife).

Because of that some of the pupils were a bit older than those that were enrolled at school at that moment. Similar to the survey for students 86% were men. This is due to the most common distribution in German engineering classes.

The survey was constructed in this way to create a comparability between both surveys. Similar to the student survey, the participants in the senior classes of the German educational system were asked whether they felt well prepared for an engineering study program by their general school education.

Similar to the student survey, most of the pupils think that the same three subjects in school are important to reach a good engineering grade (Mathematics, Physics, English). In addition to that, nearly 70% of the participants had the opinion that they were well prepared for studying engineering at a university (Fig. 4). There is no correlation to the opinion of those students that are currently studying engineering.

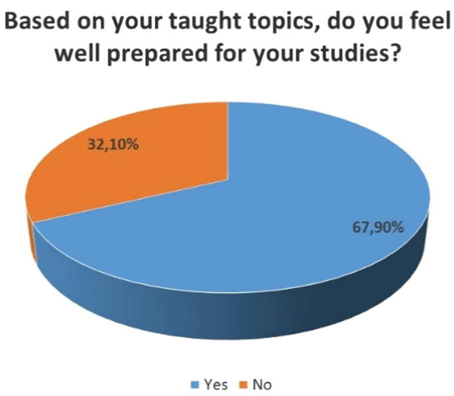


Fig. 4. Subjective preparation for studying engineering (pupils).

The next interesting aspect is that unlike to the student survey 50% of the pupils said that they knew exactly why the mediated topics are important for an engineering class and what they could use them for (see Fig. 5). Just 21% said that they didn't know what they could use their acquired knowledge for.

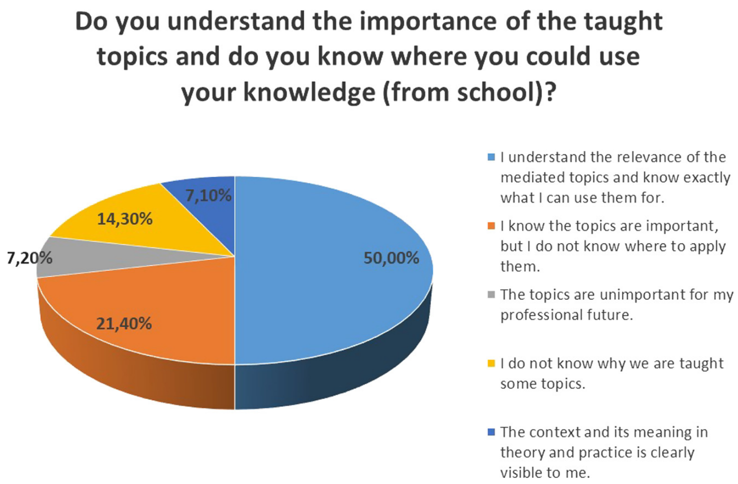


Fig. 5. Connection between theory and practice (pupils).

As could already be seen in the students' survey, 100% of the pupils said, that they wish for more practice-oriented classes in general school education. 96% of them said that it would motivate them more to progress in the acquired topics (Fig. 6).

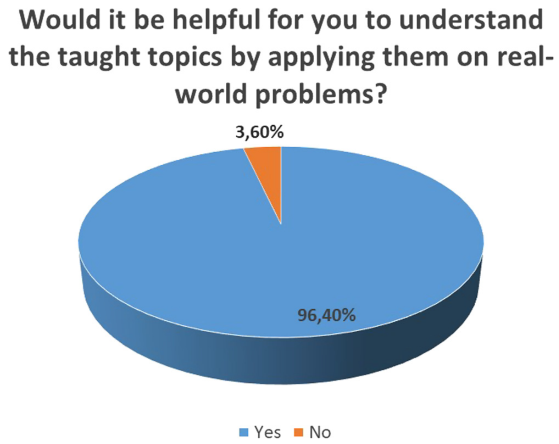


Fig. 6. Keep themselves busy with acquired topics (pupils).

4 Results

The results of the survey and the discussion with the stakeholders displayed interesting findings which will be detailed here. The pupils' survey revealed that the majority of the participants would support more practice-oriented teaching methods because applying the taught topics to real-world problems would help them to understand them in more detail. Nevertheless, all in all they felt well prepared for starting their studies. The student survey, on the other hand showed that most of the students did not feel very well prepared for an engineering study programme. They would also like to have a more practice-oriented teaching style as almost half of the participants were not sure how and where to use their gained knowledge.

Comparing both results it can be said that pupils seem to not have a specific expectation when they start their engineering study programme as they feel, as opposed to the students, well prepared for their studies. As both groups wish for a more practice-oriented and real-world based teaching style these difficulties could be overcome by trying to give pupils a more detailed insight into the practical application of their theoretical knowledge. Based on this they could make a more informed and balanced decision for certain areas of study and on which studies are suitable for them.

In accordance with the SPI Manifesto, the study was designed to create a learning organization, leading to a lifelong learning environment and enhancing successful business [18].

Both surveys were conducted in Germany. However, it is planned to question further pupils and students of European countries in the near future. This will place the research on a broader basis.

The in October 2019 finished EU-project ELIC as well as the several round tables with stakeholders, such as teachers, parents and members of STEM organization, displayed a similar picture, namely every party involved welcomes the exchange of experience and bundled knowledge. Here, a European perspective is already displayed. Figure 7 shows how the different institutions are involved in the engineering education process and their relationship with each other. These three institutions influence each other on various levels and are in a constant exchange.

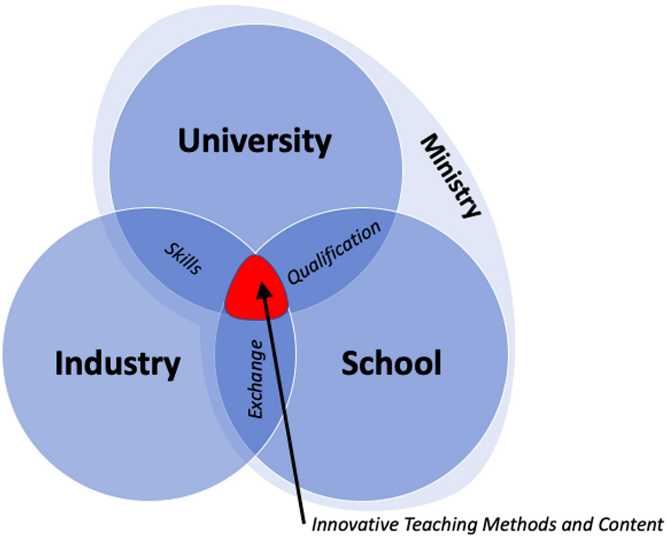


Fig. 7. Stakeholder in engineering education.

The Fig. 7 shows clearly that schools need an exchange with the industry: experts from different industrial branches can share their experiences in schools in order to show the students what to expect when they choose a career in the industry. The next step is the qualification that a pupil receives when graduating high school. With the industrial experts' expertise in mind, the pupils can choose which study field they want to pursue.

The key element of Fig. 7, located at its heart, are the innovative teaching methods and contents, which are mutually influenced by industry, university, school and the ministry responsible. If this combination of techniques and contents is implemented in the schools in an adequate and flexible way, it can satisfy every stakeholder's requirements for future engineering graduates. The lack of a practical and interdisciplinary approach in STEM-subjects can be counteracted by these new inductive teaching methods.

The review of relevant literature and the results of the survey as well as the discussion lead to the following recommendations:

1. **Problem Based Learning**
   Problem based learning, as already mentioned and described in the introduction of this paper, is an inductive teaching methods, where pupils and students mostly work in groups to solve real-world and open-ended problems [19].
2. **Flipped classrooms**
   This teaching method relies on the pupil's/student's independent learning behavior. Here they have to learn the topics on a self-learning basis and will apply them during lessons with the support of a skilled teacher [20].
3. **More experiments**
   For a better understanding more experiments can be helpful in order to visualize the taught topics.
4. **Experts from industry/university**
   Experts from industrial branches or from universities can share their experiences in school/university in order to prepare students and pupils on what to expect from their future work life [13].
5. **Curriculum update/space**
   By working together with the responsible ministry, the current curriculum has to be updated in order to give it a more practical focus. Furthermore, an interdisciplinary nature and more space for new teaching methods should be implemented.

5 Conclusion

The survey and the results of the almost finished EU project show that there is a tremendous need to adopt the engineering education on many levels. Referring to the questions asked at the beginning of the study, namely how STEM-subjects influence the study choice, the study success and how the focus on practical application can be rated, the survey and the EU project gave a clear picture, and show that changes are necessary. The survey took place at the Duesseldorf University of Applied Sciences and therefore most of the data used is influenced by the experiences made in German schools and universities and consequently cannot be adopted to other European countries one-to-one. However, data taken from EU project ELIC showed, that there are similar problems in other countries such as a lack of practical approach in the teaching content.

Further improvements could be achieved by reworking the curricula for teachers in order to leave more space for interdisciplinary content as well as a more practice-based education. The interviews with teachers of STEM-subjects showed that they already have ideas for improvements; however, in many cases a lack of time and money make improvements difficult. Some of the change, especially when dealing with the curricula, will take time and will not happen overnight. Research has shown that there are new ways of teaching, which have been proven efficient. In order to achieve the common goal of adapting the engineering education to the requirements of today's working environment several organizational, pedagogical and content-related steps have to be taken.

References

1. Niemann, J., Fussenecker, C., Schlösser, M., Ahrens, T.: ELIC – teachers as a medium to build a new generation of skilled engineers. In: International Conference on Competitive Manufacturing (COMA) Conference 2019, Stellenbosch, South Africa (2019)
2. Lee, J., Bagheri, B., Kao, H.: A cyber-physical systems architecture for industry 4.0-based manufacturing systems (2014)
3. von Solms, S., Nel, H.: STEM project based learning: towards improving secondary school performance in mathematics and science. In: 2017 IEEE AFRICON (2017)
4. Dolmans, D., De Grave, W., Wolfhagen, I., Van der Vleuten, C.: Problem-based learning: future challenges for educational practice and research. Med. Educ. **39**, 732–741 (2005)
5. Felder, R.M., Silverman, L.K.: Learning and teaching styles in engineering education. J. Eng. Educ. **78**, 674–681 (1988)
6. Chandrasekaran, S., Stojcevski, A., Littlefair, G., Joordens, M.: Project-oriented design-based learning: aligning students' views with industry needs. Int. J. Eng. Educ. **29**, 1109–1118 (2013)
7. Mahanija, Md.K.: Evaluation of an engineering program: a survey to assess fulfillment of industry requirements (2009)
8. Richert, A., Shehade, M., Plumanns, L., Groß, K., Schuster, K., Jeschke, S.: Educating engineers for industry 4.0: virtual worlds and human-robot-teams: empirical studies towards a new educational age, pp. 142–149 (2016)
9. Dahlgren, M.A., Castensson, R., Dahlgren, L.O.: PBL from the teachers' perspective. High. Educ. **36**, 437–447 (1998)
10. Capraro, R.M., Slough, S.W.: Why PBL? Why STEM? Why Now? An introduction to STEM project-based learning: an integrated science, technology, engineering, and mathematics (STEM) approach (2009)
11. FH Joanneum, ELIC – Engineering Literacy Online. In: Strategic Partnership for School Education, KA201, ELIC, Erasmus+ Programme, 2017-1-AT01-KA201-0305034 (2017)
12. European Commission: Life Long Learning Programme. http://ec.europa.eu/education/lifelong-learning-programme_de. Accessed 30 June 2019
13. ELIC, Note Taking Form, Round Table Germany. Intercontinental Düsseldorf, Germany, 21 May 2019

14. Qualitäts und Unterstützungsagentur NRW: Kernlernplan für Sekundarstufe II Gymnasium/Gesamtschule in Nordrhein-Westfalen - Physik. https://www.schulentwicklung.nrw.de/lehrplaene/upload/klp_SII/ph/KLP_GOSt_Physik.pdf. Accessed 01 July 2019
15. Qualitäts und Unterstützungsagentur NRW: Kernlernplan für Sekundarstufe II Gymnasium/Gesamtschule in Nordrhein-Westfalen - Mathematik. https://www.schulentwicklung.nrw.de/lehrplaene/lehrplan/47/KLP_GOSt_Mathematik.pdf. Accessed 01 July 2019
16. Fachbereich Maschinenbau und Verfahrenstechnik - Hochschule Düsseldorf: Gender Diversity Action Plan (2019)
17. Schneider, L.: Frauen erobern den Ingenieurberuf. https://www.ingenieur.de/karriere/arbeitsleben/frauen-erobern-den-ingenieurberuf/. Accessed 01 July 2019
18. The SPI Manifesto, Alcala, Euro-SPI 2009. http://2019.eurospi.net/images/eurospi/spi_manifesto.pdf. Accessed 18 May 2020
19. McQuade, R., Wiggins, S., Ventura-Medina, E., Anderson, T.: Knowledge disagreement formulations in problem-based learning tutorials: balancing pedagogical demands with 'saving face'. Classroom Discourse **9**, 227–243 (2018)
20. El-Senousy, H., Alquda, J.: The effect of flipped classroom strategy using blackboard mashup tools in enhancing achievement and self-regulated learning skills of university students. World J. Educ. Technol.: Curr. Issues **6**, 144 (2017)

Recent Innovations

Democratizing Innovation in the Digital Era: Empowering Innovation Agents for Driving the Change

Andreas Riel[1,2(✉)], Richard Messnarz[2], and Bruno Woeran[3]

[1] Université Grenoble Alpes, CNRS, Grenoble INP, G-SCOP, 38031 Grenoble, France
andreas.riel@grenoble-inp.fr
[2] ISCN GesmbH, 8010 Graz, Austria
{ariel, rmess}@iscn.com
[3] Oulu Business School, 90570 Oulu, Finland
bwoeran@gmx.eu

Abstract. In the digital era, innovation power and speed have become the lifeblood of every successful industrial organization. While driving forward and managing innovation has been a subject mainly related to top-level decision makers and innovation managers for a long time, increasingly many organizations realize that this is inappropriate for keeping pace with the rapid changes in today's societal, economic and ecologic context. Instead, leveraging and pushing innovation from bottom-up by mobilizing the creativity and diversity of the entire workforce has become the lingua franca of innovation in the digital era. An essential part of an organization culture fostering the democratization of innovation is the appropriate empowerment of the workforce through the understanding of essential drivers of change as well as the key innovation megatrends that are about to transform society. This position paper attempts to compile a comprehensive set of skills qualification programs ought to cover, which are aimed at empowering a workforce to act as innovation agents on all levels and organizational positions.

Keywords: Innovation training · Innovation culture · Innovation management · Business model innovation · Change management · Digitalization

1 Introduction

Against the background of the highly networked, interconnected, and therefore complex and technology-intensive ecosystems that we have been building up over the last decade, continuous innovation power has become the lifeblood of every successful industrial organization. More than ever before, short innovation cycles and rapid change driven by complexly interlinked change drivers are challenging and transforming not only every single industry sector, but the entire social, economic, and ecologic ecosystem. As several key decision makers of European automotive manufacturers and suppliers put it, "The next ten years will see more changes than the last

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 757–771, 2020.
https://doi.org/10.1007/978-3-030-56441-4_57

one hundred years have ever seen". In such a context, innovation can no longer be driven and managed by a few decision makers and innovation managers only. Instead, innovation must become an integral part of company's DNAs and with that, their organizational cultures. This is only possible by empowering a huge quantity of employees from several organizational units and hierarchy levels for taking an active, entrepreneurial part in the company's innovation initiatives.

While this is a fact largely confirmed by several studies and analyses of global innovation leaders, decision makers in Europe are lacking ways of implementing this requirement at a large scale in their organizations. One decisive reason for this is the lack of specialized training programs that teach concrete facts, challenges and good methods and practices of addressing them in their workforce's daily working practice.

In order to point this out, this paper is organized as follows. Section 2 explains the motivation of this paper, and the methodology applied to come up with the position statements and proposals it contains. Sections 3 to 8 propose the units and elements of competence we propose for Innovation Agents in Europe. Section 9 relates this work to the SPI Manifesto. Finally, Sect. 10 concludes and gives an outlook to related future activities.

2 Motivation and Methodology

The motivation for this position paper comes from the authors' close collaboration with major European industry, as well as their active participation in strategic European initiatives such as the EU Blueprint projects ALBATTS [1] and DRIVES [2]. We observed the need for an Innovation Agent qualification program that helps empower a broad range of workforce at all hierarchy levels and all industry sectors to actively inspire and push forward innovation in industrial organizations within the context of digital transformation and Circular Economy.

While innovation management has recently initiated its own body of knowledge [3], has become an internationally recognized job role with associated qualification programs over the last decade [4], the role of innovation agents has received much less attention so far. Professional training programs are still rare, e.g. [5, 6], and are mostly still focused on traditional innovation concepts and skills. Considering the fundamental transformation key industry sectors are undergoing, however, it quickly becomes evident that these traditional concepts and skills—although still relevant and valid—are not sufficient anymore to understand the full picture and to actively contribute to the key drivers of change. We motivate this statement by very briefly characterizing major changes that are ongoing in three primary industry sectors in Europe.

2.1 Automotive Industry Moving into Mobility Services

Astonishingly, one of the currently largest IT organizations in Germany is VW Group IT. This is due to the fact that they manage data and services across the entire vehicle fleet, as well as related tools and analysis results on several levels. The difference between data and knowledge is that knowledge can connect data by a context. This knowledge creation by a context can create new business scenarios. Furthermore, the cloud is not only used in large companies but rather in the open network and can therefore help connect multiple businesses and sectors to form unprecedented alliances.

In the future, vehicles will be a part of the data cloud and make decisions depending on a huge multitude of data coming from various distributed sources. This unlimited amount of data shared among cars, the environment and personal data in mobile phones and profiles of drivers can be combined to new services, such as e-mobility services, personalized services, workplace services etc. [7, 8]. In order to be deployed successfully at a European scale, all these new services need promising sales scenarios and innovative business models fitting into the societal, economic and ecological megatrends of a Sharing Economy and a Circular Economy.

2.2 Medical Industry Engaging into Healthcare Services

Medical data almost on a global scale are private, highly confidential, and therefore need to be protected. Patients giving their consent, connected medical data will open completely new models of healthcare service models. E.g. in times of pandemic, mobile devices could measure patients' characteristic health indicators such as body temperature and heart rate and feed these data to the (health) cloud. Once in there, these data could be combined with data provided by mobile devices of people in the close neighborhood, allowing to quantify the risk of exposure to already infected people. This kind of service would give a significant contribution to better controlling pandemic situations particularly in highly populated (typically urban) spaces. Such services could also largely facilitate healthcare operations of aged and handicapped patients needing a caretaker's help particularly in emergencies.

2.3 Battery Industry Moving into Energy Services

In the European strategic sector skills alliance Blue Print project ALBATTS [1], the EU currently supports the specification of competences, skills and job profiles required for the creation of energy systems and a battery production chain in Europe. This requires building up new business scenarios and finding new business models that allow Europe to compete with the Asian manufacturers. The sophisticated collection of data across all energy clusters in Europe can facilitate services and business models leveraging the creation and sustainable operation of the alternative energy value chain in Europe. In such a value chain, green energies obtained in Europe would feed green batteries produced in Europe and driving forward the new mobility paradigm across the European territory. Big data containing valuable information about stocks, supply chains, user profiles, etc. could be shared by context among all European cleantech energy clusters to form joint strategies to harvest and provide green energies all the way to the last mile to end users and incite the latter to follow the urgent megatrend of sustainability.

All three innovation domains cited above have in common that a profound transformation of the involved sectors is required and already ongoing. These transformations need the particularly strong active engagement of a huge number of employees engaged in industrial organizations, both in their professional (as creators of innovation) and private environments (as providers and consumers of data and energy).

Making them understand the key drivers of the paradigms they are supposed to push forward, as well as the main modern facilitators should precisely be the principal objective of the Innovation Agent's qualification that we propose here.

3 Qualification Curriculum Strategy

The topics that the proposed Innovation Agent qualification needs to cover are numerous, diverse, interdisciplinary and interconnected. Therefore, we consider it essential to adopt a strategy for clustering the desired competences, and to define skill profiles expected from job roles that are representative for the targeted industrial organizations. We decided to adopt specification schemes that have evolved from key European initiatives, ECQA [9] and DRIVES [7]. Both also have strong roots in the three industry domains that are in our focus as of Sect. 2.

In DRIVES [2, 7], the European ESCO [10] format for describing job roles is applied. This requires a structure of competencies and skills, and the assignment of a level to a competence. The levels of competence used are

- Awareness: People can repeat the knowledge and are aware of it in their daily work.
- Practitioner: People can apply the knowledge in their daily work.
- Expert: People can extend and transfer the knowledge to solve complex problems in their daily work.

ESCO can easily be mapped to the ECQA format that foresees a structuring of competences in units and elements. An element (e.g. U1.E1 Identifying and Understanding Drivers of Change, see Table 1 and Table 2) signifies a competence containing skills (ESCO terminology) or performance criteria (ECQA terminology) denoting an individual's capabilities, see Table 1 below.

Table 1. Performance criteria example for the innovation agent competence element U1.E1.

Performance criterion	Evidence check: the student can demonstrate
INNOAGENT.U1.E1.PC1	The student knows drivers of change and can enumerate related smart technologies
INNOAGENT.U1.E1.PC2	The student knows key initiatives identifying change drivers and can explain the drivers promoted them

Table 2 shows the set of skill units and elements we propose that Innovation Agent qualification programs should cover, as well as their desired competence levels per job roles related to one of the three key activity areas in the value creation process chain: design/development, production/supply/service provision, organization/management. Each of these skill elements will be described in Sect. 4 with respect to their essential motivations and objectives.

Table 2. The proposed innovation agent skills set.

Units (U) and Elements (E) of the skill card	Innovation agent development	Innovation agent organisation & management	Innovation agent production
Unit 1 Understanding Change Drivers and Building Scenarios			
U1.E1 Identifying and Understanding Drivers of Change	Awareness	Awareness	Awareness
U1.E2 Building an Innovation Vision and Future Scenarios	Awareness	Awareness	Awareness
Unit 2 Continuous Idea Generation and Realisation			
U2.E1 Identification and Deep Understanding of Innovation Needs and Opportunities	Expert	Practitioner	Practitioner
U2.E2 Idea Generation, Shaping and Evaluation	Expert	Practitioner	Practitioner
U2.E3 Idea Realisation in Complex Industrial Ecosystems	Expert	Practitioner	Practitioner
U2.E4 Idea Market Deployment in Data- and Service-driven Ecosystems	Practitioner	Expert	Practitioner
Unit 3 Agile Learning Organisations and Open Innovation			
U3.E1 Core Competence Identification and Evolution	Awareness	Expert	Awareness
U3.E2 Dynamic Learning Cycles in Open Innovation Ecosystems	Awareness	Expert	Awareness
U3.E3 Open Innovation Organisation and Process Design	Awareness	Expert	Awareness
U3.E4 Agile Transformation and Networked Agility	Practitioner	Expert	Practitioner
U3.E5 Risk Management and Resilience	Practitioner	Expert	Practitioner
Unit 4 Modern Innovation Principles, Paradigms, and Strategies			
U4.E1 Open Networking, Sharing Economy and Circular Economy	Practitioner	Expert	Practitioner
U4.E2 Cyber-Physical Product-Service Systems, Servitisation, and Service Innovation	Expert	Practitioner	Expert
U4.E3 Big Data and Cloud-Enabled Business Models	Practitioner	Expert	Practitioner
U4.E4 Smart Systems-of-Systems and Enabling Business Models	Practitioner	Expert	Practitioner
U4.E5 Value-added Smart Manufacturing and Enabling Business Models	Awareness	Practitioner	Expert
Unit 5 Case Studies			
U5.E1 Mobility	Practitioner	Practitioner	Practitioner

(*continued*)

Table 2. (*continued*)

Units (U) and Elements (E) of the skill card	Innovation agent development	Innovation agent organisation & management	Innovation agent production
U5.E2 CleanTech/Energy	Practitioner	Practitioner	Practitioner
U5.E3 Healthcare	Practitioner	Practitioner	Practitioner

4 Understanding Change Drivers and Building Scenarios

4.1 Identifying and Understanding Drivers of Change

Knowing, understanding and interpreting drivers of change in the context of a particular industrial organization has become the key success factor for innovation leaders in the long run. The reason is again rooted in the complexity of modern ecosystems. Nobody is able to predict even the near future anymore, because mutual interdependencies of influencing factors are overly complex. Who predicted the hilarious speed of transformation in the automotive sector? Who predicted the Covid-19 crisis? Who can predict movements at the stock exchange? Therefore, decision makers need to align their strategies and decisions with key drivers of change, since they are sure to last for at least the decade to come. It is not sufficient, however, that only decision makers at the top hierarchy levels develop this insight. In the end, their workforce needs to understand and follow. For this to happen, they need to be made aware to understand the key concept of drivers of change, as well as their significance and impact. An Innovation Agent training program shall teach those key drivers from the economical, ecological, social, as well as technological perspectives. Political and legislative trends on national and international levels shall also be discussed in the particular context of selected target domains.

4.2 Building an Innovation Vision and Future Scenarios

The decade 2020–2030 will be characterized by a fundamental transformation of society in both economic and ecological aspects that keeps going on at a tremendous speed and affects every industrial sector. In today's highly interconnected world, we have created technical, economical and societal ecosystems whose complexities have become extremely difficult to manage. Consequently, it is practically impossible for anybody to predict how even their near future is going to look like. The current Covid-19 pandemic is a tragic, however, outstanding manifestation of our modern ecosystems' vulnerabilities and volatilities. When it comes to developing and continuously adapting innovation strategies in such complex contexts, the development of scenarios has proven successful at a large scale. Teaching those, shall cover concrete innovation scenarios for selected key industries in Europe such as mobility, energy and healthcare. These scenarios shall be based on insights and results collected from recent European and international studies addressing in particular the main drivers of change, megatrends, politics, as well as technology and legislation push and market pull.

5 Continuous Idea Generation and Realization

5.1 Identification and Deep Understanding of Innovation Opportunities

Every problem, shortcoming and inconvenience represents an opportunity for innovation. Well analyzing and deeply understanding such opportunities is the fundamental prerequisite for any fruitful downstream innovation process. This element shall teach methods and tools for deep problem understanding and continuous opportunity identification, with a particular focus on Design Thinking and special scenario techniques [11].

5.2 Idea Generation, Shaping and Prioritization

Being able to exploit every employee's and partner's creativity potential on an everyday basis is one of the crucial success factors that help breed an organizational culture of permanent innovation that is essential to stay at the forefront of competition. Generating and shaping ideas should no longer be reserved to a small number of "privileged" innovation managers and think tankers, but rather a daily natural element of a workforce (or on a daily basis). The evaluation and prioritization of the collected ideas has to be transparent and according to clearly defined criteria that have to be known to the contributors at every time. Contributors need time and creative freedom to shape their ideas against these criteria in order to maximize their contributions' relevance to their organization's strategic priorities [12]. Methods, models, techniques, organizations and process elements leveraging this idea generation, shaping and evaluation process shall be at the heart of this element.

5.3 Idea Realization in Complex Industrial Ecosystems

One of the biggest challenges traditional organizations are facing, is increasing the throughput of ideas to successful market deployment. Corporate entrepreneurship (sometimes also called "intrapreneurship") signifies the need of employees to adopt and continuously develop an entrepreneurial spirit in order to push their idea forward and let them grow in an organizational context that is increasingly complex [13]. Developing and permanently updating an entrepreneurial strategy taking into account the specific organizational, political, legal, and market context is crucial in such modern environments. Without the appropriate skills for mastering related methods and tools, intrapreneurs will quickly lose motivation and drive, negatively influencing the company's organizational culture and innovation power. Creating awareness for this and teaching inter- and intra-organizational innovation methods and strategies shall be the core subject of this element.

5.4 Idea Market Deployment in Data- and Service-Driven Ecosystems

While traditional economies relied primarily on the selling of products and consumer product ownership, modern and future business scenarios need to emphasize the creation of consumer experience by a complex interplay of high-tech products with the data cloud, and related service providers. Hence, deployment strategies and processes in such modern economic and social environments differ significantly from what has been known so far. Not only new strategies, ecosystems, and business models must be created, but also the concerned enabling product and data chain needs to be designed for the provision of value-adding services creating unique customer experiences while still being provably compliant with relevant legislation and politics. While creating numerous new opportunities for higher-margin businesses, these new strategies will even become a necessity in the very near future, latest at the moment when high-tech smart products will become too expensive for mass users to afford. This will imply the slow but certain total replacement of the traditional product ownership paradigm by service models and ecosystems of a sharing economy (cf. BlaBlaCar, AirBnB, etc.).

This element shall teach essential concepts and success factors driving the integrated co-design of high-tech products and smart services in ways that services at various quality and experience levels can be sold to consumers all along the (closed) service lifecycle. The key role of harvesting and capitalizing on "Small Data" (i.e., the subset in the Big Data cloud that is of essential relevance for the provided services) shall be explained, both for enabling services and for providing (data) resources to other service providers (e.g., real-time environmental data captured by the diverse sensor network in current and future cars can be used not only for letting cars drive autonomously with the best possible driving experience, but they can also be sold to service providers outside the automotive sector, such as weather forecast centers, road network supervision and maintenance centers, etc.).

6 Agile Learning Organizations and Open Innovation

6.1 Core Competence Identification and Evolution

Core competence analysis is a traditional company strategy tool helping decision makers identify the actual core domains of knowledge, skills, and experiences which they have gained real competitive advantage in. These core competences therefore, represent the strategic pillars of the organization, which they shall build their innovation roadmap upon. Nowadays and in future, however, these core competences evolve and are linked with other competences much more rapidly than in the past. In particular, the IoT megatrend leads to electronic hardware, software, and internet connectivity integrated in almost every product, even in the most traditionally mechanical ones. In combination with the fact these products increasingly take over safety-critical and vital functions our everyday lives depend on, almost every sector will have to decide where to position related competences in hardware development, software development, as well as functional safety and cybersecurity. Innovation leaders master the continuous evolution, extension and adaptation of their core competence strategies and deployments better and in a more flexible and agile way than their less successful competitors. One of the key success factors we can derive from them is that they involve employees of several organizational units and hierarchy levels in the core competence analysis and evolution strategy rather than only high-level decision makers. Therefore, this element shall teach core competence analysis and evolution principles in a way that a larger part of the employee basis can actively take part in the related activities and processes.

6.2 Dynamic Learning Cycles in Open Innovation Ecosystems

The dynamic learning organization is a fundamental innovation principle enabling corporate organizations to dynamically adapt to the continuously evolving context in terms of market demands, consumer behavior, politics, legislation, etc. in a way like a living organism. While implementing dynamic learning cycles in a particular organization has always been a challenge, doing so nowadays is even more difficult. The reason for this is the need for creating Open Innovation ecosystems [14], which means that a multitude of external and internal partners (distributed across the entire extended, connected enterprise) must be integrated in a company's innovation strategy, models, processes and operations. In general, industries in Europe needed more than 15 years to adopt Open Innovation principles, and some are still struggling to fully capitalize on the power of innovating in networks rather than within company walls only. Open Innovation requires an open, collaborative mind-set across the entire organization. This element shall teach the main characteristics of organization cultures that successfully implemented dynamic learning cycles in an Open Innovation spirit. It shall elaborate on the principles of Open Innovation, and how these can be made an engine for the realization of disruptive ideas, as well as for boosting the organization's responsiveness to change.

6.3 Open Innovation Organization and Process Design

Innovating in an Open Innovation environment needs to be reflected both within the enterprise organization and within their processes. The workforce needs to understand where and how a close collaboration with innovation partners and suppliers is expected and how it shall happen. Process definitions shall be explicit about critical interfaces in this co-innovation ecosystem, and about where each party's responsibilities are. Related process definitions shall, however, still leave enough freedom for the networked workforce to organize themselves best according to instant needs and get empowered for using their collective creativity to respond to innovation challenges. This element shall elaborate on this aspect from the perspectives of both process owners and process agents.

6.4 Agile Transformation and Networked Agility

The agile transformation of traditional organizations has received a lot of attention in the recent few years. Companies have been facing the need of developing the capability of responding to change much more quickly than they used to do. This trend is not only going to continue but it will even get considerably stronger given the highly interconnected and complex ecosystems that we have been creating in the digital age. Despite the asset of collected experiences from the recent past, creating agile organizations remains a challenge, especially when it comes to size and legacy. Highly networked organizations practicing Open Innovation add another level of complexity to this, since agility needs to be replicated across enterprise boundaries with limited influence of a central controlling instance. In such an ecosystem, only the workforce itself can achieve connected, collective agility, driven by change agents replicating the agile mindset via their internal social networks. Established agile frameworks like Scrum and SAFe include some of these aspects. However, they lack guidance and do not scale in innovation networks [15]. This element shall elaborate on this subject and teach pragmatic solutions derived from best practice in industry.

6.5 Risk Management and Resilience

Just as in finance, sports and many other domains, moving fast and changing quickly and permanently implies a lot of risks and volatility. Managing risk and building up a solid level of resilience to permanent uncertainty, triggers of change, and potential crisis has become one of the most vital capabilities of industrial innovation leaders. While Open Innovation networks contribute to resilience, they make risk management more complex. Finding the right balance between the two has become an art and life insurance at the same time. This element shall teach core principles of modern risk management across the entire innovation life cycle and ecosystem, emphasizing the role of employees as innovation agents in this process.

7 Modern Innovation Principles, Paradigms, and Strategies

7.1 Open Networking, Sharing Economy and Circular Economy

The “Always On” society that is permanently connected through the ubiquitous internet has a huge impact on our economic ecosystems. Open, global networking principles have become a vital trait of successful organizations, as well as of consumer behavior. Platform companies such as BlaBlaCar, AirBnB, FlixBus and many others have made ubiquitous internet access the backbone of capitalizing on existing resources by making them available to everybody in the world at a fingertip. In this way, the new concept of Sharing Economy is developing rapidly and reshaping traditional consumer behavior and values. Furthermore, as climate change and resource scarcity have become tangible in people’s everyday lives everywhere on this planet, consumers’ awareness for the urgent need of changing behavior is gaining more and more influence in the way industry is supposed to create value. Closing the product lifecycle by various “Re-strategies” (re-using, re-remanufacturing, re-purposing, re-cycling) has become a top-priority even in traditional resource-intensive industrial organizations. Far beyond a legislative “must” only, the Circular Economy is about to become a major strategic workhorse for the fourth industrial revolution. Circular economy supported by data economy and an IoT infrastructure enables the extension of the traditional manufacturing operation to the entire product life cycle.

7.2 Cyber-Physical Product-Service Systems and Service Innovation

The key enabler of service-oriented economy and innovation is connectivity and smart technologies leveraging continuous adaptation and learning capabilities in service-enabling. The design of such connected, intelligent, adaptive product-service-systems requires the research and deployment of new design principles, methods and tools. Not only data-driven machine learning algorithms confront traditional development organizations with unprecedented challenges, but there is also the need for designing and re-designing such cyber-physical systems all throughout their service lives. Currently, this becomes most evident in Over-the-Air-Update functionality (OTA) that increasingly many product designers must integrate into their products in order to assure their capability of continuously upgrading and adapting their behavior throughout the entire use phase, even including the end-of-life. While apparently simple and easy to specify, just this single add-on functionality confronts development organizations with design challenges that are new for them, in particular functional safety and cybersecurity [16]. How to avoid that an OTA introduces unexpected, safety-relevant misbehavior of the product? How to avoid any kind of malicious attacks before, during, and after the OTA operation? How to assure the compatibility of the updated products with the other connected products? How to re-certify and re-homologate the product after the update? These and many other design-related questions need to be solved before any design problems and weaknesses can lead to collateral damages via the IoT and the cloud. This element shall teach the essential concepts of successfully designing cyber-physical product-service-systems, and create awareness for the particular challenges and partly yet unsolved problems. Since the keyword at the center of these challenges is multi-disciplinary, multi-competent design teams, the Innovation Agent training shall empower a large basis of designers from different disciplines (technical, legal, environmental, medical, etc.) to give their active contributions to modern product-service-systems innovation.

7.3 Big Data and Cloud-Enabled Business Models

A natural consequence of Service Economy gradually replacing Product Economy through networked Open Innovation ecosystems is that traditional business models will no longer be adequate to be successful on the market. Traditional automakers must move from selling individual cars towards mobility experience. Traditional centralized electrical power station and network operators have to transform themselves into green energy providers. Traditional medical device manufacturers must learn using patients' health data in the medical cloud to provide optimal and distanced healthcare services. Apart from the challenges to the design processes and organization outlined earlier, the design and deployment of innovative service and cloud-oriented business models brings along huge difficulties and risks as well. A new business model, once designed, can hardly be prototyped for experimentation like a product. It rather requires the establishment of a lot of legal contracts, the deployment of new organizations and IT infrastructures, and the like. Either it works, or the company (and with it, probably the entire network of suppliers and service partners) will fail. This element shall elaborate on these subjects that are absolutely crucial for Europe's economic future. Every employee needs to understand that their organizations do not even have the choice. It is imperative for them to follow this trend if Europe does not want to be left behind and lose its position at the forefront of innovation. Practical trainings on modern business model innovation still being rare, in Europe, it shall be the Innovation Agent qualification's key mission to fill this particular gap.

7.4 Smart Systems-of-Systems and Enabling Business Models

This element shall extend the previous one by taking a focus on Business Model Innovation for deploying intelligent system-of-systems on the European markets. The trend that is addressed here is the fact that numerous sectors are becoming more and more dependent on each other. An outstanding example is the introduction of E-Mobility on the European market, which is still hampered to a large extent by the missing charging infrastructure. Likewise, autonomous driving relies on a ubiquitous, safe and secure real-time cloud infrastructure linking vehicles with road infrastructure (e.g. traffic lights), but also people. How to co-develop these systems of systems? How to cope with the differences in innovation cycles? How to manage the risks?

Even if most of related challenges are of political and legal nature and therefore need to be solved on higher, strategic levels, it is a proven fact that in the highly complex interconnected world we created, top-level decision makers increasingly depend on advice coming from a large multi-disciplinary group of domain-experts: Trained Innovation Agents will be able to fulfill exactly this crucial role.

7.5 Value-Added Smart Manufacturing and Enabling Business Models

Finally, Europe must not miss the big, unique chance to re-establish their formally leading manufacturing industry thanks to technological trends such as cyber-physical, interconnected manufacturing systems, additive manufacturing, as well as smart, flexible manufacturing systems [17]. Being pushed ahead by global mega-trends such as sustainability, customization, and Circular Economy, the Fourth Industrial Revolution (Industry 4.0) is going on at a tremendous speed from a technology-push perspective. Political, organizational and business model related aspects, however, are still lagging and endanger Europe's competitiveness with respect to countries like China and the USA in particular. Unless a large basis of highly capable, competent and still motivated employees gets trained in thinking and actively pushing in this direction, companies in Europe will face more and more difficulties in becoming and staying a key player in this industrial revolution that is about to transform society on a global level.

Apart from creating awareness for all these subjects, this element shall make Innovation Agents understand the core technical, social, economic, and ecologic aspects of Industry 4.0 in very concrete down-to-the-earth terms using case studies from real industry practice. It shall help them elaborate the basis of an Industry 4.0 strategy roadmap for their particular organizations and thereby move from generic buzzwords to concrete value-adding strategies.

8 Role-Based Training Design and Delivery

As pointed out in numerous groundbreaking innovation and creativity management publications, in particular [18], an effective and successful innovation culture requires different roles to be actively represented in any innovative organization. Democratizing innovation by scaling innovation training and other enablers to a large organization scale therefore also implies carefully taking into account innovation roles and related profiles in both content and approach to empowerment and teaching. To design and deliver the proposed Innovation Agent qualification, we propose the following four building blocks:

1) *Emotional* - *The behavioral innovator*. Human factors driving innovative behavior: typologies, learning and strategies.
2) *Educational* - *The innovation toolbox.* Principles, concepts, strategies, methods and tools: existing knowledge on innovation skills.
3) *Empirical* - *The innovation experience*. Use cases where the innovation toolbox has been applied with success or failure.
4) *Experimentation* - *The innovation journey (or endeavor).* A hands-on, coached module to experiment a selected innovation use case.

The idea is to let training participants discover their own interests and characteristic traits in order for them to be able to select the role(s) they will feel most comfortable and enthusiastic in. This is fundamental, since innovation, creativity and entrepreneurial spirit come from motivation, passion, and conviction.

9 Relevance to the SPI Manifesto

According to the SPI Manifesto [19], people are at the center of every process, every organization, as well as every change process. The digital era affects each and every sector in terms of innovation and change processes are going on much more rapidly than ever before. Moreover, the digital transformation is highly complex in terms of the numerous interdependencies among stakeholders and business processes it requires. Such speed and complexity can only be driven by a pro-active support and contribution of an industrial organization's workforce. Concentrating change power and strategy to just a few top-level managers is a model that has turned out to be inappropriate in such digital, networked ecosystems, and therefore bears a lot of risk.

In this sense this contribution also tries to strengthen the existing SPI Manifesto in a way that it highlights the increasing lively importance of giving power and motivation to the workforce in order for the change to happen in a continuous and sustainable manner. In addition to that, the proposed qualification is aligned with the need for cradle to crate lifelong learning concept that is increasingly vital for empowering workforce to understand and drive the change.

10 Summary, Conclusion and Outlook

This position paper elaborates on the urgent need for the establishment and large-scale deployment of an Innovation Agent qualification program that aims at empowering a large part of the workforce of companies in Europe to actively push forward innovation in the digital era while also controlling the direction it takes with respect to sustainability. Apart from the digital megatrend, concepts such as Circular Economy, Sharing Economy, as well as Product-Service Systems have to be well known and understood by the workforce in order for them to be able to take them into account in their daily activities. To this aim, the paper suggests a curriculum structure in terms of competence units composed of focused elements. The suggested content of these units and elements has been derived from the authors' exhaustive experiences in consulting, coaching and teaching major industry in Europe, as well as from EU Blueprint project studies.

Based on the EQF, it provides the essential design basis that facilitates its integration in diverse curricula in Europe, as well as Europe-wide certification. While large-scale enterprises are the main target group for the Innovation Agent qualification, we expect that SME's will also profit from cultural leverage effect that a broad and timely innovation facilitation training will have in every organization. We also pointed out the importance of a role-specific adaptation of the Innovation Agent curriculum in terms of depth of knowledge elements, as well as the emphasis on learner's personal characteristic traits and objectives. Only motivated learners will also be innovative!

Since, to the best of our knowledge, there is currently no comparable program available on European level, we are currently launching an initiative uniting representatives from European industry clusters, VET organizations, and universities in order to set up such a program. Our objective is to launch pilot trainings in one year to come into play and subsequently, to improve and extend the program continuously through teaching practice both on-site and on-line. Additionally, the pupils of today will be the employees, employers, and entrepreneurs of tomorrow. Teaching the young will feed the knowledge workforce we need in the not so distant future.

Acknowledgements. The authors sincerely thank Thomas Peisl, who significantly inspired Sect. 8. Some of the needs included in this position paper have been derived from the EU Blueprint projects ALBATTS [1] and DRIVES [2]. Both have been financially supported by the European Commission in the Erasmus+ Programme under the project numbers 612675-EPP-1-2019-1-SE-EPPKA2-SSA-B and 591988-EPP-1-2017-1-CZ-EPPKA2-SSA-B, respectively. This publication reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

References

1. ALBATTS EU Blueprint project. https://www.project-albatts.eu/. Accessed 22 Apr 2020
2. DRIVES EU Blueprint project. https://www.project-drives.eu/. Accessed 22 Apr 2020
3. International Standardization Organization, ISO 56000:2020. Innovation Management (2020)

4. Riel, A.: Innovation managers 2.0: which competencies? In: O'Connor, R.V., Pries-Heje, J., Messnarz, R. (eds.) EuroSPI 2011. CCIS, vol. 172, pp. 278–289. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-22206-1_25
5. Danish Agency for Science and Higher Education. The Innovation Agent Program. https://www.dti.dk/specialists/innovation-agent-program/31424. Accessed 22 Apr 2020
6. The Innovation Agent Academy. https://www.innovationagent.academy/. Accessed 22 Apr 2020
7. Stolfa, J., et al.: DRIVES—EU blueprint project for the automotive sector—a literature review of drivers of change in automotive industry. J. Softw.: Evol. Process **32**(3) (2020). Special Issue: Addressing Evolving Requirements Faced by the Software Industry. https://doi.org/10.1002/smr.2222
8. Neumann, M., Riel, A., Dismon, H.: Technology strategy planning and innovation management at Rheinmetall automotive to face future mobility challenges. In: Larrucea, X., Santamaria, I., O'Connor, Rory V., Messnarz, R. (eds.) EuroSPI 2018. CCIS, vol. 896, pp. 607–618. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-97925-0_51
9. Messnarz, R., Ekert, D., Reiner, M., Sicilia, M.A.: Europe wide industry certification using standard procedures based on ISO 17024. In: 2012 Technologies Applied to Electronics Teaching (TAEE), Vigo, pp. 342–347 (2012)
10. ESCO European Skills/Competences, Qualifications and Occupations. https://ec.europa.eu/esco/portal/home. Accessed 28 Apr 2020
11. Riel, A., Flatscher, M.: A design process approach to strategic production planning for industry 4.0. In: Stolfa, J., Stolfa, S., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2017. CCIS, vol. 748, pp. 323–333. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64218-5_27
12. Riel, A., Tichkiewitch, S., Neumann, M.: Structuring the early fuzzy front-end to manage ideation for new product development. CIRP Ann. – Manuf. Technol. **1**(62), 107–110 (2013)
13. Gebhardt, K., Maes, T., Riel, A.: A project management decision support tool for keeping pace with the dynamics of corporate innovation projects. In: Larrucea, X., Santamaria, I., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2018. CCIS, vol. 896, pp. 619–630. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-97925-0_52
14. Chesbrough, H.W.: Open Innovation. Harvard Business School Press, Boston (2003)
15. Poth, A., Kottke, M., Riel, A.: Scaling agile – a large enterprise view on delivering and ensuring sustainable transitions. In: Przybyłek, A., Morales-Trujillo, M.E. (eds.) LASD/MIDI-2019. LNBIP, vol. 376, pp. 1–18. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-37534-8_1
16. Riel, A., Kreiner, C., Messnarz, R., Much, A.: An architectural approach to the integration of safety and security requirements in smart products and systems design. CIRP Ann. – Manuf. Technol **67**(1), 173–176 (2018)
17. Armengaud, E., Sams, C., von Falck, G., List, G., Kreiner, C., Riel, A.: Industry 4.0 as digitalization over the entire product lifecycle: opportunities in the automotive domain. In: Stolfa, J., Stolfa, S., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2017. CCIS, vol. 748, pp. 334–351. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64218-5_28
18. Kelley, T., Littman, J.: The Ten Faces of Innovation: IDEO's Strategies for Beating the Devil's Advocate and Driving Creativity Throughout Your Organization. Crown, New York (2006). ISBN 9780385517010
19. Korsaa, M., et al.: The people aspects in modern process improvement management approaches. J. Softw.: Evol. Process **25**(4), 381–391 (2013)

Spin-Off Strategies of Established Companies Due to Digitalization and Disruption

Dominik Kretschmar(✉), Jörg Niemann, and Felix Müller

University of Applied Sciences Düsseldorf,
Münsterstr. 156, 40476 Düsseldorf, Germany
{dominik.kretschmar,joerg.niemann}@hs-duesseldorf.de

Abstract. Due to digitalization and disruptive innovations, many established companies are faced with the task of adapting their business models and processes to new requirements. To avoid risks and exploit opportunities, spin-offs are founded for this purpose.

The research question to be answered is if spin-off strategies can be based on a general form and if measures can be clustered. For this purpose, literature was evaluated and interviews with experts were conducted. In this way a structured approach could be developed, which is based on four result clusters: reasons, arguments, differentiation and success factors.

The results are particularly interesting for the management of spin-offs and for companies that want to set up one.

Keywords: Spin-off strategies · Digitalization and disruptive innovations · Business model innovation

1 Introduction and Research Approach

Digitalization and disruptive innovations are increasingly putting companies ahead of challenges. The megatrends of the 21st century - such as digitization, globalization, shortened product life cycles and a general deregulation of markets - contribute to the blurring of boundaries between industries [1]. Established companies in particular are forced into action by inertia and bureaucracy. On the one hand, digitization can optimize a wide variety of business processes and thus exploit great potential [2]. On the other hand, companies can quickly lose touch with the flexible markets of the digital age. In this context, Porter has predicted that in the future traditional business models will be eroded by the sum of new technological possibilities [3]. "Accordingly, business models will have to be adapted or even completely redeveloped in almost all industries [4]. Already in 1997 this problem was titled the "innovation dilemma" [5].

Today, established companies therefore often face this dilemma with the tactic of beating the innovative competitors on the market with their own weapons by founding spin-off companies including promising start-up ideas [6].

For this purpose, the research question is to be clarified if spin-off strategies of established companies, as an approach for digitization and innovations can be based on general forms and measures can be clustered?

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 772–782, 2020.
https://doi.org/10.1007/978-3-030-56441-4_58

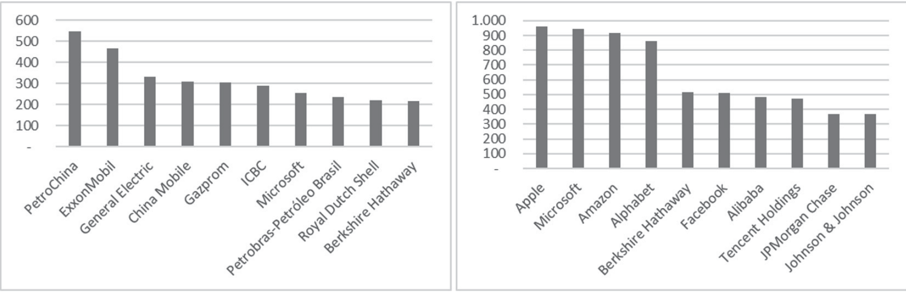


Fig. 1. Market value of the world's largest public companies in $B in 2008 (left) [7] and in 2018 (right) [8]

If we look at the most successful companies worldwide (see Fig. 1), it becomes clear that especially those companies that have technology and digitization as their main business are particularly successful. Technological progress is changing the conditions for success. In the last decades, the majority of the successful companies have shifted away from the gas and oil industry to the data industry. If we look at the absolute market value of the technology groups, an extreme development can be seen here. This shows that digitization and innovation are revenue drivers and that an adequate business model has a significant impact on the development of a company.

These developments suggest that there is a need for established companies to be innovative and technology oriented. Established business models are already under pressure from disruptions or innovators. Beside this new business models [9–11], industry 4.0 [2], shorter product life cycles [12, 13], servitization [14], and neo-ecology [15] are major challenges.

In accordance with the SPI Manifesto, the study was designed to enable change while managing risk to support sustainable business success [16].

1.1 Research Approach

The procedure is mainly divided into three sections (see Fig. 2): The analysis of the existing literature and approaches, the verification by expert interviews to conclude a procedure and a cluster of methods based on the results.

The approach of a systematic literature review is based on the recommendations of Henriette, Feki, and Boughzala [17]. For this, six steps must be completed: research identification, research strategy, study selection, quality assessment, data extraction, data synthesis and analysis. The criteria for including a literature source are that the full text is available in German or English in Google Scholar, Wiley Online Library or the library of the University of Applied Science Duesseldorf, published from 2018 onwards, it have to be an academic journals, books or publications, the content requirements are given by specific search terms referring to the conducted literature review. It looks on the implementation and adaptation by established companies through spin-off structures for innovations like products, services and business models. The search terms are listed in the respective chapter.

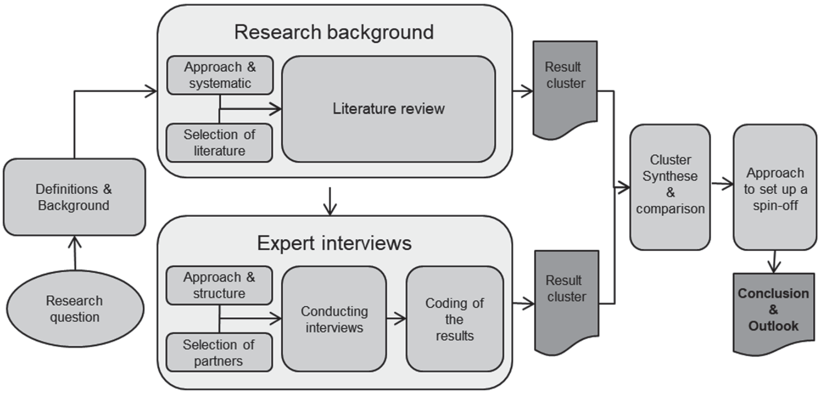


Fig. 2. Research approach

The interviews were conducted in a semi-structured, qualitative format and the interview guidelines provided the opportunity to answer freely within a given framework. The survey was divided into three blocks of questions: 1. person-related information and topic-specific background, 2. spin-off process and growth/start-up phase, and 3. advantages and disadvantages for the spin-off and the parent company.

To evaluate and analyze the interviews, they were transcribed and examined in terms of content, with the aim of forming result clusters that can be compared with the results of the literature research. For this purpose, coding guidelines according to Charmaz were applied [18].

The term “established company” is not clearly defined, but is frequently used (see also [19, 20]). For the purpose of this paper, companies are understood to be those that have been successful on the market for at least ten years and offer their products with an almost unchanged business model.

1.2 Literature Review: Spin-Off Structures of Established Companies for Development and Adaptation of Innovations

The literature review looks at spin-off structures of established companies for development and adaptation of innovations like, products, services and business models. The used search terms (incl. German translations) are: Structure, Spin-Offs, Established Companies, Corporate Spin-Off, Development, Adaption, Product Development, Service Development and Business Model Development. Five of the twenty-two sources identified met all criteria and could be summarized in four clusters (Table 1):

Table 1. Four cluster of measures for spin-offs

Title	Description
Cluster 1	*Reasons for setting up a spin-off* [4, 21–23]
New challenges	New business models, products and services must be developed
Ambidextry	– Exploration and exploitation must be realized simultaneously – (non-core business) Spin-offs can guarantee exploration
Cultural backgrounds	– Unused, entrepreneurial possibilities – Inertia and bureaucratic hurdles in established companies
Cluster 2	*Advantages of setting up a spin-off* [4, 22–24]
Enabler creativity	– Offer room for openness and creativity – Enable legal and physical separation from parent companies – Benefits from the network and name of the parent company – Necessary spin-offs through necessary restructuring or similar – Internal projects can be capitalized (Luc, Filion, & Fortin, 2002) – Enable the use of their technical know-how
Cluster 3	*Opportunities for differentiation* [4, 23, 24]
Organization and cooperation	– Company structure, company culture – Cooperation with other companies
Degree of independence	– Sponsored (Low-selective, supportive, incubator) – Not sponsored
Necessity and origin drivers	– Necessity vs. opportunity – Top-down vs. bottom-up – Restructuring vs. entrepreneur
Transfer of know-how	– Transfer of intellectual property rights – Mobility of employees
Cluster 4	*Success factors* [4, 21–25]
Relationship to the parent company	– Clear communication and regulation with the parent company – Allocation of autonomy – Mutual provision of know-how – Provision of financial resources and networks by parent companies
Location and infrastructure	– Digital and technical infrastructure must be up-to-date – Distance to parent company must be chosen appropriately
Composition of the team	– Innovative ability of the team – Elimination of routines – Know-how of the management – Open and agile culture
Strategic behavior	– Creation of adaptable organizations – Agile and flexible reactions – Decentralized and flat organization – Business model innovation through unique selling proposition – Continuous validation of the USP and the business model

Apart from the small amount of suitable literature, some are not significant and shows that research in this area is still incomplete and therefore does not allow a clear recommendation for action. The following expert interviews will complement this in order to obtain a more comprehensive picture.

1.3 Expert Interviews

The one hundred largest companies in Germany were reviewed to find appropriate spin-offs using the criteria and characteristics in Table 2.

Table 2. Criteria and characteristics for an appropriate spin-off

Criteria	Specification
Size of the parent company	In Germany's Top 100
Type of innovation department of the parent company	Independent spin-off
Tasks and core competence of the spin-off	Leading the parent company into digitization

If the parent company is in the top 100 [7] and the spin-off acts independently[1], it was checked if the spin-offs provide the core tasks/competences appropriate to this paper. This primarily includes whether the spin-off will only sell its own products or whether it will support the parent company in digital transformation and challenges posed by digitization and disruptive innovation. This resulted in a list of 15 spin-offs (Table 3):

Table 3. Appropriate spin-offs

– Rewe Digital – Henkel X – OBI Next – Ergo Digital – Oetker Digital – Evonik Digital	– SMS Digital – Allianz X – Talanx Systeme AG – PXG Health Tech – Grow (Bosch) – Openspace (Comerzbank)	– Lufthansa Innovation Hub – Next 47 (Siemens) – Spielfeld Digital Hub (Visa/Roland Berger)

Due to the very specific requirements and the rejection of the companies to be interviewed, (consulting) companies, which work with spinoffs, were also included. The reasons for rejection were lack of time and too many interview requests, which underlines the relevance of the topic. Thus, a COO of a spinoff [26] and the managing director of a consulting company [27], which supports spinoffs, could be interviewed. Both have many years of professional experience and are entrusted with the development of strategy. Through the second interview, insights into many different spinoffs could be given.

[1] Due to manual evaluation of the information of the respective internet website.

1.4 Result Comparison of the Literature Clusters with the Interviews

The statements and criteria of the literature review are largely confirmed by the interviews and supplemented in various aspects. Especially cluster three and four are completed by multiple additions. Below, the four clusters from the literature review are compared and supplemented with the synthesized and aggregated results of the expert interviews [26, 27] with the respective criteria in order to develop a usable methodology in the following chapter.

Cluster 1 – Reasons for Setting Up a Spin-Off: In the first cluster, the statements of the interviews are close to the literature. New challenges result mainly from the pressure on established companies to develop new products, services or business models. The interviews confirm this point in several statements. In addition to this, it is often difficult for established companies to be innovative in conservative markets and the exploration with the help of a spin-off significantly reduces the risk for the parent company (see also Cluster 2). The next characteristic deals with the corporate cultural backgrounds that prevent established companies from acting innovatively. The consistent statement says that mainly bureaucratic hurdles and corporate inertia are the main reasons why entrepreneurial opportunities are not exploited.

Cluster 2 – Advantages of Founding a Spin-Off: The advantage that spin-offs offer more room for openness and creativity can be expanded by the experts by adding the attribute “speed”. The literature also argues that spin-offs can ensure legal and spatial separation from the parent company, which the experts also confirm and call “risk hatching”. A further advantage, which is frequently found in the literature, is the possibility for the spin-off to use the already existing networks and the name of the parent company. This point is supplemented by the experts by the possibilities of using personnel, funds, infrastructure and market access. The next advantage is the possibility for established companies to capitalize on internal projects. The interviews can add to this the fact that established companies can complete their portfolio of products and services through the spin-off and thereby develop new sales potentials. The final point of the second results cluster is the possibility of innovative use of the technical know-how of the parent company through the possible spin-off.

Cluster 3 – Opportunities for Differentiation: In the third cluster, the literature review identified four differentiation criteria. The first criteria describe the possibility of differentiating spin-offs according to their form of organization and form of cooperation. The primary focus is on the structure and culture of the company, but also on the cooperation of the spin-off with other companies. The interviews add the differentiation form of the types of development partnerships with customers. The second criteria look at the degree of independence of the spin-offs, with a primary distinction being made between whether the spin-offs are sponsored or not. Here, too, the type of sponsoring (financial, administrative or controlling) is supplemented and it is noted that sponsoring restricts the spin-offs’ degree of freedom, which is critical for success. The third criteria, which describes the necessity and origin driver as a differentiating feature of spin-offs, is fully confirmed in the expert interviews without additions. Criteria number four focuses on the transfer of know-how between spin-off and parent company. This primarily involves the transfer of intellectual property rights in the form of patents and licenses. Additionally, the mobility of employees between both is mentioned. The interviews list the possibility of transfer agreements.

Cluster 4 – Success Factors: The first criteria include the success factor of clear communication and regulation of the relationship between the spin-off and the parent company, for example through fixed contracts. But there are also contrary statements when autonomy is granted. Although the allocation of autonomy and freedom is recognized as important, a certain proximity in the founding period of the spin-offs is essential. Positive aspects are the mutual provision of know-how, financial resources and networks (market and customer access) as well as cooperation in the areas of accounting, controlling, administration, product development and sales. In addition, the interviews result in the allocation of a long planning horizon by the parent company for the healthy development and growth of the spin-off.

Regarding the location and infrastructure of the spin-off, the main statements are that the digital and technical infrastructure is essential for innovative success and must be up to date. The interviews show that highly qualified employees also take this equipment for granted and use it as a criteria for or against the spin-off. The geographical, cultural and relational distance of the spin-off from the parent company is confirmed and also differentiated between the start-up phase (further away for finding culture and vision) and the growth phase (closer to the parent company from a market access perspective).

The team structure must ensure an open and agile culture. This is highlighted by the fact that a collegial, open culture with distinctive discussion and feedback possibilities as well as the heterogeneity and degree of networking of the team re seen as essential. In addition, the systematic development, qualification and support of the employees is seen as essential for the team's innovative ability. A further success factor is the elimination of routines within the team and emphasizes that micromanagement should be strictly avoided. The last success factor is the know-how of the management, supplemented by their experience with spin-offs and the living example of their visions and goals. A target-oriented methodology is titled "Servant Leadership" and a meeting at the same level as well as the management's motivation to learn.

In the strategy of spin-offs, the creation of adaptable, decentralized and flat organizations is described as a clear success factor. These organizations must be able to react agile and flexible, which are supplemented by interviews with the attributes of agility and speed, but neither growth nor professionalization should suffer from this.

The literature review shows that spin-offs succeed best in business model innovation through unique selling propositions. The experts confirm this and attribute the greatest possible success to both. In contrast, spin-offs, which are founded as "lifelines" with high pressure of expectations, are often not successful. Regarding to this, the continuous validation of the unique selling proposition and the business model are supported by the interviews and the help of development partnerships is important.

As result, the fourth cluster could be significantly supplemented by the experts. Therefore, the recommendations in the next chapter are not purely theoretical but have been additionally confirmed and expanded by experience from practice.

2 Approach to Set up a Spinoff

It is hardly possible to generalize the results of the previous chapter into a generally applicable methodology for established companies that are exposed to the risks of digitization and disruption. Nevertheless, recommendations for procedures for established companies can be derived from the synthesized results from literature and expert interviews. The recommendations are based on the result clusters from the synthesis described above and lead through them in four steps (Fig. 3).

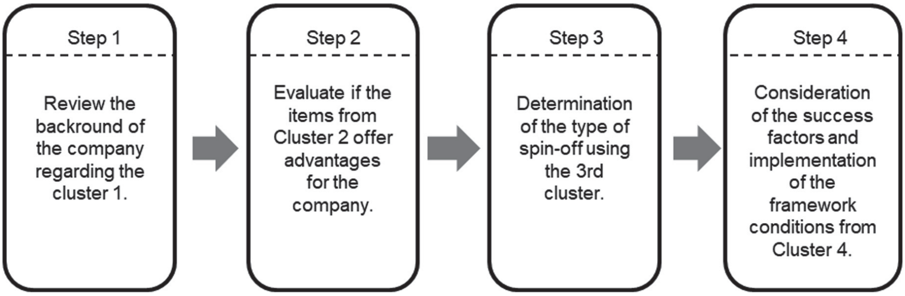


Fig. 3. Recommendation for an approach to set up a spinoff

In the first step, the results of the first cluster raise the question if a spin-off is a general option. Here it is necessary to clarify if the business model of the company is endangered, ambidextrous action cannot be taken or if the innovative action within the company is negatively influenced by bureaucratic hurdles and inertia In this case, a spin-off could be the solution.

The second step is to determine if a spin-off is suitable. For this, the arguments from the second results cluster that list the benefits of founding a spin-off should be examined. If these arguments are valid and convincing for the company and promise solutions, the recommendation can be made to consider the foundation of a spin-off.

The third step is to define the differentiation of the spin-off. The organizational form of the spin-off must be defined and the co-operation form must be planned. Therefore, the degree of freedom, sponsorship and dependencies should be defined. In addition, the necessity and the motivation for the spin-off and thus the scope of tasks must be defined in order to avoid a reactive orientation and promote proactivity. Finally, the regulation on the transfer of know-how should be planned.

In the fourth step of the recommendations, the identified success factors and framework conditions for successful spin-offs are applied. For this purpose, the four criteria on “relationship to the parent company”, “location and infrastructure”, “composition of the team” and “strategic behavior” are adapted from the fourth cluster to the spin-off.

3 Conclusion and Outlook

On the one hand, there is a shift of the most profitable companies to those that run their main business with which technology and digitization. On the other hand, aspects like these as well as new business models, industry 4.0 shorter product life cycles, servitization and neo-ecology, are major challenges.

The question how spin-offs, which were founded for this reason, are structured and how measures could be clustered, could be answered in this research, by supplementing the literature review with expert interviews from the industry. Recurring patterns were identified in the form of result clusters, which contain clearly defined prerequisites, structured procedures and success factors as well as general conditions (see Fig. 4).

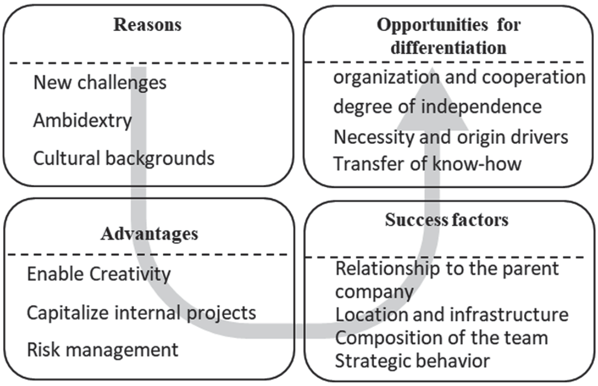


Fig. 4. The four cluster for a successful spinoff

By systematically applying the four clusters, a methodology was developed which evaluates the set-up of a spin-off as a possible solution and defines the framework conditions as well as the type of the spin-off. Following this, the identified success factors are given as guidelines for strategy and management.

The results of this research are limited in certain points. On the one hand, the existing scientific literature on set up a spin-off is incomplete and can be expanded, even to other databases and libraries. Accordingly, the literature review was able to identify the current state of the art, but also showed that the topic has not been sufficiently addressed so far. Furthermore, although the expert interviews provide a holistic picture of the research results, a larger number of expert interviews and the analysis of relevant, practice-related case studies can guarantee a more profound representativeness.

The trend suggests that companies will face increasing pressure from digitization and disruptive innovations in the future, but this can be also used as an opportunity. In order to find solutions by founding spin-offs, existing research must be expanded and long-term analyses of their success must be conducted. Another approach would be to compare it with spin-offs which have a different objective and act e.g. as accelerators or compare it with startups with the same objective but different background.

The research shows that especially the foundation of a spin-off is a promising strategy if it is long-term and the necessary rules are observed.

References

1. Doz, Y.L., Kosonen, M.: Fast Strategy. How Strategic Agility Will Help You Stay Ahead of the Game, 1st edn. Wharton School Publ, Harlow (2008)
2. Potentiale der Digitalisierung für mehr Ressourceneffizienz nutzen, Berlin (2018)
3. Krieger, W., Hofmann, S.: Digitalisierung – die neuen Anforderungen. In: Krieger, W., Hofmann, S. (eds.) Blended Learning für die Unternehmensdigitalisierung. essentials, pp. 3–7. Springer, Wiesbaden (2018). https://doi.org/10.1007/978-3-658-19204-4_2
4. Eisenbeis, U., et al.: Spin-off als Organisationskonzept. Springer, Wiesbaden (2020). https://doi.org/10.1007/978-3-658-28524-1
5. Christensen, C.M.: The Innovator's Dilemma. When New Technologies Cause Great Firms to Fail. The Management of Innovation and Change Series. Harvard Business School Press, Boston (2008)
6. Sauberschwarz, L., Weiß, L.: Das Comeback der Konzerne. Wie große Unternehmen mit effizienten Innovationen den Kampf gegen disruptive Start-ups gewinnen. Verlag Franz Vahlen, München (2018)
7. Forbes Media LLC: The World's Largest Public Companies. 2008 Ranking (2008). https://www.forbes.com/lists/2008/18/biz_2000global08_The-Global-2000_Rank.html. Accessed 1 Apr 2020
8. Forbes Media LLC: The World's Largest Public Companies. 2019 Ranking (2019). https://www.forbes.com/global2000/list/#header:marketValue_sortreverse:true. Accessed 1 Apr 2020
9. Digitale Geschäftsmodelle. Themenheft Mittelstand-Digital, Berlin (2017). https://www.bmwi.de/Redaktion/DE/Publikationen/Mittelstand/mittelstand-digital-digitale-geschaeftsmodelle.pdf?__blob=publicationFile&v=7. Accessed 1 May 2020
10. Gassmann, O., Frankenberger, K., Csik, M.: Geschäftsmodelle entwickeln. 55 innovative Konzepte mit dem St. Galler Business Model Navigator, 2nd edn. Hanser, München (2017)
11. Meinhardt, S., Popp, K.M.: Digitale Geschäftsmodelle. HMD Praxis der Wirtschaftsinformatik **55**(2), 229–230 (2018). https://doi.org/10.1365/s40702-018-0417-7
12. Kadam, S., Apte, D.: A survey on short life cycle time series forecasting. Int. J. Appl. Innov. Eng. Manag. (IJAIEM) **4**, 445–449 (2015)
13. Niemann, J.: Ökonomische Bewertung von Produktlebensläufen – Vom Life Cycle Costing zum Life Cycle Controlling. In: Spath, D., Westkämper, E., Bullinger, H.-J., Warnecke, H.-J. (eds.) Neue Entwicklungen in der Unternehmensorganisation. VDI-Buch, pp. 247–261. Springer, Heidelberg (2017). https://doi.org/10.1007/978-3-662-55426-5_27
14. Kohtamäki, M., Parida, V., Patel, P.C., Gebauer, H.: The relationship between digitalization and servitization: the role of servitization in capturing the financial potential of digitalization. Technol. Forecast. Soc. Change (2020). https://doi.org/10.1016/j.techfore.2019.119804
15. Papasabbas, L., Pfuderer, N., Muntschick, V.: Neo-Ökologie. Der wichtigste Megatrend unserer Zeit, 1st edn. Trendstudie Zukunftsinstitut (2019)
16. The SPI Manifesto, Alcala, EuroSPI 2009 (2010). http://2019.eurospi.net/images/eurospi/spi_manifesto.pdf. Accessed 25 May 2020
17. Henriette, E., Feki, M., Boughzala, I.: The shape of digital transformation: a systematic literature review. In: MCIS, vol. 9 (2015)

18. Charmaz, K.: Constructing Grounded Theory. Introducing Qualitative Methods, 2nd edn. SAGE Publications Ltd., London (2014)
19. Osterwalder, A., Pigneur, Y.: Business model generation. Ein Handbuch für Visionäre, Spielveränderer und Herausforderer, 1st edn. Campus Verlag, Frankfurt, New York (2011)
20. Wirtz, B.W.: Business Model Management. Design - Instrumente - Erfolgsfaktoren von Geschäftsmodellen, 4th edn. Springer Gabler, Wiesbaden (2018)
21. Wiesinger, R.: Welcher Unternehmenskultur bedarf es, damit Innovationen gelingen können? In: Herget, J., Strobl, H. (eds.) Unternehmenskultur in der Praxis. Grundlagen – Methoden – Best Practices, pp. 93–105. Springer, Wiesbaden (2018). https://doi.org/10.1007/978-3-658-18565-7_6
22. Maldaner, L.F., Fiorin, F.S.: An analysis framework of corporate spin-off creation focused on parent company: a case study of a traditional industrial company from the state of Rio Grande Do Sul, south of Brazil. Revista Base (Administração e Contabilidade) da UNISINOS [en linea], pp. 56–67 (2018)
23. Corsino, M., Giuri, P., Torrisi, S.: Technology spin-offs: teamwork, autonomy, and the exploitation of business opportunities. J. Technol. Transfer **44**(5), 1603–1637 (2018). https://doi.org/10.1007/s10961-018-9669-1
24. Vollmar, J.: Spin-offs, Diversifikation und Shareholder Value. Eine theorie- und hypothesengeleitete empirische Analyse europäischer Unternehmensabspaltungen. Springer Gabler, Wiesbaden (2014). https://doi.org/10.1007/978-3-658-06559-1. Zugl.: Sankt Gallen, Univ., Dissertation
25. Clarysse, B., Wright, M., van de Velde, E.: Entrepreneurial origin, technological knowledge, and the growth of spin-off companies. J. Manag. Stud. (2011). https://doi.org/10.1111/j.1467-6486.2010.00991.x
26. Interview partner 1 (COO of a spinoff): Spin-Off-Strategien etablierter Unternehmen in Zeiten von Digitalisierung und Disruption, personal communication, Düsseldorf (2020)
27. Interview partner 2 (managing director of a consulting company): Spin-Off-Strategien etablierter Unternehmen in Zeiten von Digitalisierung und Disruption, personal communication, Düsseldorf (2020)

Toolbox of Methods for the Digital Business Transformation

A Blueprint for the Education and Training of Service Engineers

Jörg Niemann(✉), Claudia Fussenecker, and Martin Schlösser

FLiX Research Centre, Duesseldorf University of Applied Sciences, Münsterstrasse 156, 40476 Düsseldorf, Germany
joerg.niemann@hs-duesseldorf.de

Abstract. Modern business models are increasingly seen as a source of outstanding organizational performance and competitive advantage that either synergizes with the previous business model or completely replaces the previous strategy. For enabling this transition the paper describes an advanced model to master the digital business transformation. Based on a large literature review the paper will identify useful methods and tools which are used in modern industrial companies. By this the findings serve as blueprint for the education and training of future service engineers and deliver the basis for the elaboration and design of according skill cards.

Keywords: Service · Education · Tools · Digital business transformation

1 Introduction – Business Models and Trends

1.1 A Subsection Sample

Today's trends such as lean supply chain, smart manufacturing, cloud platforms, big data management, artificial intelligence, augmented and virtual reality, mobility, smart e2e transparency, additive manufacturing, customization, service-orientated business models and outsourcing etc. are all based on the changes of customer mentality and technological advancements. Through the faster means of communication caused by the introduction of worldwide accessible internet, trends form and can spread much faster than ever before in history. Figure 1 shows the different driving factors on business model trends.

Shorter product life cycles, continuous changes of business processes and higher customer expectations lead to a new kind of relationship between the customer and business partners within the value chain. Customers expect faster business transactions, one-stop-shop solutions and transparency in the supply chain. This is only possible, if companies can digitalise information and data about products, customers, processes and services and thereby digitally transform their business model. With this new working method, a high amount of data is collected about business procedures and

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 783–793, 2020.
https://doi.org/10.1007/978-3-030-56441-4_59

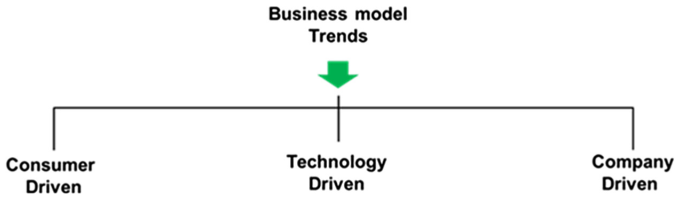


Fig. 1. Business model trends

production processes, customer demands, as well as data about internal and external communication, requiring a high amount of management and data analysis.

Digital transformation means a reorientation of products, services, processes and business models towards the continuously digitalised world and results in faster transactions and more reliability through quality and security and therefore leads to higher customer satisfaction [1]. The digital transformation of business models can be implemented in three general phases (see also Fig. 2):

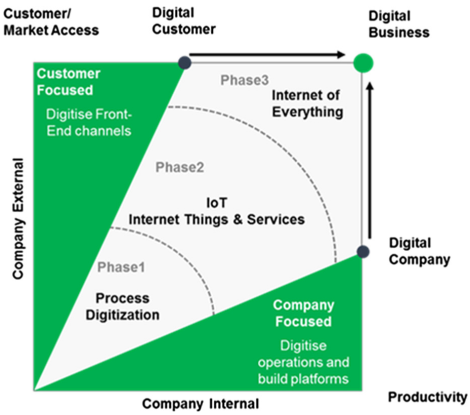


Fig. 2. Digital transformation (modified and enhanced according to [2])

- Phase 1: Digitise the current business and build a platform for digital processes.
- Phase 2: Integrate Internet of Things (IoT) functionalities into the platform and develop digital services.
- Phase 3: Close e2e loop of the entire business operations and modularize platform services [2].

Figure 2 depicts the journey from the traditional to the digital business. The model is divided into company internal and external elements, as the digitalisation of a business model can only work, if both the customer side and the own company can be "digitalised". This begins through the digitalisation of the channels and processes used to create or provide value. Afterwards, the digitalisation of products, services and other objects included in the value chain. Finally, full digitalisation of all transactions and procedures with a high automation level leads to a fully digitalised business model [3].

2 Transition from Traditional to Modern Business Models

Modern business models are increasingly seen as a source of outstanding organizational performance and competitive advantage that either synergizes with the previous business model or completely replaces the previous strategy.

New business models such as pay-per-use (usage-based payment e.g.: Car2go), peer-to-peer (trade between private individuals e.g.: Airbnb) or performance-based contracting (payment for the final performance e.g.: Rolls Royce) have revolutionised entire industries. Therefore, many companies have changed their model to move from pure product sales to the sale of problem solutions and services. When servitization moves a manufacturer all the way to becoming a solution provider there are major changes on the business model (Fig. 3).

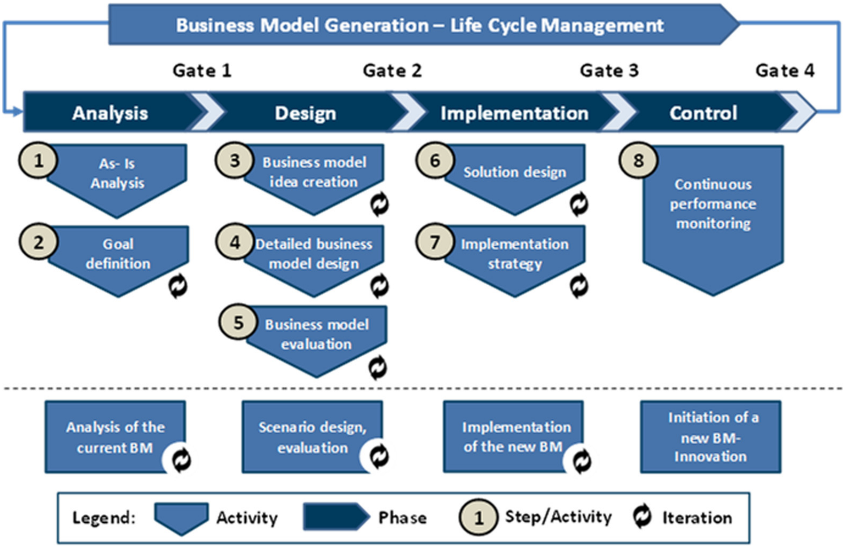


Fig. 3. Process model for business model innovation (modified and enhanced according to [6])

For enabling this transition, several frameworks are described in the literature [4, 5]. Figure 4 shows a modified and advanced model based on Bucherer which is applied for the development of a new business model on the basis of an existing model. It consists of several phases in which different activities are proposed. After each phase there is a gate which requires a verification, if the planned solutions and the meaningfulness of the concepts are given. When this is fulfilled the next phase starts, otherwise there is a need to start from scratch with the previous phase. This model is to be understood as a cycle and should serve to question and optimize the business model during the entire life cycle.

		Steps	Explanation	Proposed Methods
Analysis	1	As-Is-Analysis	The current state of the company is analysed to find out what issues exist and what the best practises are. The goal is to ensure common understanding on why a new business model is necessary and to find out what went wrong.	1. SWOT analysis 2. Benchmark 3. Ishikawa
	2	Goal Definition	Before developing ideas, the enterprise must decide which direction should be taken and what the goals are.	1. GAP analysis 2. Goal Pyramid 3. Scenario analysis

Fig. 4. Analysis – steps and tools

3 Methodology for Transformation

To develop a new business model and to transform the current business model, numerous methods exist to support the process. To provide extensive applicability to any kind of business sector, the stepwise approach presented was based on a combination of methodologies from various authors.

A methodology is defined as “a system of broad principles or rules from which specific methods or procedures may be derived to interpret or solve different problems within the scope of a particular discipline. Unlike an algorithm, a methodology is not a formula but a set of practices” [7]. The following chapters will explain the created methodology and shows the recommended methods that can be used. The set of methods in use have been selected based on a literature research and interviews with practitioners from industrial companies [8–13].

However, this process and its tools must be followed and used by a competent team with a suitable set of skills. The selection of the team members is highly important, as the business model they will find could be crucial to success. The aim of the business model transformation is to improve the existing business or change it with the result of success and higher profitability. One of the keys to success is not only the team, but also the dedication of the top management to increase innovation and change [9].

3.1 Phase of Business Analysis

Companies are a complex of various elements and interdependencies. Three main parts of the company should be analysed: The own business model (customers, value proposition, value chain and profit model), the stakeholder (customer incentives, partners, competitors) and the external influences on the business (ecosystem) [14].

Questions to be answered during the as-is-analysis of the business are:

- What can my company provide to be attractive and accepted in the future?
- What must change to ensure sustainable survival and competitiveness?
- What mechanisms must be implemented to recognise opportunities, risks and the need for change early on? [15]

Step 1 - As-Is Analysis

The As-Is-analysis includes all functions and departments of a company, the product spectrum, technology, production depth, quantity framework, financial data, customer and supplier data, organization data, all methods and tools used, as well as the employees and their relationships toward each other and the company. During the As-Is-analysis deficits will be found, which are a result of various reasons. For example, wastage of resources or potential of the employees or not seizing opportunities by acting in a non-future orientated way [15]. A literature review shows that the methods have been listed to be the most suitable and applicable in operational business:

- SWOT [4, 5, 8]
- Benchmarking [9, 15]
- Ishikawa diagram [4, 5, 16].

Step 2 - Goal Definition

The highest level of a company goal is the vision. The vision is the long-term goal of a company that describes the general purpose of the company [8]. An example for a vision from Procter & Gamble is the following:

"We will provide branded products and services of superior quality and value that improve the lives of the world's consumers, now and for generations to come. As a result, consumers will reward us with leadership sales, profit and value creation, allowing our people, our shareholders and the communities in which we live and work to prosper" [17].

When developing goals the company should decide what they want to achieve, where they want to stand in the future and in which amount of time the goal should be achieved. According to a literature review the following tools are essential in practical use:

- Goal pyramid [4, 8]
- Gap analysis [4, 8]
- Scenario analysis [4, 5, 8].

3.2 Phase of Design

In this phase ideas for the new business model will be systematically developed. This is mainly done by work groups applying creativity techniques (Fig. 5).

The identified business models will then be designed and evaluated according to the company goals and capabilities. Figure 6 shows the single steps.

Step 3 - Business Model Idea Creation

The idea finding phase marks the beginning of the business model development process. The goals have been defined and the development direction set. Either an existing product or service is to be improved or a completely new idea is to be developed. In either case a key aspect is to take the customers into account. It is crucial to understand what the customer needs and where his inconveniences lie, as well as what the customer expects from the company. Another question to be answered is, how the company can position itself in a way to satisfy the customers need in the best manner in

		Steps	Explanation	Proposed Methods
Design	3	Business model idea creation	Ideas are collected in groups and systematically developed. This phase lays the foundation for the further steps	1. Destroy your business 2. Empathy Map 3. St. Galler business model navigator
	4	Detailed business model design	The ideas are transformed into business models and described in detail.	1. Canvas 2. SIPOC
	5	Business model evaluation	The new-found business models are evaluated to enable a comparison and also to help decide on the most suitable one.	1. PESTEL Model 2. Porters Five Forces 3. Value Benefit

Fig. 5. Design – steps and tools

		Steps	Explanation	Proposed Methods
Implementation	6	Solution design	Considering all relevant details of the new-found business model on which the planning and implementation phase is based	1. Detailed Process design 2. Resource planning 3. Investment planning 4. Business Case
	7	Implementation strategy	Definition and development of the implementation strategy. Setup of a transformation plan regarding change management aspects	1. Project Charter 2. Action plan 3. RACI plan 4. Project Plan 5. Transformation Plan

Fig. 6. Implementation – steps and tools

comparison to the competition. All methods to develop creative ideas and to emphasise creativity underlie the same principles: understanding of the challenge, loosening of transfixed stereotypes and assumptions, recombining existing approaches and solutions and refining of ideas through criticism and improvement [18].

This phase helps to create new ideas through creative thinking, without being influenced by existing business models and ideas, as well as the current business model in place [19]. The following sections will explain some methods to develop creative, innovative ideas. The most commonly used idea generation methods are mind mapping and brainstorming. However, these will not be looked at further, as they are very simple and well known. For operational business applications the literature references the following tools:

- Destroy your business [11]
- The empathy map [13]
- St. Galler business model navigator [14].

Step 4 - Detailed Business Model Design

After some ideas have been gathered, they should be thought through systematically. All aspects should be considered, to enable a holistic perspective of the business model and to ensure its functionality. For this the most proven tools used in practice are

- Business Canvas [1, 13]
- SIPOC [9, 20].

Step 5 - Business Model Evaluation
Having described the business models through the canvas and SIPOC method, the business models can now be evaluated. However, this can be a difficult task due to incomparability or through incomplete perspectives on the business models. Therefore, first of all an environmental analysis for each business model that is estimated to be promising should be made and then an objective evaluation with a systematic procedure and reasonable evaluation criteria should be followed. Literature reports to apply the following tools:

- PESTEL [19]
- Porters five forces [19]
- Value benefit analysis [21].

3.3 Phase of Implementation

In this phase the solution starts the design for the new found business model. All relevant aspects and details have to be considered and included into the new solution.

After the final approval of the solution design the business model has to be implemented into the daily operation and business routines. For this a detailed implementation strategy has to be elaborated to master this transition with regard to all kinds of management aspects.

Step 6 - Solution Design
This step focuses on the business model that is planned to be used and helps to prepare and consider all aspects that are relevant for the implementation of the new business model. According to several authors this step is being performed by the application of the following tools.

- Detailed process design [4, 5, 9]
- Resource & investment plan [4, 5, 9, 22]
- Business case [9, 22–25]
- Performance management [12, 26, 27].

Step 7 - Implementation Strategy
When all the framework has been set, the implementation strategy can be developed. The goal of the implementation strategy, is that the current processes can be transformed without major delays and downtimes. Furthermore, the resource availability is to be considered and the influence on the running operations or departments before setting up a project. It is recommended to start a pilot first, to stabilise the processes and find gaps and potentials for improvement. The old and the new process should be operated in parallel, so that the new process can gain maturity and stability [9]. The implementation of a new business model within a firm should be done through a project. Beforehand, some planning must be done to define the framework. Therefore, project management elements will be used to support the transformation [19].

Beginning with the project charter, that defines all important elements of the project a second important document will be presented, that is used throughout the entire transformation phase: the project plan. Furthermore, a RACI plan is put together, to define the roles and responsibilities during the transformation project. Finally, a budget plan for the project is set up. Applicable tools to finish this step are:

- Project charter [9].
- Action plan [9].
- RACI matrix [9, 28].
- Project plan [9, 28].
- Transformation plan [11, 12, 29–32].

3.4 Phase of Control

The final phase aims at the continuous evaluation of the newly implemented business model. For this the operational performance figures have to continuously monitored and controlled in a structured manner (Fig. 7).

		Steps	Explanation	Proposed Methods
Control	8	Continuous performance monitoring	Continuous evaluation of business model performance. In case of deviations activation of measures to keep business on track	1. Balanced scorecard 2. Break-even analysis 3. Rolling forcast

Fig. 7. Control – steps and tools

Step 8 - Continuous Performance Monitoring

The objective of this final step is to ensure a durable and sustainable development of the new business model. Therefore the current operational performance is permanently monitored and benchmarked against previously defined key performance figures. By this deviations and according counter measures can be taken at an early stage. Various authors recommend the following tools to master this step:

- Balance scorecard [12, 14, 32, 36, 37]
- Break even analysis [12, 14, 32–37]
- Rolling forecast [12, 14, 32, 36, 37].

4 Summary and Outlook

The paper describes a process and a toolbox of methods for the digital business transformation. The single phases have been subdivided into smaller single steps. To master the steps a literature review has been performed to identify useful methods for the practical execution. This toolbox of methods can be used as a blueprint for the education and training of future service engineers. Therefore it is recommended to integrate the identified methods and tools into the curriculum and elaborate according skill cards for training and education.

References

1. Kreutzer, R.T.: Konzeption und Grundlagen des Change- Managements. In: Wirtschaftswissenschaftliches Studium, vol. 1, no. 46, pp. 10–17 (2017)
2. Schallmo, D., Rusnjak, A., Anzengruber, J., Werani, T., Jünger, M. (Hg.): Digitale Transformation von Geschäftsmodellen. Grundlagen, Instrumente und Best Practices. Springer Gabler, Wiesbaden (2017). https://doi.org/10.1007/978-3-658-12388-8
3. Dehmer, J., Kutzera, A.-A., Niemann, J.: Digitalisierung von Geschäftsmodellen durch platformbasiertes Value Chain Manage-ment. ZWF - Zeitschrift für wirtschaftlichen Fabrikbetrieb **112**(4), 253–256 (2017)
4. Niemann, J.: Die Services-Manufaktur, Industrielle Services planen –entwickeln – einführen. Ein Praxishandbuch Schritt für Schritt mit Übungen und Lösungen. Shaker Verlag, Aachen (2016)
5. Niemann, J., Tichkiewitch, Serge: Westkämper Engelbert: Design of Sustainable Product Life Cycles. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-540-79083-9
6. Bucherer, Eva: Business Model Innovation-Guidelines for a Structured Approach. Shaker Velag, Aachen (2010)
7. Business Dictionary "Methodology": Methodology. WebFinance Inc. businessdictionary.-com. http://www.businessdictionary.com/definition/methodology.html. Accessed 04 Apr 2020
8. Herrmann, A., Huber, F.: Produktmanagement. Grundlagen - Methoden – Beispiele, 3., vollst. überarb. u. erw. Aufl. Springer, Wiesbaden (2013). https://doi.org/10.1007/978-3-658-00004-2. Accessed 04 Apr 2020
9. Staudter, C., Hugo, C., Bosselmann, P., Mollenhauer, J.-P., Meran, R., Roenpage, O., et al. (Hg.): Design for Six Sigma + Lean Toolset. Innovationen erfolgreich realisieren, 2 vollst. überarb. und erw. Aufl. Springer, Wiesbaden (2013)
10. Schallmo, D.R.A.: Design Thinking erfolgreich anwenden. So entwickeln Sie in 7 Phasen kundenorientierte Produkte und Dienstleistungen. Springer, Wiesbaden (2017). https://doi.org/10.1007/978-3-658-12523-3
11. Kohne, A.: Business Development. Kundenorientierte Geschäftsfeldentwicklung für erfolgreiche Unternehmen. Springer, Wiesbaden (2016). https://doi.org/10.1007/978-3-658-13683-3. Accessed 04 Apr 2020
12. Gerberich, C.W. (Hg.): Praxishandbuch Controlling. Trends, Konzepte, Instrumente, 1 Aufl. Gabler, Wiesbaden (2005)
13. Osterwalder, A., Pigneur, Y.: Business Model Generation. Ein Handbuch für Visionäre, Spielveränderer und Heraus-forderer, 1 Aufl. Campus Verl., Frankfurt am Main (2011). http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=832895. Accessed 04 Apr 2020
14. Gassmann, O., Frankenberger, K., Csik, M.: Geschäftsmodelle entwickeln. 55 innovative Konzepte mit dem St. Galler Business Model Navigator. Hanser, München. http://www.hanser-elibrary.com/action/showBook?doi=10.3139/9783446437654. Accessed 04 Apr 2020
15. Nührich, K.P., Hauser, A. (Hg.): Unternehmensdiagnose. Ein Führungsinstrument zur Sicherung der nachhaltigen Existenzfähigkeit von Unternehmen. Springer, Heidelberg (2001). https://doi.org/10.1007/978-3-642-56751-3
16. Munro, R.A.: Six Sigma for the Office. A Pocket Guide. ASQ Quality Press, Milwaukee (2003)
17. P&G: Our Purpose. http://us.pg.com/who-we-are/our-approach/purpose-values-principles. Accessed 04 Apr 2020

18. Hoffmann, C.P., Lennerts, S., Schmitz, C., Stölzle, W., Uebernickel, F. (Hg.): Business Innovation: das St. Galler Modell. Springer, Wiesbaden (2016). https://doi.org/10.1007/978-3-658-07167-7. Accessed 04 Apr 2020
19. Schallmo, D., Brecht, L.: Geschäftsmodell-Innovation. Grundlagen, bestehende Ansätze, methodisches Vorge-hen und B2B-Geschäftsmodelle. Springer, Wiesbaden (2013) Zugl.: Ulm, Univ., Diss. (2012)
20. Melzer, A.: Six Sigma – Kompakt und praxisnah. Prozessverbesserung effizient und erfolgreich implementieren. Springer, Wiesbaden (2015). https://doi.org/10.1007/978-3-658-09854-4. http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&AN=1050548. Accessed 04 Apr 2020
21. Kühnapfel, J.B.: Nutzwertanalysen in Marketing und Vertrieb. Springer, Wiesbaden (2014). https://doi.org/10.1007/978-3-658-05509-7
22. Heesen, B.: Investitionsrechnung für Praktiker. Fallorientierte Darstellung der Verfahren und Be-rechnungen, 3 Auflage. Springer, Wiesbaden (2016). https://doi.org/10.1007/978-3-658-10356-9. http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&AN=1087143. Accessed 04 Apr 2020
23. Brugger, R.: Der IT Business Case. Kosten erfassen und analysieren, Nutzen erkennen und quantifizieren, Wirtschaftlichkeit nachweisen und realisieren, 2 Aufl. Springer, Wiesbaden (2009). https://doi.org/10.1007/978-3-540-93858-3. http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10297037. Accessed 04 Apr 2020
24. Wöltje, J.: Finanzkennzahlen und Unternehmensbewertung, 1 Aufl. Haufe Verlag (Haufe TaschenGuide - Band 00381, v.381) (2012). http://www.wiso-net.de/document/HAUF,AHAU__9783648025130127. Accessed 12 May 2017
25. Olfert, K.: Investition. 13., aktualisierte Auflage. Kiehl (Kompendium der praktischen Betriebswirtschaft), Herne (2015)
26. Gabler Kompakt-Lexikon Wirtschaft: 11 aktualisierte Aufl. Springer, Wiesbaden (2013). https://doi.org/10.1007/978-3-658-00008-0
27. Mühlencoert, T.: Kontraktlogistik-Management. Grundlagen - Beispiele - Checklisten. Gabler Verlag, Wiesbaden (2012). https://doi.org/10.1007/978-3-8349-3733-9. Accessed 05 May 2017
28. Bergmann, R., Garrecht, M.: Organisation und Projektmanagement, 2 Aufl. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-642-32250-1. Accessed 04 Apr 2020
29. Mervelskemper, L., Paul, S.: Unternehmenkultur als Innovationstreiber? Ein Einblick in die Praxis. In: Zeitschrift für das gesamte Kreditwesen, no. 15, p. 746 (2016). https://www.wiso-net.de/document/ZFGK__081601014. Accessed 04 Apr 2020
30. Barsh, J., Capozzi, M.M., Davidson, J.: Leadership in innovation. McKinsey Quarterly. http://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/leadership-and-innovation. Accessed 04 Apr 2020
31. Schein, E.H.: Organizational Culture and Leadership. The Jossey-Bass Business & Management Series, 4 edn. Jossey-Bass, San Francisco (2010). http://www.esmt.eblib.com/patron/FullRecord.aspx?p=588878. Accessed 04 Apr 2020
32. Gerberich, C.W., Teuber, J., Schäfer, T.: Integrierte Lean Balanced Scorecard, 1 Aufl. Gabler Verlag, Wiesbaden (2010). http://gbv.eblib.com/patron/FullRecord.aspx?p=748707. Accessed 04 Apr 2020
33. Niemann, J.: Ökonomische Bewertung von Produktlebensläufen- Life Cycle Controlling. In: Spath, D., Westkämper, E., Bullinger, H.-J., Warnecke, H.-J. (Hrsg.) Neue Entwicklungen in der Unternehmensorganisation. VDI Buch, pp. 247–261. Springer, Heidelberg (2017). https://doi.org/10.1007/978-3-662-55426-5_27

34. Niemann, J.: Life Cycle Management- das Paradigma der ganzheitlichen Produktlebenslauf-betrachtung. In: Spath, D., Westkämper, E., Bullinger, H.-J., Warnecke, H.-J. (Hrsg.) Neue Entwicklungen in der Unternehmensorganisation. VDI Buch, pp. 179–193. Springer, Heidelberg (2017). https://doi.org/10.1007/978-3-662-55426-5_22
35. Niemann, J., Fussenecker, C., Schlösser, M., Ahrens, T.: ELIC – teacher as a medium to built a new generation of skilled engineers. In: Proceedings of the International Conference on Competitive Manufacturing (COMA 2019), Stellenbosch, South Africa, 30 January–1 February 2019, pp. 234–238 (2019)
36. Gassmann, O., Frankenberger, K., Csik, M.: The Business Model Navigator. 55 Models That Will Revolutionise Your Business. Pearson, Harlow (2014)
37. Glauner, F.: Zukunftsfähige Geschäftsmodelle und Werte. Strategieentwicklung und Unternehmensführung in disruptiven Märkten. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-49242-0. Accessed 04 Apr 2020

Integrating Stakeholder and Social Network Analysis into Innovation Project Management

Karolin Gebhardt[1,2], Andreas Riel[1(✉)], and Tom Maes[2]

[1] Université Grenoble Alpes, CNRS, GrenobleINP, G-SCOP, 38031 Grenoble, France
{karolin.gebhardt,andreas.riel}@grenoble-inp.fr
[2] Bayer AG, 51368 Leverkusen, Germany

Abstract. This paper addresses a way to analyze the dynamically evolving social context of innovation projects complementary to traditional project management methods and tools. This social context being decisive for any project's success or failure, it proposes a Social Network Analysis (SNA) based approach to identify, visualize, and systematically analyze project stakeholders, their personal roles and attitudes towards the project, as well as their mutual interrelationships. This approach has been implemented as an extension to the naviProM methodology and toolset published earlier by the authors. Based on the analysis of an innovation project case study in a complex corporate environment, the proposed approach is explained and validated.

Keywords: Innovation project management · Social Network Analysis (SNA) · Corporate entrepreneurship · Decision aid

1 Introduction

Projects are embedded in a social context of human beings that is centered in the project team. All these people have their own context in terms of attitudes, interests, culture, etc. They also have relationships among each other as well with others, direct or indirect, personal or professional ones. In general, this social context has an important influence on a project's success, since it is the foundation on which project decisions are made. Innovation projects are characterized, among other properties, by a dynamic evolution of this social context. Typically, an idea is born and nurtured within a small core team. For the realization of the idea the project team has to grow in size and connect with many other players. As a result, it finds itself embedded in a complex network of stakeholders influencing the further course of the project. Traditional project management methods and tools do not support the continuous specification, visualization and systematic analysis of such highly influential stakeholder networks. This leads to the fact that they are often neglected or even ignored and can be a reason why some projects fail. In this article, we propose a social network analysis (SNA) extension to the naviProM innovation project management methodology we introduced in earlier publications [1, 2]. The key characteristic of this methodology is that it systematically considers the uncertainty that is linked to a particular project work package or task as the driving parameter to decide how this work package or task shall

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 794–804, 2020.
https://doi.org/10.1007/978-3-030-56441-4_60

be executed and managed. The social stakeholder dimension affects uncertainty by the stakeholder's capabilities, attitudes, influences, as well as mutual relationships. In order to point this out, this paper is organized as follows. Section 2 explains SNA principles and analyses related published work. Section 3 explains the particular project roles we decided to focus on, as well as the motivation behind. Section 4 reiterates on the case study presented in [2]. Section 5 complements this case study by a social context analysis at the end of two key project phases. Section 6 discusses the results achieved and their added value. Section 7 relates this work to the SPI Manifesto. Finally, Sect. 8 concludes and gives an outlook to related future activities.

2 SNA Principles and Related Work

A project team that is embedded in an organizational, and therefore social and political, context can be modelled as a network graph that has a complexity measure. Networks can be examined in terms of their structure and composition [3, 4]. The former provides a quantitative and/or graphical representation of the interconnections between the network nodes (i.e., the "network syntax"). The network composition describes the characteristics of the nodes and their links (edges) and quantifies the diversity of those attributes (i.e., the "network semantics"). [5] provides an outstanding overview of network metrics for structure and composition. For structural metrics, size, density and centralization are key on network level. For nodes and edges, different centrality measures convey measurable information related to the connectivity of particular nodes with the rest of the network or parts of it. As for quantifying composition, metrics such as the Index of Qualitative Variation (IQV) provide measures of e.g. network heterogeneity.

Representing stakeholders and their relationships within a project or process with such network graphs and using related metrics opens up the opportunity for the applying of the huge body of network graph analysis algorithms to analyzing of social networks. This particular type of graph analysis methods and applications has given rise to the scientific field of Social Network Analysis (SNA) [4]. SNA methods are based on an extensive yet flexible set of metrics that can be applied on multiple levels of analysis, both of purely social networks, and of projects, processes, products, and services (e.g. [6]). In a particular project context, such analyses have a key role in predicting and understanding interdependencies and influences both on and from the project to the task level [7].

Besides network theory, role theory helps to better characterize nodes (stakeholders) based on their specific role, expectation, norm, or skill profile. In the context of innovation projects, research emphasizes the single role of the champion or entrepreneur promoting the project from idea to implementation [8] or the sponsor providing support based on commitment or resources. Roles are temporary as networks, which means that they may change and have different relevance for specific tasks during a project life cycle. This is even more important considering the individual relationships between these roles and project's tasks. IPs must be formally integrated into the permanent organization to which innovative solutions generally do not fit. Thus, well-established practices and processes must be adjusted or eliminated, leading to active and passive organizational resistance. Furthermore, under conditions of high uncertainty, when there is no formal organizational acceptance or affiliation of an idea, dedicated resources are limited. Teams have to create value related to their vision in order to build a network of supporters that will lead them in the targeted direction (e.g. with their network, expertise, resources) [9]. It is even more important to not isolate the project and rather be aware of and understand the established structures, processes, or habits that constitute the overall system [9]. In this case, people do not fulfil their formal functions or roles, but act on the basis of personal conviction, interest and motivation. They have a positive attitude towards the project. Besides this attitude, the influence of stakeholders is relevant. The project's integration into established organizational structures makes the success dependent on dedicated and committed resources and therefore on individual stakeholders' attitudes and influences [10]. However, people with a negative attitude can have a negative influence on the project, e.g. through resistance, fear or conflict.

3 Innovation Project Roles and Stakeholder Relationships

With respect to naviProM [2], it is assumed that the stakeholder network influences all the elements, i.e., input, outcome, as well as the context. As specifically informal roles strongly influence the development of tasks [11], we focus on informal relations without explicitly differentiating hierarchy network relations. Based on scientific findings in role and network theory [11], we identified major roles with respect to the task perspective and IPs in the corporate context. Two central roles have an essential influence on task-related uncertainty, particularly with regard to the task's objective:

1. The user or customer, who potentially uses or buys the final solution, as well as the early adopter who embraces new solutions before other market players.
2. The sponsor or principle, who commissioned the task, and who provides resources that are not formally implemented as well as project support and protection in demonstrating task feasibility [12].

These roles determine the requirements and shape the objective of a task. To specify, capture, transfer and operationalize this objective in concrete actions, further roles are critical for tasks:

3. The core team that drives the project.
4. Functional experts who have complementary and required expertise to specify fulfil a task.
5. Coaches who may support the project core team on the methodical level.

In the corporate context, relationship management between the task and the existing organization is of high relevance. In particular, throughout their entire life cycle, projects need access to resources within this organization, human and non-human ones. To overcome the related challenges, another role is critical:

6. Gatekeepers can provide access to the required resources on the organizational level.

Figure 1 shows the different roles (1–6) represented as nodes in the naviProM-Net tool extension that we propose. A red dotted framed node makes it possible to differentiate between internal and external network actors.

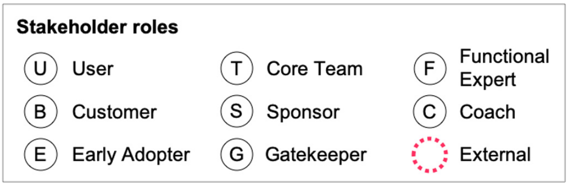


Fig. 1. Stakeholder roles [2]

As the influence and attitude of actors may vary, the model further captures positive, negative, and neutral attitude as well as low, medium, and high influence (Fig. 2). The relationship can be undirected, unidirectional, and bidirectional. The core team obviously has the highest and ideally positive impact on the task progress; however, the idea is to visualize the team as a whole in order to capture all the potential project influences coming from stakeholders.

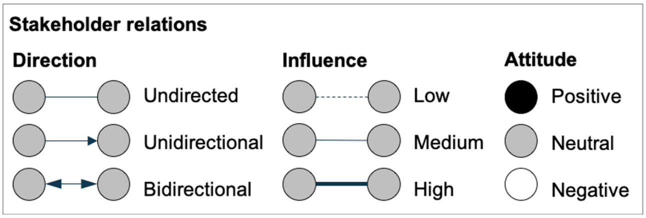


Fig. 2. Stakeholder attitude and influence [based on [2])

Figure 3 shows a sample network providing an example of how the proposed visualizing elements can be used and interpreted. In this network, the team is directly linked to one gatekeeper ("G"-node), one sponsor ("S"-node), two internal as well as one external functional ("F"-node, red dotted line) experts. These direct relations are considered in the investigation of personnel or egocentric networks. However, core teams may not have all the required network contacts in their direct network. As depicted by the three external customers ("B"-black node, red-dotted line), actors can also be indirectly linked through a chain of relations which may hamper access to required resources. The gatekeeper is of particular interest, since he has the only access to the external market, i.e., the customer. This important relationship with the customer is represented by a fat double-sided arrow. The gatekeeper's negative attitude towards the project (indicated by the white-color in the node) could indicate potential obstacles (e.g. suppression of knowledge flow) and refers to special attention. In other words, whether the massively influencing stakeholder (here the stakeholder) is using its influence to support, hamper or even stop the task is affected by its attitude towards the task. The functional experts as well as the sponsor have a neutral attitude toward the project (grey-colored nodes) and have low (dotted line) to medium influence (thin line) on the project.

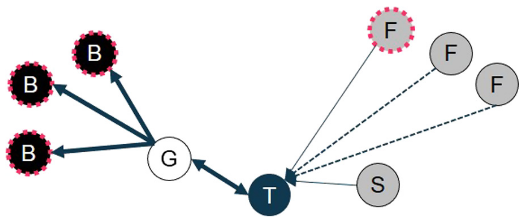


Fig. 3. Sample network

In order to apply the naviProM-Net appropriately, we propose capturing all directly and indirectly related actors to the core team in a first step. Second, the attitude towards the project must be assessed. Third, the influence and the direction of this has to be determined. The size (measured by the number of nodes; here: 8) and diversity (number of different nodes; here: 6) of networks vary and typically increase during a project [13]. Thus, all relevant actors and relations shall be captured and assessed during project setup and adjusted at the beginning of each project phase. This makes it possible to capture dynamics, to analyze the stakeholder influence in retrospective, and derive necessary strategies in order to e.g. close network gaps, win lacking sponsors or convince critics.

Based on its visualization elements, naviProM-Net makes it possible to systematically capture relevant project actors and their relationships to the project in order to detect critical or missing nodes or relationships. The key advantage of encoding these roles and their relationships in the form of a connected graph is to make available the huge set of graph analysis algorithms to understand stakeholder influences in order to apply this knowledge to task planning and analysis as well as strategy development. For example, a node's centrality is a measure of the corresponding stakeholder's capacity for reaching other stakeholders. Stakeholders with high centrality have the potential to fulfill the role of gatekeepers.

4 Case Study

In order to illustrate the practical application and added value of the naviProM-Net tool set extension, we further elaborate on the case study we presented in [2]. It is based on a corporate innovation project at a leading global Life Sciences company, where a new and innovative product-service system [14] shall be developed. This product-service system shall be facilitated by a mobile application whose design requirements were quite unknown from the beginning, since the target users' expectations had not been well analysed. A key success factor for this project is therefore to integrate pilot

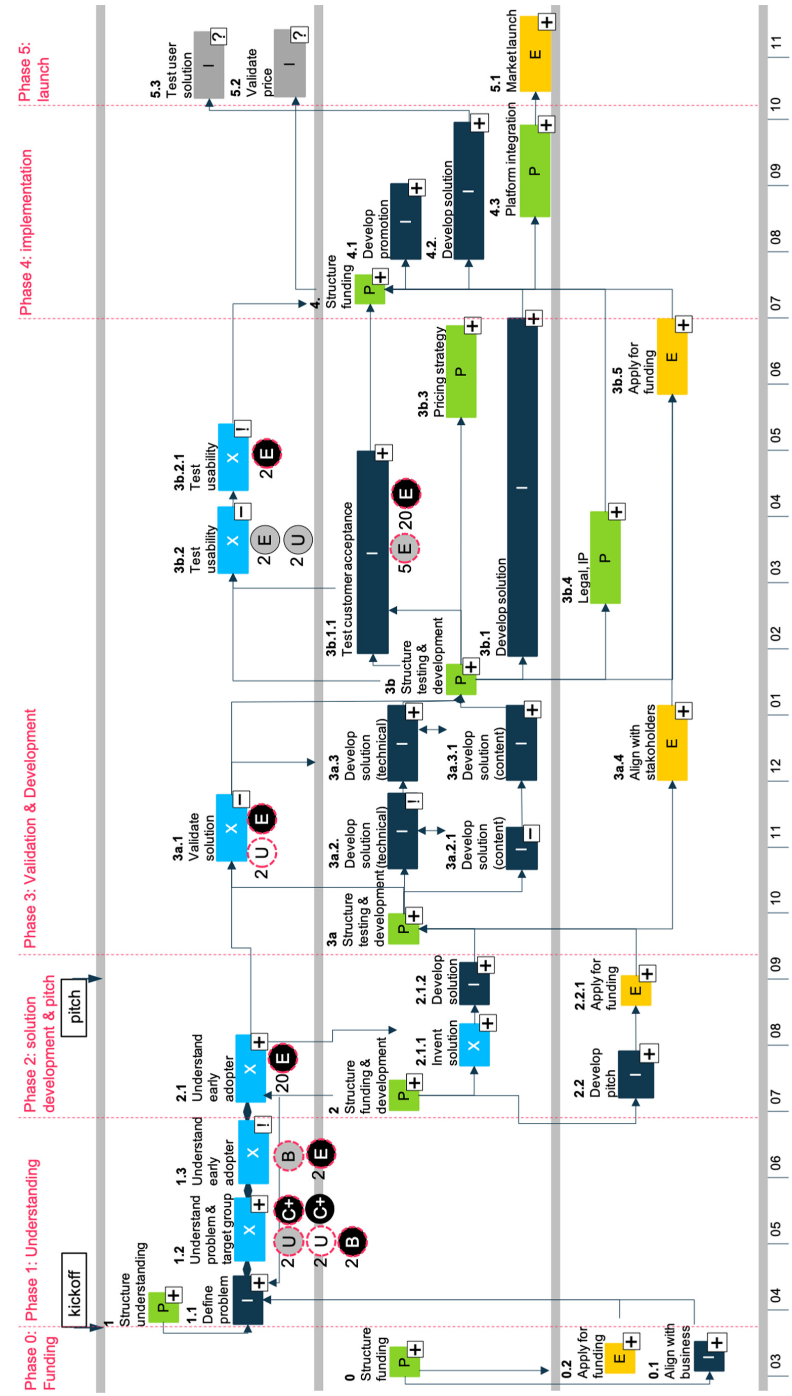


Fig. 4. naviProM-Flow representation of the case study project over five phases

customers as early adopters in the development process from the very beginning in order to assure well identifying their needs and expectations. During system and software design, it will be crucial to include regular preliminary validation steps at these early adopter's premises to include the target profile in the validation cycles and be able to adjust requirements and design decisions where necessary. The team did not execute the validation cycles well as can be seen in the naviProM-Flow (Fig. 4). The naviProM-Net (Fig. 5) helped the team to understand some of the reasons.

Figure 4 shows the course of the project across the four essential development phases, with the fifth phase (market launch) only started at the time when we finished accompanying the project. The language and tool we used to prospectively plan (during the accompaniment period) and retrospectively analyse the project based on individual tasks and their uncertainty levels is naviProM-Flow [2]. For brevity, we will not explain the entire project course here, but rather ask the interested reader to study [2]. We also provide a legend reminding the meaning of the graphical language elements we use in naviProM-Flow in Fig. 5.

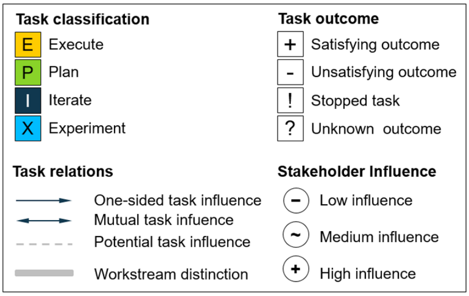


Fig. 5. naviProM-Flow project planning language elements

We clustered the project tasks into the three work streams Customer Development Work stream (CDW, top swim lane of Fig. 4), Solution Development Work stream (SDW, middle) and Business Development Work stream (BDW, bottom) in order to keep thematically related tasks together. In [2] we already pointed out that the project results available at the beginning of phase 5 suffer from unsatisfying results in phase 3. Therefore, we want to highlight how naviProM-Flow makes the major challenge of this IP visible before focusing on phases 1 and 3 in order to analyze the project's success path from the stakeholder network perspective. The long chain of experiments in the CDW, indicates that the team struggled to reduce the uncertainty related to customer needs from phase 1 to 3. Although a product-service-system project requires close interrelationship between CDW and SDW, real-life validation ensuring the service's usability in the market was not integrated. As mentioned in [2], the experiments did not generate the facts needed to validate the developed solution.

5 Integration of Stakeholder Network Analysis

The naviProM-Nets of phases 1 and 3 visualize the missing connection between the CDW and SDW in a different way. In phase 1 the naviProM-Net (left part of Fig. 6) shows a strong focus on future users and costumers. It was the neutral to negative attitude of one of the target groups noted during the activities 1.2 and 1.3 (see Fig. 4) that lead to the definition of who is an early adopter, customer, or user in this figure. The social network of this project phase also shows three smaller clusters of sponsors and gatekeepers. Their strong relationship with the core team and attitude towards the project (either positive or neutral attitude) signals the political support in the early phase of the innovation project.

In phase 3 (see right-hand side of Fig. 6), the social network grew. Additional functional experts who had the necessary skills to build the application, were added to the network. This gave rise to the formation of the new cluster content development and expanded the existing network clusters technical app development, market access, marketing strategy and business alignment. Note that some functional experts are external to the organization as some activities were outsourced. The team also expanded its external network and included additional customers, early adopters and users. The relationship between the team and the sponsors, gatekeepers did not change.

The core team made a clear internal responsibility split (represented by giving each core team member its own “T”-node.) Each team member managed a single network cluster. There was only one cross link between the clusters. One core team member together with a functional expert from the marketing access cluster were responsible for carrying out experiments with customers and users. These experiments proved to be time intensive. The limited capacity of these key players was one of the explanations why experiments were not executed as designed and had an unsatisfying outcome. Because of the missing cross links, it was up to the core team to share information.

6 Evaluation and Discussion

NaviProM helps to understand why the project was more successful in phases 1 and 2 [2] as in phases 3 and 4. The main reason was that the required experiments of phase 3, designed to validate the product-service-system solution under real life conditions were not executed as recommended by naviProM. They failed due to several reasons such as lack of capacity and improper experiment.

The social context analysis using naviProM-Net helped the team to detect and analyze decisive stakeholder networks and gaps, as well their evolution over the project course. They identified that from project phase 3, their network changed significantly, bringing along a network that clearly enlarged the core team. This is a very positive aspect in itself (given the development challenges ahead in all the three work streams). However, naviProM-Net also uncovered that this enlarged network effectively consisted of several individual network clusters that were only linked via individuals in the core team. Other links between those individual networks were missing, which alerts on a lack of communication, collaboration and integration, knowledge diffusion, as well as on a certain vulnerability and fragility of the entire network.

Even without the determination of graph-related metrics such as node centrality, connectivity, etc. the social network visualization at different stages of the project helped understand the sources of the project's key strengths and weaknesses. Most of these root causes are not visible at all in a traditional project management charts such as Gantt and PERT. Applied regularly and prospectively throughout a project, the social network analysis therefore allows identifying sources of potential problems early, and consequently adjust the project management strategy. Here, the biggest challenge is the objective collection of data required to correctly classify stakeholders and their interrelationships. While the evaluation of formal interaction is technically easy, however, subject to data protection rights, informal interaction is difficult to capture and quantify.

7 Relevance to the SPI Manifesto

People are represented by the very first letter in the SPI Manifesto's "ABC" [15]. People are at the center of every process, every organization, as well as every change process. Innovation projects are intrinsically about change, about a dynamically growing and moving network of stakeholders each having their particular direct or indirect influence on the project. From this perspective, this work gives an important contribution to current project planning and management practice in terms of fostering the pro-active integration of the human and social context perspective into project planning and strategy development. If applied from the outset, and regularly throughout the entire project duration, the influences on the project coming from people can be much more easily and systematically included in any project decision.

8 Summary, Conclusion and Outlook

We have proposed and applied an SNA-based method and tool in order to support project managers and teams to systematically include the consideration of social context in any project related decision at any time. As an extension to the naviProM methodology and tool set, naviProM-Net allows capturing key stakeholders, the roles they fulfill, their attitudes and influences in the project, as well as their interrelationships as connected network graphs. This enables the application of graph-related algorithms for the systematic evaluation of social network metrics revealing important information such as a stakeholder's centrality, network distances, etc. Applied regularly throughout a project, naviProM-Net allows the modeling and visualization of the dynamic evolution of a project's social context, which is particularly helpful and relevant in innovation projects where project managers have to develop strategies for growing the project and its spread within an organization and/or on the market.

We have also shown the integration of naviProM-Net into other naviProM project management tool elements, which is an important contribution to making project management more holistic in terms of systematically including uncertainty as well as the social dimension in every project management consideration and decision. Initial applications of naviProM-Net in four diverse innovation projects in a leading Life

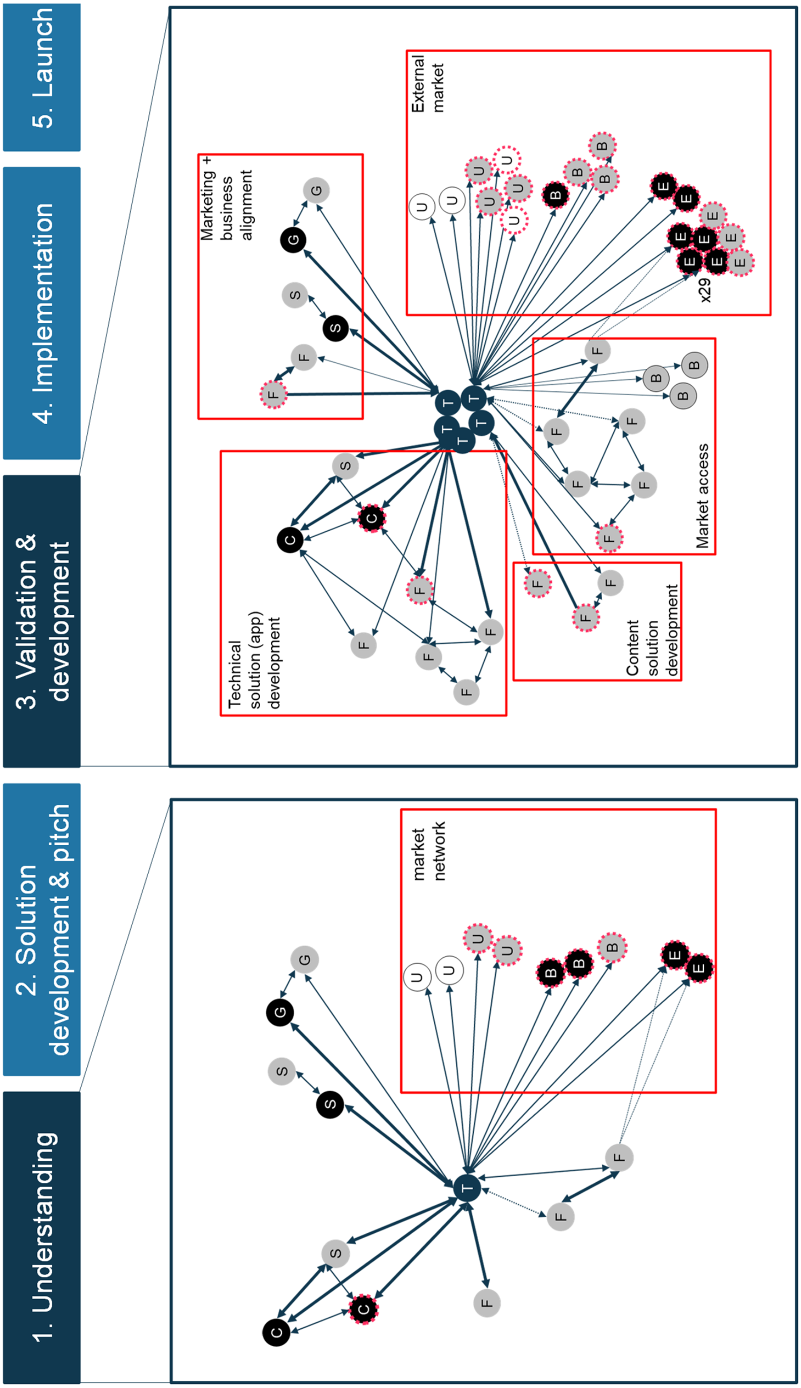


Fig. 6. naviProM-Flow representation of the case study project over five phases

Sciences company has proven very successful. Our future objectives are a broader application in various industries, as well as more detailed evaluation of the managerial impact of our method and tool set.

References

1. Gebhardt, K., Maes, T., Riel, A.: A project management decision support tool for keeping pace with the dynamics of corporate innovation projects. In: Larrucea, X., Santamaria, I., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2018. CCIS, vol. 896, pp. 619–630. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-97925-0_52
2. Gebhardt, K., Riel, A., Maes, T.: A new approach to analysing and visualizing the management of corporate innovation projects. In: Walker, A., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2019. CCIS, vol. 1060, pp. 756–768. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28005-5_59
3. Phelps, C.C.: A longitudinal study of the influence of alliance network structure and composition on firm exploratory innovation. Acad. Manag. J. **53**(4), 890–913 (2010)
4. Wasserman, S., Faust, K.: Social Network Analysis - Structural Analysis in the Social Sciences. Cambridge University Press, Cambridge (2012)
5. Parraguez, P.: A networked perspective on the engineering design process: at the intersection of process and organisation architectures. Doctoral dissertation, DTU Management Engineering (2015)
6. Collins, S.T., Yassine, A.A., Borgatti, S.P.: Evaluating product development systems using network analysis. Syst. Eng. **12**(1), 55–68 (2009)
7. Sosa, M.E.: Realizing the need for rework: from task interdependence to social networks. Prod. Oper. Manag. **23**(8), 1312–1331 (2014)
8. Howell, J.M., Higgins, C.A.: Champions of technological innovation. Adm. Sci. Q. **35**(2), 317 (1990)
9. Kelley, D.J., Peters, L., O'Connor, G.C.: Intra-organizational networking for innovation-based corporate entrepreneurship. J. Bus. Ventur. **24**(3), 221–235 (2009)
10. Hoang, H., Yi, A.: Network-based research in entrepreneurship: a decade in review. Found. Trends Entrepr. **11**(1), 1–54 (2015)
11. Markham, S.K., Ward, S.J., Aiman-Smith, L., Kingon, A.I.: The valley of death as context for role theory in product innovation. J. Prod. Innov. Manag. **27**(3), 402–417 (2010)
12. Tighe, G.: From experience. Securing sponsors and funding for new product development projects - the human side of enterprise. J. Prod. Innov. Manag. **15**(1), 75–81 (1998)
13. Dedehayir, O., Mäkinen, S.J., Roland, O.J.: Roles during innovation ecosystem genesis. A literature review. Technol. Forecast. Soc. Change **136**, 18–29 (2016)
14. Meier, H., Roy, R., Seliger, G.: Industrial product-service systems—IPS^2. Ann. – Manuf. Technol. **59**(2), 607–627 (2010)
15. Korsaa, M., Johansen, J., Schweigert, T., Vohwinkel, D., Messnarz, R., Nevalainen, R., Biro, M.: The people aspects in modern process improvement management approaches. J. Softw.: Evol. Process **25**(4), 381–391 (2013)

Virtual Reality

Analysis of Improvement Potentials in Current Virtual Reality Applications by Using Different Ways of Locomotion

Natalie Horvath[1], Sandra Pfiel[1], Florian Tiefenbacher[1], Rene Schuster[1], and Michael Reiner[1,2](✉)

[1] IMC FH Krems, Krems, Austria
michael.reiner@fh-krems.ac.at
[2] ECQA GmbH, Krems, Austria

Abstract. It has become common knowledge that the use of virtual reality (VR) applications can very often cause typical symptoms of motion sickness such as nausea and dizziness. For this reason, however, there are also more and more attempts to contain or completely avoid the side effect, which is meanwhile called cybersickness. This paper elaborates on a pre-study of a large-scale study investigating whether the use of treadmills for real, physical movement in virtual worlds can reduce the incidence of motion sickness compared to other common VR locomotion techniques and therefore enlarge the time participants can use VR for training or simulation in companies.

Keywords: Improvement · Virtual reality · VR training · Motion sickness · Virtual locomotion techniques · OmniDeck

1 Introduction

Many companies are currently programming or using VR solutions for different situations. Quite a lot of them are using VR in training or production. This means that employees need to use VR for these tasks more and more in the future. VR for trainings, for example, is currently trying to solve several problems for companies, including needed space, needed environments, risk management, preparation of facilities before building them.

These problems can be solved very well by VR. The problem in most of these applications is the movement and the haptic part of VR. The haptic part is tried to be solved by gloves by different producers, where most of them provide only small feedback and the solutions seem to have a need of improvement. More problems arise by the movement problem. If people need to move in VR, depending on the head mounted displays (HMD) the user can either currently walk by himself (only very limited space normally available and cable length is quite limiting), use teleportation (aiming in VR at a certain point and being teleported to that place) or "joystick" movement (movement like in a game with the typical ways, forward, backward, left and right). Except for one's own movement, all of the others do lead to sickness or dizziness for participants after a short period of time. This is a big problem, as this

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 807–819, 2020.
https://doi.org/10.1007/978-3-030-56441-4_61

would mean that the typical time in VR for training is limited to only a few minutes. The effectivity of a training only lasting some minutes is very small, so the cost benefit ratio is very bad. To improve this, other forms of movement need to be used. Treadmills do provide a form of "natural" movement and within the part of treadmills, the omnidirectional treadmills do provide a sense of free movement in all directions. Most people using these treadmills report that the feedback of own movement reduces the sickness or dizziness to a minimum or even zero.

This paper is about possible solutions (besides pure programming) and compares the hardware (omnidirectional treadmill) and software solutions (teleportation and joystick movement) in a prospective study. It also describes the design of a larger scale study to be carried out after the Covid-19 shutdown. Quite often the software tools for training are good but the immersion can be greatly improved when the movement is used correctly. The outcome should be usable for companies to decide on future setups for VR applications (training, simulation, etc.) so that the time spent in VR can be longer and the benefit larger for the companies in order to have a better ROI.

2 Background

Not only is the technology developing rapidly, but so are the application possibilities. VR is no longer used in games only, but increasingly in training and education, medicine, tourism and many other fields. It is therefore even more important that the technology develops at a correspondingly rapid pace. It is particularly important to create a simulation that is as close to reality as possible - especially when it comes to training for real - for example medical - missions. However, this involves far more than the quality of the visualization of what is seen. For example, a coupling with real movements - both arms and legs - is important.

However, due to the limited freedom of movement in VR, natural locomotion has so far posed a particularly great challenge. Physical walking has so far been hardly or not at all possible due to various limitations such as lack of space, cables, risk of injury, etc. [1]. This has often been replaced by simulations such as teleporting, flying or being passively moved. Since the decoupling of virtually experienced and one's own physical movement very often led to symptoms of motion sickness, it was important to find solutions for this problem.

In addition to other new techniques of simulated movement by means of controllers, there are now also possibilities for physical movement of the legs - either at a stand or with the help of kind of treadmills. In Sweden, for example, the OmniDeck was developed - a round treadmill on which you can move almost without limits in all directions due to the omnidirectional movement. The idea that the fact that one moves physically in the virtual worlds leads to the conclusion that this also leads to a reduced occurrence of motion sickness. If and how strong the difference to previous modes of locomotion is, was investigated by this field study.

The OmniDeck is a kind of round treadmill that always takes you back to the centre. This enables you to move in all directions without limits by walking independently in virtual worlds. Therefore it is, one of the few possibilities that enables a very realistic motion sequence in VR. Accordingly, the idea is obvious that its use leads to less motion sickness. However, since it has only recently been used in commercial applications, there are no studies yet to confirm this. Therefore, the goal of this study is to evaluate how different the effects of motion sickness are for different locomotion possibilities in VR - especially in comparison to the OmniDeck.

3 Fundamentals

3.1 Motion Sickness

The term "motion sickness" was first used by Irwin [2]. He was concerned with the phenomenon of seasickness but introduced the term "motion sickness" as a more comprehensive term. More comprehensive, because the symptoms associated with motion sickness occur not only when moving around on water, but also in other situations involving movement. Seasickness is a classic example of motion sickness. In addition, many people experience other forms of kinetosis, such as when flying (air sickness) or when travelling in a bus or car (travel sickness).

3.2 Simulator Sickness

In literature there are various theories regarding the origin of the simulator sickness. The most widely accepted one is the so-called sensory conflict theory [3]. The following terms are also used synonymously for this approach: "mismatch", "neural mismatch", "cue conflict" and "sensory rearrangement theory" [4]. The theory of postural instability is somewhat more recent (Postural Instability Theory; [5]). Both explanatory approaches consider the influence of the visual and vestibular system, whose interaction is of great importance for the perception of movement. It is generally accepted that symptoms of motion sickness can only be triggered in people with an intact vestibular system (e.g. [6, 7]).

3.3 Cybersickness

An important point regarding virtual reality is cybersickness, a kind of motion sickness and an effect that can occur during and after visiting virtual realities. This is one of the most important and limiting side effects, especially for many patients - not only in neurological rehabilitation - who use virtual reality systems. The effects described are dizziness, tears and other eye disorders (such as oculomotor disorders), postural instabilities, headaches, drowsiness, lack of relationships. Up to 12 h after using virtual realities, disorientation, sudden memory loss, hand-eye coordination and balance disturbance may occur [8].

3.4 Symptoms of Motion-/Simulator-/Cyber-Sickness

The simulator disease contains a constellation of symptoms like motion sickness. However, the patterns or profiles of symptoms that occur differ from those of the "real" motion sickness. Cardinal symptoms and signs of motion sickness include nausea and vomiting as symptoms, and paleness and cold sweating as signs [3]. In addition to these most common symptoms and signs, there are others that are only occasionally reported or observed. According to Reason and Brand [3], these include increased salivation, sighing, yawning, hyperventilation, gastric signs (belching and bloating), headache, head pressure, confusion, anxiety, weakness, loss of appetite, rise in body temperature, a feeling of cold in the face and extremities, a feeling of tightness in the neck or chest and continued drowsiness or drowsiness. In simulator disease and cyber sickness, vomiting and choking are rare, while other obvious signs such as paleness, sweating and salivation are common. The main symptoms reported include nausea, drowsiness, general malaise, apathy, headache, disorientation, exhaustion and fatigue [4].

4 Study Design

This study will be conducted in a randomized controlled trial design to evaluate and compare motion sickness at the OmniDeck with other VR locomotion types.

As described in the scientific literature [11], the participants will be randomly assigned to treatment groups or a control group. The process of randomly assigning a subject to a treatment or control group is called "randomization". Various tools can be used for randomisation (closed envelopes, computer-generated sequences, random numbers). Randomisation consists of two components: the generation of a random sequence and the implementation of that random sequence, ideally in a way that is not known to the participants. Randomisation excludes the potential for distortion.

For this study, the same number of pieces of paper per experimental group will be placed in a container. The test persons then have to pick a sheet of paper on which their demo type will be written. Accordingly, they will be allocated to one of the five experimental groups. This ensures that the same number of people are randomly assigned to each of the experimental groups.

There are different types of randomised study designs [12]. For this study, parallel group randomization was chosen. Each participant remains in the treatment arm to which they will be assigned for the duration of the study. Parallel group design can be used for many diseases, it allows experiments to be conducted in several groups at the same time, and the groups can be in different locations [13]. In this way it will be possible to create the same conditions as much as possible despite different modes of locomotion that require different technical equipment, and conclusions can be drawn about the respective connection with motion sickness.

4.1 Participants

An attempt will be made to cover the widest possible range of subjects. Accordingly, almost as many female as male test persons will take part in the experiments. The age of the test persons will vary from 15 to 60+ years, as we think that typical training in VR can occur already in apprenticeships until the end of the career (lifelong learning).

Depending on age and gender, they will also be distributed as homogeneously as possible among the various test groups. The participants are randomly selected. We do have lots of visitors in the VR Lab, we tend to ask all of them if they are willing to become participants. Additionally, we will, of course, use students at our university and because of them being very diverse (age from 18 to 50+, full time and part time students) this will increase the number of people interviewed. The number per group to be tested is aimed at 50+, because of the five groups this will provide us with approx. 250+ people being tested and questioned.

4.2 Environment

In order to create the same test conditions for all test groups the environment will be standardised.

Due to the fact that the OmniDeck (for real movement in virtual worlds) with its size of more than four meters in diameter and several hundred kilograms is permanently installed in one room - the eVRyLab - the other experiments will also take place in this room. In addition to this, the eVRyLab - the VR lab ate IMC FH Krems - has enough space for all kinds of VR movements, the room is darkened, and further VR equipment is stored there. Another advantage is that two television monitors and loudspeakers installed in the eVRyLab allow the researcher to see and hear live what the test person is experiencing in virtual reality.

4.3 Test Procedure

The test procedure will always be carried out in the same way regardless of the test group: First of all, the test persons will randomly pick a demo type written on paper (OmniDeck, teleporting, etc.) and be assigned to one of the five test groups. Then the questionnaire (see chapter 3.6 Motion sickness questionnaire) will be handed out to them to collect general information and to assess their current status. After a short introduction and explanation of the respective demo and type of movement, the test persons will be asked to complete the VR task. Immediately after finishing the demo (by reaching the goal or aborting the test) the test persons will be asked to answer the part about the survey of the actual state. This makes it possible to directly analyse and compare the persons' state of well-being in terms of motion sickness in conjunction with the respective demo/motion type.

One day later, the test persons will again be given the part of the questionnaire that ascertained the current status in order to be able to investigate the after-effects of the demos with regard to a possible motion sickness. Since the questionnaire will always be filled out online, there will also be no problems to send the questionnaire to the test persons by e-mail and have them fill it out, while respecting data protection (assignment of the test persons by numbers).

4.4 Tested Locomotion Techniques

The demos and tasks are the same for all types of movement: moving through the room, moving to the numbers on the side of the corridor, reading and remembering them, and finally delivering them as a kind of PIN.

OmniDeck Group

The persons in this experimental group all have the same task: they are placed in the middle of the OmniDeck, put on the VR glasses and then slowly walk step by step to complete the task in the demo and reach the goal. They are free to choose the speed at which they move, as the OmniDeck automatically adapts to this.

The functionality of the OmniDeck itself is described in detail in chapter 4.5 Technical equipment.

Flying

There is the possibility to "fly" in VR or to be passively flown/moved through virtual worlds.

In this case, the test persons are able to control their flight movement - in speed, direction and altitude - completely independently with a joystick.

Teleport VR Group

In this demo, the test persons are able to navigate standing or sitting through the corridor and to the numbers without physically moving themselves. In contrast to flying, the test persons can use a joystick to aim at a place on the ground and, by pressing a button, be immediately transferred/teleported there. Thus, there is always a change of location without the virtual world visibly passing by.

Guided Video

The test persons will be completely passively guided through the demo and have no possibility to influence speed, direction or altitude.

Control Group/Without VR Demo

The introduction of a control or comparison group is a proven means to establish high internal validity. The control group will be treated in the same way as the training group, except that they will not receive/perform the demos to be evaluated. The comparison groups as control groups will make it possible to control non-programme-bound effects to be able to actually trace changes back to the different VR locomotion types. By comparing the training group with the control group, it is possible to rule out, among other things, that a change in the testing group can be attributed solely to the time elapsed (e.g. employees have had time to practice and thus improve) or to external events (e.g. economic downturn, dismissals, summer holidays), because such causes affect the control group to the same extent [14].

In this study, the test persons will be given one and the same questionnaire twice, without any VR application in between. Only a time delay of at least 4 h will be applied to ensure that the test persons cannot remember the answers of the first questionnaire exactly but re-evaluate their current state of health.

4.5 Technical Equipment

Regarding the technical equipment, an attempt was also made to create the same test conditions for all test groups. Apart from OmniDeck, which was only used for the experimental group that should physically walk in virtual worlds, the same VR glasses were used in all experiments.

VR-Glasses

There are numerous different types of VR glasses - they differ in shape, colour, brand, manufacturer, application area, technology and accessories. What they all have in common is that they offer the possibility of diving into virtual worlds at close quarters. No generalised statement can be made about which VR glasses are most suitable, as this always depends on the respective application [15].

Since OmniDeck was used in this study for real, limitless movement in virtual worlds in a test group, it was of great importance that the VR glasses do not restrict the range of movement. Accordingly, only wireless VR glasses were considered. Since the OmniDeck's manufacturer, Omnifinity SE [10], provides software to connect to HTC VR glasses, the HTC Vive Pro Eye [16] was chosen.

These VR glasses are a wireless version with one dual OLED display per eye and a total resolution of 2880 × 1600 pixels (Fig. 1).

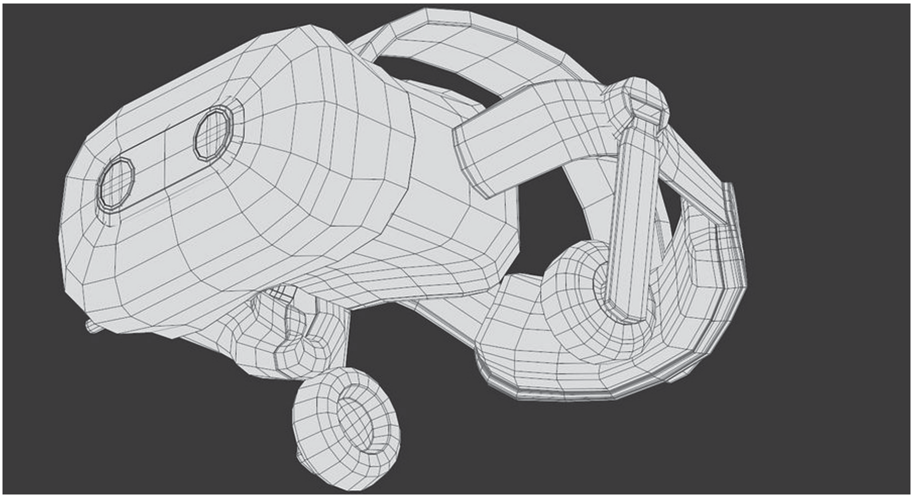


Fig. 1. Sketch of the HTC Vive Pro Eye [52]

The built-in headphones, which reproduce both volume and pitch and the reverberation of the real world, further enhance the feeling of immersion.

Since the conditions in all experimental groups should be as identical as possible, the HTC Vive Pro Eye was used equally in all experimental groups for all types of locomotion.

The Omnifinity OmniDeck

OmniDeck is - as briefly described in the basics of this paper - a motorized omnidirectional treadmill that is equipped with optical tracking technology and translation software. The deck, which has a diameter of approximately four and a half meters, consists of 16 triangular sections, each of which consists of a roller surface and is driven by its own motor and associated belt. In use, the components move at a constant speed towards the centre of the approx. 16 cm high treadmill disc. This gives the user a stable floor feeling without worrying about slipping or running off the OmniDeck. OmniDeck's software recognizes the exact position of the player by registering the six axes of movement, enabling it to move the player to the centre of the platform again and again. The optical tracking procedure thus also translates the player's movements into the virtual world [10] (Fig. 2).

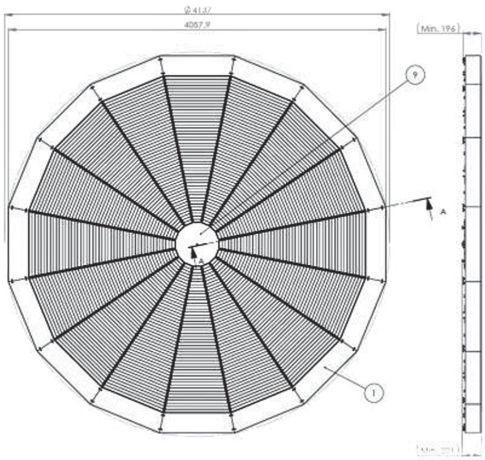


Fig. 2. Sketch of the view from above of the OmniDeck [53]

Acceleration, braking power and top speed (2 m/s or 17.2 km/h) are so high that the engines can easily manage if the user stops abruptly from the run and vice versa. Furthermore, the player can jump or crawl. The OmniDeck can be used with various VR glasses such as those from Oculus or HTC, both wired and wireless [17] (Fig. 3).

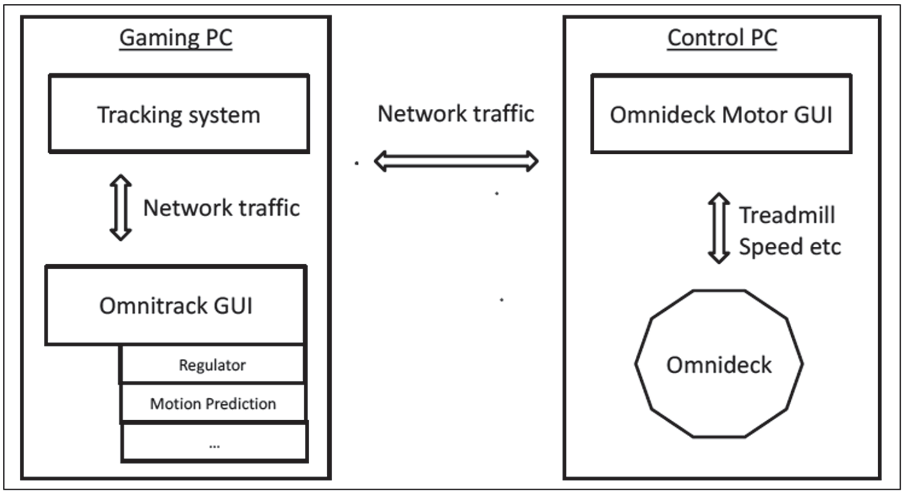


Fig. 3. Overview of the OmniDeck-System [53]

The OmniDeck, including VR equipment, PC and other accessories, is installed at the campus of IMC FH Krems, Austria, in the so-called eVRyLab, which was specially converted for this purpose. Around the OmniDeck itself, a floor was built at the same height, which makes it impossible for the OmniDeck to fall off, even if the OmniDeck edge is violated [18].

As described above, the OmniDeck was used in one of the four experimental groups. The demo and task were the same for all of them, only the way of moving around in the virtual world was different. Thus, in comparison to other types of VR locomotion, it could be analysed how differently the locomotion techniques affect the motion sickness. Furthermore, it could be investigated whether the real movement on the OmniDeck leads to a reduced occurrence of motion sickness.

4.6 Motion Sickness Questionnaire

"Cybersickness" has already been investigated in many different studies [9].

For the project "Simulation of missions on behalf of the Presidium of the Bavarian Riot Police" [19] a detailed literature analysis on the survey of motion sickness has already been carried out. The most frequent questionnaires to date were also examined:

- Motion History Questionnaire (MHQ – Kennedy, Lane, Grizzard et al., 2001) [20]
- Motion Sickness Questionnaire (MSQ – Kennedy and Graybiel, 1965) [21]
- Simulator Sickness Questionnaire (SSQ – Kennedy, Lane, Berbaum and Lilienthal, 1993) [22]

Based on these investigations, a new, timely questionnaire was then created. Because of this, the survey for this study was created by means of Unipark - an online tool for creating surveys and the academic program of Questback [23]. The questionnaire can be found in the appendix.

4.7 Screening Experiment

The screening experiment with 16 participants showed some minor hints. Of the 16 Participants 9 used the Omnideck, one teleportation, one flying and two the video mode.

Four out of the 9 participants of the Omnideck felt at least in one symptom worse than before, all of the participants with flying and all of the participants with video felt worse than before. No real change occurred to the one person using the teleportation mode.

The following number of changes (16 questions to this topic) occurred:

- Omnideck between 0 and three changes per participant
- Flying between 4 and 11 changes per participant
- Video between 2 and 5 changes per participant

This could hint to the assumption that flying and video is making people feel worse than using the Omnideck, still we need to have a larger number of people and go for a more detailed analysis to make first judgements.

For this a big study will be realised within the next 6–8 month where the we will aim to have at least 50+ participants per group (if possible even more).

5 Conclusion

The conclusion will be finally delivered after the study has properly finished and all questionnaires have been analysed. Yet. the pre-test already shows some first directions.

If the sickness and dizziness of participants can be reduced while training in VR, by using hardware ways of movement in VR, future training applications can be massively improved in terms of time within VR and by this in the amount of training that can be given to participants. Papers do clearly state that the “gut feeling” while training is a very important part for success, as people not feeling well are not good learners. This means that companies planning or working on ideas for training or simulations in VR currently should already think of the movement part.

This means that we think a lot of principles of the SPI manifesto can be applied to this improvement. The improvement here will hopefully affect the daily activities of trainees using VR in the future. The improved training and longer time in VR should improve the performance of the trainees and, thus, should also improve the success of the trainings process. Also, the discussions in the future about movement in VR can lead to a model where the different movement possibilities could be applied to different roles and, thus, making programming and money being spent effectively for more companies. The culture of training already sees a change today, as the surrounding of trainings is improved by companies more and more to have a better or even faster result while training people. This will be even more important while training the employees in VR. As this is a learning cycle for companies, it will lead to improved trainings with more topics to focus on, as VR enables new ways of training. And, of course, this will lead to new solutions of locomotion in VR.

Appendix

Questionnaire:
BEFORE the VR demo was held, the following questions were asked:

Age: x
Gender: x
Country of origin: x
Previous experience with VR: x
Good spatial imagination? yes/no

State of health: (based on 2.3.4.1 Individual factors) [55, p. 20]
When did you get up this morning? Approx. around x o’clock

How many hours did you sleep last night?
When did you eat your last meal? Approx. around x o'clock
Do you take medication regularly? yes/no
Have you recently been vaccinated against flu? yes/no

Suffer from

- Exhaustion: yes/no
- Cold: yes/no
- Fatigue: yes/no
- Stomach discomfort: yes/no
- Emotional stress: yes/no
- Ear infections/ear blockages
- Aftermath of excessive alcohol consumption: yes/no

Motion disposition/inclination to motion sickness:
How well do you contract these situations? (very bad/bad/medium/good/very good)

- Ride a roller coaster
- Driving along in the back seat of a car
- Sitting backwards to the direction of travel on the train
- Carousel driving
- Travelling in a car in childhood
- Long rides in the car or bus
- Reading while driving in a car
- Boat trips

Initial situation - Questionnaire for the detection of SSQ symptoms:
Please check how much the following statements apply to you (not at all/something/medium/strong)

- General discomfort
- Fatigue
- Strained eyes
- Increased saliva flow
- Difficulty seeing sharply
- Nausea
- Difficulty concentrating
- Head pressure
- Blurred vision
- Dizziness (eyes open)
- Dizziness (eyes closed)
- Belching
- Sweating
- Stomach makes itself noticeable
- Headache
- Balance disorders

Directly AFTER the execution of the respective demo the following questions were asked:
Demo questions:
What demo did you see?/What form of VR movement have you experienced? No Demo/Passive (Video)/Teleport/"Flying"/Active Walking on OmniDeck

Could you see/execute the demo to the end? yes/no

If not, when was the demo cancelled? After about x minutes

Simulation-related questionnaire for the detection of SSQ symptoms:
Please check how much the following statements apply to you. (not at all/something/medium/strong)

- General discomfort
- Fatigue
- Strained eyes
- Increased saliva flow
- Difficulty seeing sharply
- Nausea
- Difficulty concentrating
- Head pressure
- Blurred vision
- Dizziness (eyes open)
- Dizziness (eyes closed)
- Belching
- Sweating
- Stomach makes itself noticeable
- Headache
- Balance disorders

Other comments/special features:

References

1. Nabiyouni, M., Saktheeswaran, A., Bowman, D.A., Karanth, A.: Comparing the performance of natural, semi-natural, and non-natural locomotion techniques in virtual reality. In: 2015 IEEE Symposium on 3D User Interfaces (3DUI), Arles, France, pp. 3–10 (2015). https://doi.org/10.1109/3dui.2015.7131717
2. Irwin, J.A.: The pathology of sea-sickness. Lancet **118**(3039), 907–909 (1881). https://doi.org/10.1016/S0140-6736(02)38129-7
3. Reason, J.T., Brand, J.J.: Motion Sickness. Academic Press, Cambridge (1975)
4. Kennedy, R.S., Hettinger, L.J., Lilienthal, M.G.: Simulator sickness. In: Motion and Space Sickness, chap. 15, pp. 317–341. CRC Press, Florida (1990)
5. Riccio, G.E., Stoffregen, T.A.: An ecological theory of motion sickness and postural instability. Ecol. Psychol. **3**(3), 195–240 (1991). https://doi.org/10.1207/s15326969eco0303_2
6. Crampton, G.H.: Motion and Space Sickness. CRC Press, Boca Raton (1990)

7. Cheung, B.S., Howard, I.P., Money, K.E.: Visually-induced sickness in normal and bilaterally labyrinthine-defective subjects. Aviat. Space Environ. Med. **62**(6), 527–531 (1991)
8. Weiss, P., Kizony, R., Feintuch, U.: Virtual reality applications in neurorehabilitation. In: Textbook of Neural Repair and Rehabilitation. Cambridge University Press, New York (2014)
9. Davis, S., Nesbitt, K., Nalivaiko, E.: A Systematic review of cybersickness. In: Proceedings of the 2014 Conference on Interactive Entertainment - IE2014, Newcastle, NSW, Australia, pp. 1–9 (2014). https://doi.org/10.1145/2677758.2677780
10. Omnifinity. http://omnifinity.se/. Accessed 16 Mar 2020
11. Friedman, L.M., Furberg, C.D., DeMets, D.L., Reboussin, D.M., Granger, C.B.: Fundamentals of Clinical Trials. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-18539-2_23
12. Wang, D., Bakhai, A. (eds.): Clinical Trials: A Practical Guide to Design, Analysis, and Reporting. Remedica, London (2006)
13. Hulley, S.B. (ed.): Designing clinical research, 4th edn. Wolters Kluwer/Lippincott Williams & Wilkins, Philadelphia (2013)
14. Hager, W. (ed.): Evaluation psychologischer Interventionsmaßnahmen: Standards und Kriterien: ein Handbuch, 1st edn. Huber, Bern (2000)
15. Zobel, B., Werning, S., Berkemeier, L., Thomas, O.: Augmented- und Virtual-Reality-Technologien zur Digitalisierung der Aus- und Weiterbildung – Überblick, Klassifikation und Vergleich. In: Thomas, O., Metzger, D., Niegemann, H. (eds.) Digitalisierung in der Aus- und Weiterbildung, pp. 20–34. Springer, Heidelberg (2018). https://doi.org/10.1007/978-3-662-56551-3_2
16. VIVE Pro Eye | VIVE. https://www.vive.com/de/product/vive-pro-eye/. Accessed 19 Mar 2020
17. MSE Omnifinity AB: Manual - Omnideck User Guide C-00968-B. Omnifinity, 05 March 2019
18. IMC FH Krems Web: eVRyLab - Project Scan2VR. https://www.fh-krems.ac.at/en/research/department-of-business/. Accessed 17 Mar 2020
19. Neukum, A., Grattenthaler, H.: Abschlussbericht – Teil II: Kinetose in der Fahrsimulation (Projekt: Simulation von Einsatzfahrten im Auftrag des Präsidiums der Bayerischen Bereitschaftspolizei). Inter-disziplinäres Zentrum für Verkehrswissenschaften an der Universität Würzburg, July 2006
20. Kennedy, R.S., Fowlkes, J.E., Berbaum, K.S., Lilienthal, M.G.: Use of a motion sickness history questionnaire for prediction of simulator sickness. Aviat. Space Environ. Med. **63**(7), 588–593 (1992)
21. Kennedy, R.S., Graybiel, A.: The dial test - a standardized procedure for the experimental production of canal sickness symptomatology in a rotating environment. U. S. Naval Schoole of Aviation Medicine, Pensacola (1965)
22. Kennedy, R.S., Lane, N.E., Berbaum, K.S., Lilienthal, M.G.: Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. Int. J. Aviat. Psychol. **3**(3), 203–220 (1993). https://doi.org/10.1207/s15327108ijap0303_3
23. Unipark: Unipark. https://www.unipark.com/en/. Accessed 05 Mar 2020

A Concept for Virtual Reality Based Industrial Maintenance Training Preparation

Manoch Numfu, Andreas Riel(✉), and Frédéric Noël

Université Grenoble Alpes, CNRS, Grenoble INP, G-SCOP,
38031 Grenoble, France
{manoch.numfu, andreas.riel}@grenoble-inp.fr

Abstract. This paper presents ongoing research about the use of Virtual Reality (VR) to facilitate training for industrial maintenance. The objective is the help industry training personnel for maintenance operations at the distance without being obliged to make trainers and training objects travel. Despite the availability of increasingly many and mature VR devices, it is still difficult to achieve the level of realism needed for the effective training of particular manipulation gestures that are vital for specific assembly and disassembly procedures. We propose a systematic concept for a process in a VR-based experimentation environment that facilitates the selection, calibration, and evaluation of different VR devices for the training of specific maintenance operations. We present a concrete implementation of this process for demonstrating the feasibility of our concept. Via virtual hands, users can interact directly with virtual objects via various devices (mouse, keyboard and Leap Motion). Maintenance gestures can be executed as naturally as possible and recorded for facilitating the preparation of work instructions and training sessions.

Keywords: Virtual reality · Virtual training · Maintenance training

1 Introduction

Despite the trend towards remote maintenance tasks, largely facilitated by global high-bandwidth internet, a lot of mechanical installation still require frequent regular as well as occasional and urgent on-site maintenance and repair interventions. In particular, heavy and locally fixed installations such as ropeways and power stations are affected by this need. Driven by the globalization megatrend, these installations have seen a rapid worldwide expansion over the last decades. Along with this comes the need for rendering the maintenance tasks more sustainable in terms of safety and security for the maintenance personnel, qualification and employment of local people, knowledge and skills transfer to developing countries, huge savings thanks less travelling and preventive maintenance, the reduction of ecological footprint due to long and frequent travelling activities of skilled personnel, as well as better and more frequently maintained installations. In addition to that, the increasing density and complexity of mechanical parts and assemblies makes maintenance tasks increasingly difficult and hazardous [1].

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 820–829, 2020.
https://doi.org/10.1007/978-3-030-56441-4_62

In order to address these challenges, we present a methodology for transferring the training for such maintenance operations to the virtual space thanks to Virtual Reality (VR) technology. Thereby, our focus is on creating a concept that facilitates the easy experimentation of many different VR devices for preparing work instructions and the training process itself, in order to adequately address the challenge of providing the appropriate level of realism to the virtual training operations. This article was mostly concentrated on VR technology to prepare the training preparation process without considering the economic analysis. We also aimed to create an environment allowing the preparation of industrial maintenance task trainings in VR environments. The major driver for this challenge in the mechanical operations we need to simulate are specific hand gestures, which, if not performed correctly, will not lead to success. A typical category of such operations requiring very specific hand gestures is assembly and disassembly by manipulating screws, bands, clips, etc. Based on a library of classified hand gestures that we have built within our virtual training platform, we show how training sceneries and sequences can be modeled such that different training setups can be specified and permanently stored in the platform. Using the case study of a folding bicycle we demonstrate the key challenges of our research as well as how we currently address them.

The remainder of this paper is structured as follows: Sect. 2 elaborates on related works and relates them to ours. In Sect. 3 we present the virtual training preparation process concept we published previously in order to provide the necessary foundation for explaining our ongoing work. Section 4 introduces the data meta-model used in order to specify and store different training sequences in terms of elementary hand gestures. Section 5 presents the case study mentioned above in order to show the feasibility of our concept and process. Finally, Sect. 6 concludes and gives an outlook on planned future research activities.

2 Related Work

There is a lot of research trying to apply the virtual reality (VR) technology to increase efficiency, reduce costs as well as the risk in the training process in maintenance task [2–9]. Most of this research focusses on the overall picture of the training process, rather than digging into the details of the particular gestures applied during disassembly, assembly and any kind of special maintenance task. Appropriate gestures are the main factor for achieving the maintenance work successfully and safely. In fact, there are many research studies about the use of virtual reality technology for creating a virtual hand and applying it to various applications. Gan Lu et al. [10], proposed development and evaluation of an immersive human-computer interaction system based on stereoscopic viewing and natural hand gestures. The system is enable a user to use a number of simple hand gesture to perform basic object manipulation task involving selection, releasing, translation, rotation and sealing. Prachyabrued et al. [11], studied a virtual grasp and release method. They used heuristic analysis of finger motion and a transient incremental motion metaphor to manage a virtual hand during grasp and release. Ullmann et al. [12], described the basic algorithm for grasping with one or two hands and presented a method to manipulate the grasped objects between the virtual fingers. Xiao et al. [13], proposed a hand gesture-based design interface for the review of CAD models in virtual environments. Their applications provide a natural and intuitive user experience in the CAD model manipulation. Shamsuzzoha et al. [14], proposed the generic concept application of VR technique in industrial maintenance. They applied the application of VR techniques for operating and maintaining power plants. Their technique was able to explain relevant maintenance work, which can be helpful for the engineer\technician to acquire sufficient knowledge and expertise before conducting the real life maintenance works. This technique also provided the support needed to develop cost-effective production systems with complex maintenance management. VR technology also can provide more outcomes than many traditional training methods. Many research works found that the VR technology can reduce learning time, decrease the number of mistakes made by trainees, increase the learning efficiency, as well as help trainees retain knowledge longer than traditional methods [15]. Although these researches described the characteristics of the gestures and hand usage, they do not mention the application of these gestures for designing and creating a work instruction. Most of these researches focused on using virtual hand technology only in during the training session. They do not address neither the training preparation phase nor the designing process of the work instruction. However, these phases are crucial for assuring the appropriateness of the virtual training, both in terms of VR technology used, and gestures actually trained. Legardeur et al. [16], propose a new kind of tangible interface based on the handling of physical objects that they call “interactors”. They affirm that the handling of interactors can produce two kind of result. The first is to give matter for reflection on the parts assembly operations. The second can lead to the proposition of different assembly solutions for the mechanical system studied.

3 Virtual Reality Training Preparation Process Concept

In [17] we introduced our Virtual Reality Training Preparation Process (VR-TPP) leveraging a systematic approach to creating a VR-based experimentation environment that facilitates the selection, calibration, and evaluation of different VR devices for the training of a specific maintenance operation. As depicted in Fig. 1, the VR-TPP consists of the four tasks Preparation, Execution, Implementation and Training, clustered in two phases, the Preparation phase and the Training phase. The preparation phase consists of the two tasks preparation and execution. The preparation task consists of the two modules Model Preparation and Work Instruction (WI) Making, while the Execution task consists of the four modules Recording, Translation, VR Using and Assessment. Within this holistic concept, the special focus of the research presented in this article is on the data model and the WI making process step. In [18], we elaborated on building a library of elementary hand gestures needed to manipulate mechanical joints. The gestures specified in this library in terms of the most appropriate pattern for executing them provide the basis for creating WI and training procedures for higher level operations.

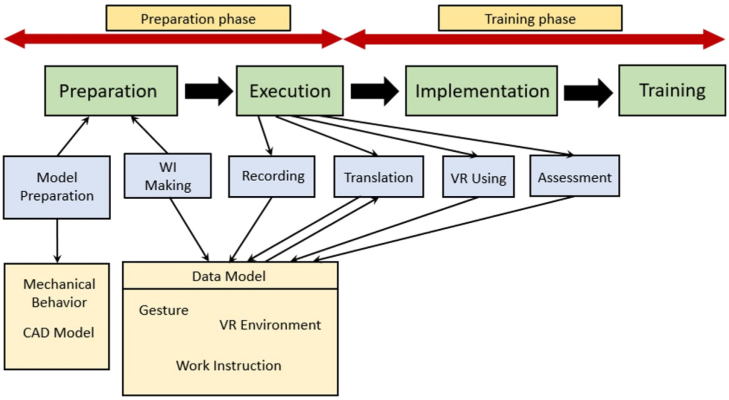


Fig. 1. VR-TPP Virtual Reality Training Preparation Process concept [17]

4 Abstract Training Sequence Specification Model

The UML class diagram of the generic interaction model consists of two main models, the BOM (Bill of Materials) and WI (Work Instruction) models as show in Fig. 2. Both models have a relationship through the classes representing a mechanical part and a gesture. The BOM consists of assemblies and parts. On this level, the assembly is a group of components or parts. Each assembly can be linked to other assemblies. The part class represents a mechanical part that constitutes a product component or any subset of parts of a product. Gestures can interact with any parts in order to constitute the WI model that describes the work steps in terms of gestures, tools and connected joints. Gestures, tools, and joints can be instantiated from libraries containing various information items required for simulating these objects appropriately in the VR environment, as well as creating work instructions.

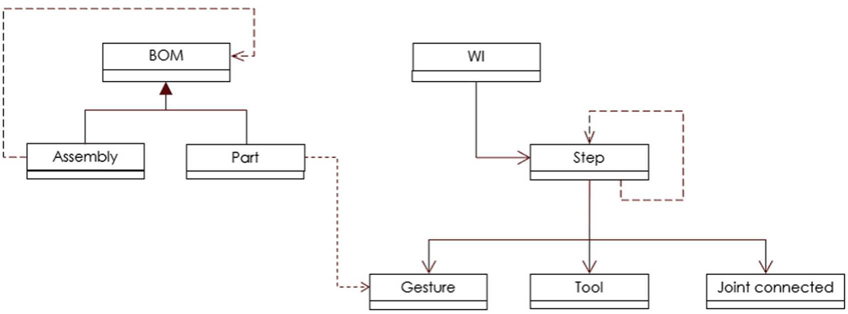


Fig. 2. UML model for specifying interactions during maintenance operations

Based on this abstract generic interaction model, we created models for the specification of BOM and WI, as shown in Fig. 3. The BOM contains several parts that can be manipulated using different gestures. Parts are composed of one or several 3D objects. A WI is composed of a sequence of steps each characterized by the specific gestures required to perform the maintenance task correctly. Instantiations of this generic WI model will be used to specify training sequences, which can subsequently be replayed, provided that the corresponding image sequences have been recorded before (e.g. by filming an expert performing the sequence), and the logical model entities assigned.

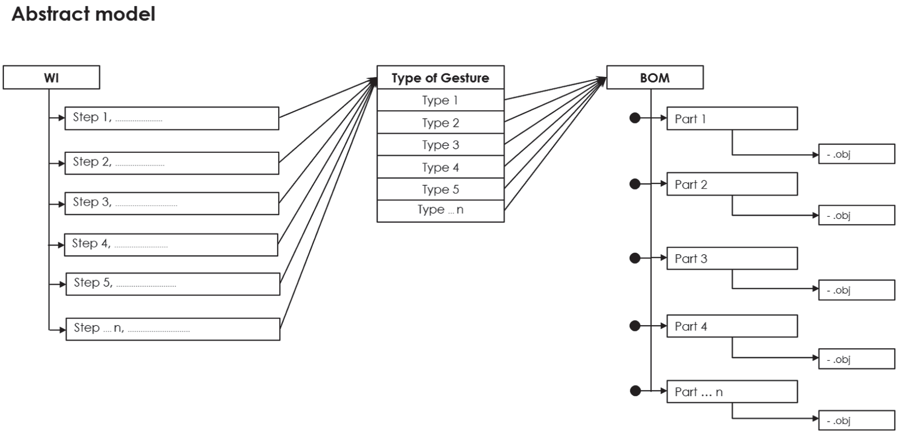


Fig. 3. Abstract WI sequence model

Figure 4 shows two examples of hand gesture specifications, in this case for removing of specific types of nuts from bolts. Only data essential to the inclusion of instructions in the work instructions are shown, as well as data required to be able to determine if the trainee applied the gesture correctly. At this stage, this judgement requires human intelligence.

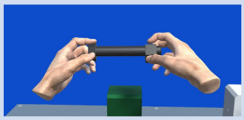
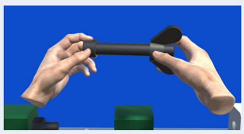
Type of maintenance task	Pattern of hand gesture	Definition	Hand gesture
Removing Hex Finish Nuts out from Hex Head Bolts	Pinch with TIM T = Thumb I = Index M = Middle	This hand gesture use three fingers to catch the bolt and nut. Thumb, Index and Middle will be used to catch the head of bolt and nut. Using three fingers is easy used to catch and rotate.	
Removing Wing Nuts out from Socket Cap Screws	Pinch with TIMP T = Thumb I = Index M = Middle P = Ring	This hand gesture is a special type, it use four fingers to catch the wing nut (Thumb, Index, middle and Ring). Thumb and Middle use to pinch the middle of head of wing nut. Index and Ring use to control upper wing and lower wing because it will use to rotate clockwise or counterclockwise.	

Fig. 4. Example gesture specification in the gesture library [18]

In the future, however, one can imagine that human intelligence can be replaced by machine learning techniques relying on pattern matching between the model gestures from the library and those recorded from the training session [18].

5 Case Study

In order to demonstrate our concept's feasibility, as well as our abstract model's validity, we investigated several scenes, objects and training tasks. In this section, we will give a brief insight into the specification of a work instructions for the most important gestures required to fold a foldable STRIDA bicycle correctly. The particular bicycle we had at our disposal for this study is depicted in Fig. 5, both in folded (left) and unfolded form (right). In fact, this bicycle can be folded conveniently only by applying the right gestures of holding the bicycle and manipulating the right joints in the correct sequence. This requires studying the work instruction and practice on the real object. Although the printed, "static" work instruction delivered with this bicycle is helpful, there is clearly a gap to overcome between the theory shown there, and its practical application to the real object. This is mainly due to the fact that the manipulation of the individual joints is not evident and requires investigation and practice. This is also why we have chosen this bicycle as being representative for the difficulties typical mechanical maintenance tasks can bring along.

Fig. 5. Foldable STRIDA bicycle folded (left), and unfolded (right)

Figure 6 shows the instantiation of the generic WI/BOM model for the specification of the individual steps required for folding the bicycle. Each step is associated with the gestures that are appropriate for manipulating the concerned parts of the BOM. This specification's elements were related to the image sequences recorded in the VR environment depicted in Fig. 7.

During training on the VR-TPP-MT (MT stands for Maintenance Tasks) platform, users must wear the Head Mounted Display (HMD) and stand in front of the laptop at a distance of 0.3 to 0.5 m. The virtual environment is displayed in real time on the HMD screen. The user raising his both hands in front of the HMD screen leads to the

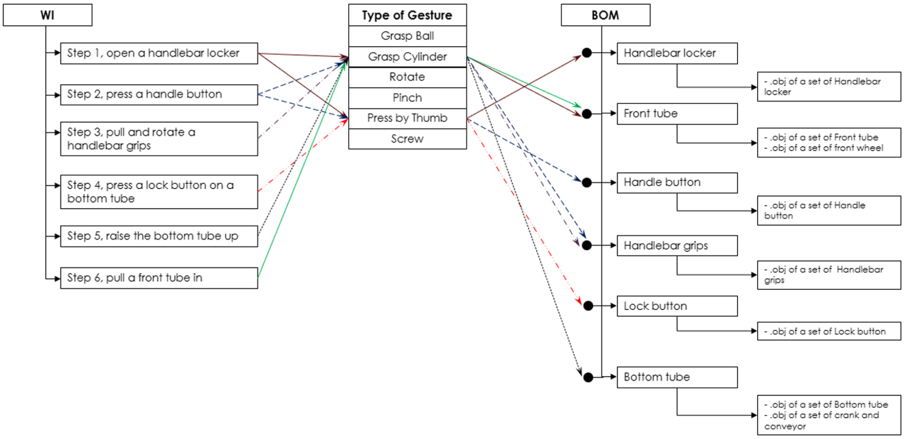


Fig. 6. Work instruction specification for folding the STRIDA bicycle

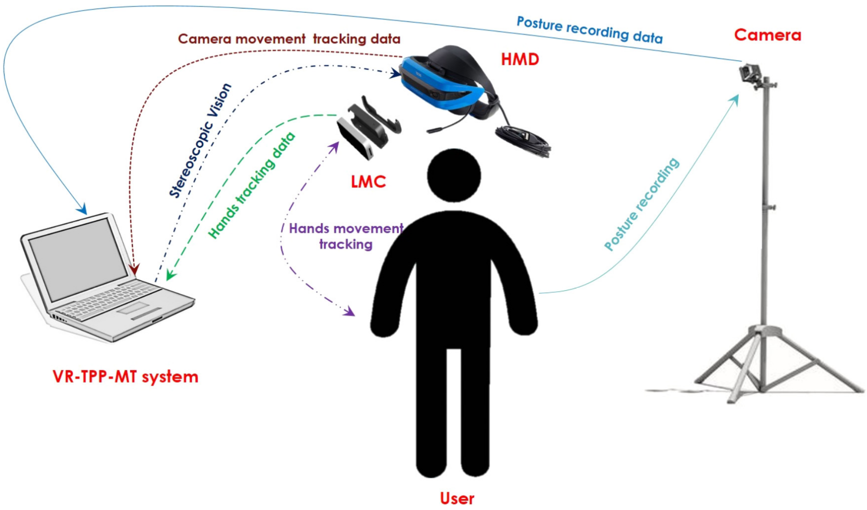


Fig. 7. VR environment for recording maintenance gestures

calibration of the virtual hands. The HMD screen integrates the Laser Distance Measuring device (LMC), allowing the system to detect and display the virtual hands in the virtual environment. After calibrating, the user can walk around in the real space to discover the virtual environment in an area of approximately 1.5 m^2, a limitation given by the cable lengths of HMD and LMC. While the user rotates and moves the head, the camera will display the image on the HMD screen in the first person view, which is like the own human eyes' view.

In this environment, the parts of the virtual bicycle are identified by the system as the user touches them with his virtual hands (using his real hands). The particular characteristics of the gestures are added from the library of gestures by relating the manipulated joints with the appropriate gestures. In this way, we avoid the need for a highly realistic simulation of gestures in the VR environment.

The bicycle folding process sequence consists of the following six steps:

1. open the handlebar locker,
2. press the handle button,
3. pull and rotate the handlebar grips,
4. press a lock the button on the bottom tube
5. raise the bottom tube up,
6. pull the front tube in.

The recorded sequence of these steps executed in the VR-TPP-MT platform are shown in Fig. 8.

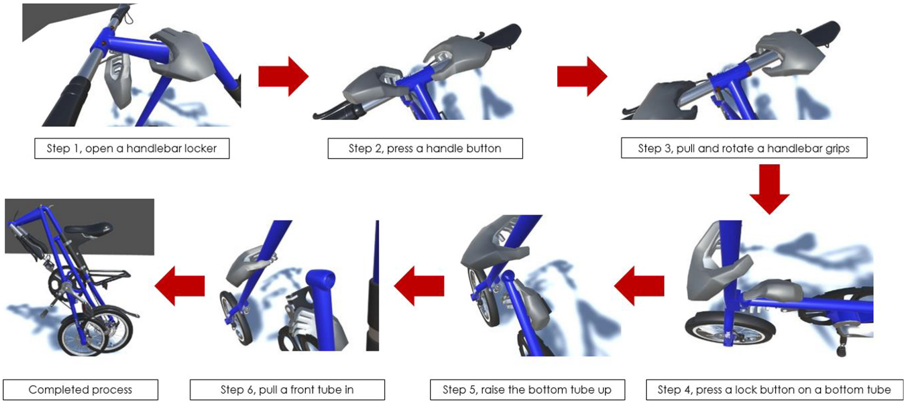


Fig. 8. STRIDA bicycle folding gesture sequence recorded in the VR environment

This recorded sequence can be used for multiple purposes, in particular for creating a visual work instruction, for instructing trainees both in real and virtual environments, as well as for studying typical mistakes trainees make during their manipulations. We can also think about analyzing the recorded sequences of different experts' or trainees' approaches to performing the tasks, and extract from them alternative ways of achieving the same result.

6 Summary, Conclusion and Outlook

In this article we presented intermediate results of an ongoing research aiming at creating an environment for the allowing to prepare industrial maintenance task trainings in VR environments. This preparation phase is particularly important, since investment in VR facilities need to be done in a very target-oriented way. Providing a training preparation environment in which the suitability of several different VR devices and training approaches can be tested and evaluated is our ultimate objective. So far, we presented a conceptual approach to such an environment, as well as a focus on the work instruction preparation part of it with the library of hand gestures being an essential building block. We showed the generic data model facilitating this task and demonstrated its feasibility in a case study that is comparable to a typical maintenance assembly/disassembly task in industry. The case study illustrates the relevant maintenance assembly/disassembly tasks in the preparation phase, which can be helpful for the trainee to acquire appropriate knowledge and expertise before performing them in the training phase. The case study was performed in a VR environment driven by virtual hands controlled by the user's real hands, as well as HMD with integrated LMC. Based on this initial proof of concept, we plan further extending the hand gesture library by more types of mechanical joints. We will investigate various ways of exploiting this library for the creation of work instructions and training plans, as well as the validation of the correctness of gestures during virtual trainings. In order to implement haptic perception, we plan integrating the CyberGlove into our platform. We also expect that using CyberGlove will allow recording the hand gestures more accurately and clearly. We will also integrate the work instructions generated from the gesture library as virtual work instructions (VWI) in the VR environment. In addition, we plan analyzing the economic, ecologic, and social benefits of virtual maintenance training compared to real-life maintenance training. We believe in the relevance of this work for the SPI Manifesto [19] in terms of providing a contribution to continuous learning process improvement and empowerment of people with cutting-edge ICT.

References

1. Liu, X., Cui, X., Song, G., Xu, B.: Development of a virtual maintenance system with virtual hand. Int. J. Adv. Manuf. Technol. **70**(9), 2241–2247 (2013). https://doi.org/10.1007/s00170-013-5473-0
2. Christiand, Yoon, J.: Assembly simulations in virtual environments with optimized haptic path and sequence. J. Robot. Comput.-Integr. Manuf. **27**(2), 306–317 (2011)
3. Bailey, S.K.T., Johnson, C.I., Schroeder, B.L., Marraffino, M.D.: Using virtual reality for training maintenance procedures. In: Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC), Paper No. 17108, pp. 1–11 (2017)
4. Guo, Z., Zhou, D., Chen, J., Geng, J., Lv, C., Zeng, S.: Using virtual reality to support the product's maintainability design: immersive maintainability verification and evaluation system. Comput. Ind. **101**, 41–50 (2018)
5. Randeniya, N., Ranjha, S., Kulkarni, A., Lu, G.: Virtual reality based maintenance training effectiveness measures – a novel approach for rail industry. In: IEEE 28th International Symposium on Industrial Electronics (ISIE), pp. 1605–1610 (2019)
6. Linn, C., Bender, S., Prosser, J., Schmitt, K., Werth, D.: Virtual remote inspection – a new concept for virtual reality enhanced real-time maintenance. In: 23rd International Conference on Virtual System & Multimedia (VSMM) (2017)
7. Vélaz, Y., Arce, J.R., Gutiérrez, T., Lozano-Rodero, A., Suescun, A.: The influence of interaction technology on the learning of assembly tasks using virtual reality. J. Comput. Inf. Sci. Eng. **14/041007**, 1–9 (2014)

8. Zhou, D., Zhou, X.-X., Guo, Z.-Y., Lv, C.: A maintenance time prediction method considering ergonomics through virtual reality simulation. SpringerPlus **5**(1), 1–22 (2016). https://doi.org/10.1186/s40064-016-2886-x
9. Ganier, F., Hoareau, C., Tisseau, J.: Evaluation of procedural learning transfers from a virtual environment to a real situation: a case study on tank maintenance training. Ergonomics **57**(6), 828–843 (2014)
10. Lu, G., Shark, L.K., Hall, G., Zeshan, U.: Immersive manipulation of virtual objects through glove-based hand gesture interaction. Virtual Reality **16**(3), 243–252 (2012). https://doi.org/10.1007/s10055-011-0195-9
11. Prachyabrued, M., Borst, C.W.: Virtual grasp release method and evaluation. Int. J. Hum.-Comput. Stud. **70**(11), 828–848 (2012)
12. Ullmann, T., Sauer, J.: Intuitive virtual grasping for non-haptic environments. In: Proceedings the Eighth Pacific Conference on Computer Graphics and Applications, pp. 373–379 (2000)
13. Xiao, Y., Peng, Q.: A hand gesture-based interface for design review using leap motion controller. In: Proceedings of the 21st International Conference on Engineering Design (ICED 2017), vol. 8: Human Behaviour in Design, pp. 239–248 (2017)
14. Shamsuzzoha, A., Helo, P., Kankaanpää, T., Toshev, R., Tuan, V.V.: Applications of virtual reality in industrial repair and maintenance. In: Proceedings of the International Conference on Industrial Engineering and Operations Management Washington DC, USA (2018)
15. Fletcher, J.D., Belanich, J., Moses, F., Fehr, A., Moss, J.: Effectiveness of augmented reality & augmented virtuality. Presented at MODSIM World 2017 Conference (2017)
16. Legardeur, J., Garreau, L., Couture, N.: Des interacteurs pour l'assemblage mécanique en CAO. In: Integrated Design and Production/Conception et Production Intégrées, CPI 2005, Casablanca, Morocco (2005)
17. Numfu, M., Riel, A., Noël, F.: Virtual reality based digital chain for maintenance training. In: 29th CIRP Design 2019, Procedia CIRP, pp. 1076–1081 (2019)
18. Numfu, M., Riel, A., Noël, F.: Virtual reality based digital chain for creating a knowledge base of hand gestures in maintenance tasks. In: 27th CIRP Life Cycle Engineering Conference 2020, Procedia CIRP (2020, in print)
19. Korsaa, M., et al.: The people aspects in modern process improvement management approaches. J. Softw. Evol. Process **25**(4), 381–391 (2013)

Evaluation of Gamification Elements in a VR Application for Higher Education

Florian Tiefenbacher(✉)

IMC University of Applied Sciences Krems, Krems an der Donau, Austria
florian.tiefenbacher@fh-krems.ac.at

Abstract. Virtual reality (VR) and gamification are two topics that are of interest to the higher education. This study was conducted with 102 students of the health department of IMC university of applied sciences. Two versions of a VR game have been designed, one with highscore and achievements and the other without those elements but with time pressure instead. Each group consisted of a cohort of students. The results showed that the group without highscore and achievements were more motivated to play the game and felt more supported by the game while studying than the group with those gamification elements. Even though the majority of students with time pressure stated that they found it disturbing, it did not have a negative effect on their overall evaluation of the game itself and their motivation. The outcome suggests that a highscore and achievements lead to higher competition and evaluation which is known to have negative impact on intrinsic motivation.

Keywords: Virtual reality · Gamification · Highscore · Achievements · Motivation

1 Introduction

The term gamification is often misused in media, when it appears that it can be used synonymously with the term gaming. But it is an independent concept used in a wide range of contexts. Gamification can be defined as the use of game design elements in non-game contexts [1]. There are a lot of different questions surrounding this concept – e.g. measuring the learning outcome, measuring the impact on motivation. In this paper the focus lies on the subjective perception of students in higher education and how gamification elements such as a highscore and achievements impact the motivation of students and which version make them feel more supported in their studies through a VR application. The hypothesis before this study was that a VR game for higher education should contain a highscore and achievements and no time pressure.

2 Background

There are topics in the curriculum that students need to learn by heart, mostly declarative knowledge. After speaking to two lecturers each selected a topic which would be suitable for additional support in form of a VR game for students. Therefore, the aim was to satisfy the needs of lecturers and students in supporting the educational process.

M. Yilmaz et al. (Eds.): EuroSPI 2020, CCIS 1251, pp. 830–847, 2020.
https://doi.org/10.1007/978-3-030-56441-4_63

There have been various studies on the use of instructional and educational games in higher education, but there are not many that tried to combine these elements with the emerging technology of virtual reality. Therefore, after intensive literature reviews, two versions of VR games have been programmed and designed in Unity. During the design process the needs of those two cohorts and their curriculum has always been in mind.

3 Research and Theory

3.1 Gamification

On the one hand there are a meta-analysis that looked on various studies related to gamification. J.J. Vogel [2] did an analysis in 2006 where 32 studies were analysed. The outcome was that there seems to be an increase in cognition and memory when using gamification in comparison to conventional learning scenarios. Another fact that this analysis brought forward was, that this effect is not dependent on age and gender. This finding is relevant for this study, as the participants are not equally divided in gender and age. Another meta-analysis from Fengfeng Ke [3] in 2009 looked on 89 articles about the effects of game-based learning methods. The outcome was that 52% of those articles found a significant positive effect of gamification in comparison to conventional learning methods. Only 18% found no effect. Another fact that was brought forward was that the motivation of participants seems to increase after using gamified methods. A topic that is relevant in this paper is the use of game-based learning scenarios in higher education. The following 5 recent studies have been found relevant:

- Li et al. [4] did a qualitative and quantitative study with 14 participants on a gamified way to learn AutoCAD. It involved highscore, achievements and time pressure. The outcome was an increase in participant engagement and motivation.
- McDaniel et al. [5] did a quantitative study with 138 participants about achievements and highscores in learning games. The outcome was just a small effect on motivation and women were more spurred on from achievements than men.
- Dominguez et al. [6] did a quantitative study with 211 participants involving badges and highscores. The outcome was a positive effect on practical exercises but a negative effect on written exercises.
- Hakulinen et al. [7] did a quantitative study with 281 participants involving just badges. The outcome was that just a part of the participants showed an increase in motivation whereas most participants did not react to badges.
- Banfield and Wilkerson [8] did a qualitative and quantitative study with 96 participants on the effects of a highscore to intrinsic motivation. The outcome showed an increase of intrinsic motivation.

After analysing those studies, it becomes evident that there are different outcomes regarding the effect of game-based learning on motivation and memory.

3.2 Motivation

A very important factor in the concept of gamification is motivation. There are a lot of different theories and studies around this topic. In this paper there will be a concise look on some important theories.

Theory of Intrinsically Motivation Instruction

T. Malone [9] did a study in 1981 which was the basis of his theory that wants to explain why games are motivating. He defined 3 key aspects: challenge, fantasy, and curiosity. Those aspects must be present in a game to be motivating to the player. He also divided fantasy in intrinsic (e.g. active negotiations with a game character) and extrinsic (e.g. preventing a hurricane) fantasy. Malones theory stresses that intrinsic motivation is key, and that curiosity is dependent on the complexity of information.

Instructional Design Principles for Intrinsic Motivation

M. Lepper [10] from Stanford university defined design principles for lecturers to arouse intrinsic motivation. The 4 principles are: Control, Challenge, Curiosity and Contextualization. Students should have the feeling of control in every situation, they should also be constantly challenged. Regarding the setting Lepper is advising a realistic or fantasy setting with the main goal of being authentic.

Flow Theory

When it comes to game-based learning, there is almost no way around the flow theory from Mihaly Csikszentmihalyi [11]. The theory describes a state called flow in which the time is flying, and the cognitive load is ideal. The two main factors to reaching this state is a perfect balance between challenge and skill. He describes that too much challenge and skill requirements lead to anxiety, while too less of that leads to boredom (Fig. 1).

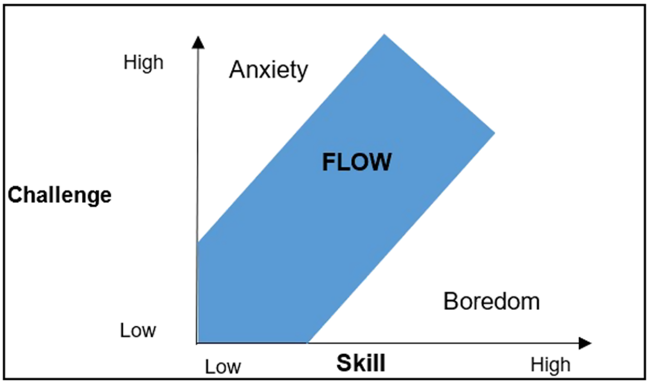


Fig. 1. Flow theory

Self-Determination Theory

A very essential meta-theory of motivation is the self-determination theory of Deci & Ryan [12, 13]. There are six mini theories that each cover another field of motivational research. All of these are relevant when trying to design a game for higher education, but to keep it short for this paper the focus will lie on the following two sub theories:

Cognitive Evaluation Theory

In the CET from Ryan & Deci [14] the focus lies on the social contexts on intrinsic motivation. It is looking at how factors like rewards impacts the intrinsic motivation. It stresses the importance of autonomy and competence when trying to increase intrinsic motivation. This theory is a key to education and game-based learning.

Goal Contents Theory

This theory is focusing on the difference between intrinsic and extrinsic motivators or goals and their impact on motivation [15]. This is relevant due to the fact that in game design some goals are intrinsic while other are extrinsic. The fundamental assumption in this theory is that intrinsic goals correspond to human basic needs while extrinsic goals do not correspond. It is assumed that intrinsic goals therefore have greater impact on wellbeing than extrinsic goals.

To complete the six theories of self-determination from Deci & Ryan here are the other four theories:

- Organismic Integration Theory
- Causality Orientations Theory
- Basic Psychological Needs Theory
- Relationships Motivation Theory

3.3 Psychological Needs

The self-determination theory is stressing the importance of psychological needs when it comes to motivation. Therefore, it is important when designing a game for higher education with the aim of motivating the students. There are 3 key factors to psychological needs that are relevant:

Competence

An essential factor of competence is interaction, as humans want to understand and control his surroundings [15]. When designing a game, it is important to make sure that the player does not have the feeling of being controlled by game mechanisms, because that could lead to negative effects on intrinsic motivation. Game design elements should challenge the player to show his skills and the ideal situation would be to induce a flow state as described before. Another important factor is feedback – as Rigby and Ryan [16] stated there are three different kind of competence feedback:

- Granular feedback: immediately connected to the action – in this study that would be the points gained for the highscore
- Sustained feedback:
  is connected to a certain time span in which an action is carried out.
- Cumulative feedback:
  is connected to the skills of the player and summarize the actions of the player. This feedback is persistent even after finishing the game – in this study that is represented in the form of achievements and the highscore.

Autonomy
The main factor here is the psychological freedom or will to carry out a certain action [15]. It is important that the player has the feeling of being the cause of certain actions and that he can make decisions autonomous. It could lead to negative effects if the game design is causing pressure or force some actions onto the player. This assumption leads to the fact that one group in the study will have time pressure, while the other group will not have a time limit when answering the questions. To increase autonomy there should be autonomous decision alternatives and informative feedback [17].

Social Relations
To fulfil this human need it is important that there is a feeling of a bond to other people [13]. There are different ways to reach this goal – praise, recognition and being able to witness actions of others are elements that can fulfil this need. In the context of game design this would mean including other players in one way or the other. In this study there will be the social component of competition within a group. As Rigby and Ryan [16] found, competition can increase the social relations and individual skills.

3.4 Cognition

There are two main types of cognition, one is experimental cognition and the other is reflexive cognition [18].

Experimental Cognition
This type of cognition describes thinking processes while interacting with the surroundings. There are three categories of experimental cognition – they will be described in the context of game-based learning:

- Attention
  It is important to gain attention from the user through visual and auditive signals. There are different ways to do that, like colorized highlighting, animations, and a clear structure. When designing a game, it is important to find the right balance, as too many signals can lead to over-excitement.
- Perception
  The focus when it comes to perception is the reception of sensory stimuli. Especially in virtual reality this factor plays a huge role in game design. Font size, contrasts, auditive and haptic signals to differentiate are design methods to find the right amount of perception.
- Memory
  Memory is the process of saving information for a future retrieve. This process is tightly woven with the attention, as more attention leads to a better memorization of information.

Reflexive Cognition
This type of cognition describes thinking processes while studying, brainstorming, deciding etc. When designing a game, it should be taken care off, that the cognitive load is not too big, and that the player is able to decide freely and without too much effort.

3.5 Player Types

When it comes to game design, there is a categorization that must be considered. Richard Bartle defined four types of players [19].

- Achiever
  Players of that type want to get achievements and gain a status to show others what they have done. Their main satisfaction is challenge, other activities that are less challenging are less attractive to this type of players.
- Explorer
  This type of players wants to discover the virtual worlds and be fully immersed in those. Their main satisfaction is getting to discover something new, while achievements are less appealing to them.
- Socializer
  The socializer type of player describes those that wants to build relationships with other players. They cherish the social interactions and communications within the virtual world.
- Killer
  This type of players wants to overcome other players.
  Each type of players and their characteristics must be considered when designing a game.

3.6 Virtual Reality

In a recent paper Jensen and Konradsen [20] did an analysis of 22 studies on this topic. 14 studies were user-centered, investigating factors like presence and immersion and the users' attitude towards VR in education. They found that the big issue with most studies are that it is difficult to evaluate the learning outcome impact of VR in comparison to other mediums.

Janssen et al. [21] found in a study that the personality of the user has an impact on immersion and therefore on the VR experience. The outcome of their study suggest that introverted people are less immersed and tend to have less positive experiences in VR. This would have to be considered when designing a game for students, as there are different types of personality throughout the cohort.

In a study conducted by Alhabi [22] the outcome showed that there was a positive effect on motivation that was directly dependent on immersion. The students were more motivated to spend time in the VR application when they felt more immersed. There is a small margin between immersion and over-excitement, which is especially important considering when designing a VR experience for educational purpose.

As mentioned in 2.4 under attention, the application should gain attention, but too many signals would undermine the positive effects. In a study Fernandes et al. [23] found that too much sensory stimuli through realistic noises, simulated hands etc. can have negative effects on users and therefore have a negative effect on their cognitive ability to concentrate on learning activities while using VR.

Cognitive Effects of VR in Learning Scenarios

In a field-study by Ray and Deb [24] with students they compared the effects of VR learning with those of traditional learning methods. The outcome showed, that in the first two weeks there were could be no difference found, but after those two weeks students using VR had slightly better results than those who did not.

A very interesting study from Makransky et al. [25] used EEG measurement to determine whether VR leads to a higher cognitive load while learning than traditional learning methods. The outcome showed that VR can increase the cognitive load and therefore could be beneficial when studying.

Another study that compared VR with other learning methods was carried out by Stepan et al. [26] They did their experiment with students of neuroanatomy and they let one group study with VR glasses and the other with online books. The outcome was no significant difference between the two groups.

4 Study Design

This study was conducted in a non-randomized controlled trial design. Two cohorts of students of the health department of the IMC university of applied sciences Krems participated in this study. Each cohort received questions from their respective current curriculum, therefore the lecturers gave feedback on question design and elaborated a list of questions and answers for their students. In each cohort the students participated voluntarily, and the experiment was taken out during course times.

4.1 Participants

Since two cohorts have been chosen to participate with their specific questions it was not possible to select a sample that is equally distributed in gender and age. In the two picked cohorts of the health courses there are a lot more female students, which is reflected in the participants of the study (Table 1).

Table 1. Age structure

Age	Participants
18 to 24	66
25 to 32	19
33 to 44	9
45 or older	8

- Participants overall n = 102 − female: 93/male: 9

4.2 Environments

Both groups had the exact same environment. The study took place in the eVRyLab – the VR Lab of the IMC university of applied sciences Krems. There were two workplaces with a pc and an Oculus Rift S for the participants.

4.3 Test Procedure

There have been different measures to ensure the same test procedure for every participant. As stated before, the environment was the same for both groups and every participant. Before playing the VR game every participant had to take the first of two questionnaires in which general information, attitude towards gamification and VR and prior experiences with those have been examined. Every participant received a short introduction after finishing the first questionnaire. Then the participants got the VR device and has been supported during the individual adjustment of the device. During the game there has been no further help and immediately after finishing the game the participants were asked to fill out the second questionnaire about their experience during the VR game.

4.4 Tested Gamification Elements

- Group A:
  This group consisted of 45 participants all from the same cohort. They played the VR game without highscore and achievements but with time pressure.
- Group B:
  This group consisted of 57 participants all from the same cohort. They played the VR game with highscore and achievements and without time pressure.

VR Game Development
Both versions of the game were programmed and designed in Unity. They were made with the OpenVR SDK, but in the experiment they were only played with Oculus RiftS headsets. The game has three levels with three different methods of answers/questions. The game starts in a Start room where the students had their two magic wands and were asked to type in their names by touching the respective letter (Fig. 2).

Fig. 2. Interaction with magic wands

Fig. 3. Start & Endroom with highscore & achievements

For Group B there was a start room which is also the end room and displays the actual highscore and the achievements that are unlocked by the player (Fig. 3).

Level 1 had a simple question and answer system – displaying the question on the board while the student had to choose which of the two answers are correct.

Leve 2 displayed a medical term (e.g. Meiose) and the students had to choose which explanation fits.

Level 3 displayed a statement on the board and the student had to choose whether it was right or wrong.

The whole interaction was done with two magic wands, to interact with the answers they simply had to touch it with one of the wands. They did not have to press any key throughout the whole game. This was a design choice to simplify the interaction so that the students could concentrate on the learning materials rather than learning the interaction first. The feedback whether an answer was correct or not was instantly given by auditive signals and by visual signals – changing the colour (red or green) – see Fig. 4.

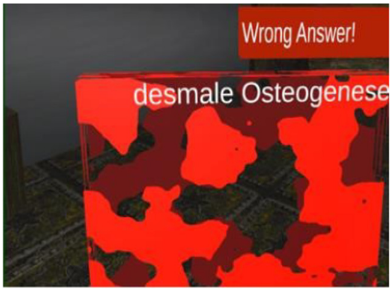


Fig. 4. Changing to colour to red (Color figure online)

After finishing a level, the player had to interact with a “Next Level”-button, which was a design choice to increase autonomy of the player.

Group B could unlock certain achievements – e.g. 1000 points reached – as seen in Fig. 6. When the player unlocked an achievement, it was celebrated with a special auditive signal, and visual signals like animated fireworks around a pop-up achievement bar (Fig. 5).

Fig. 5. Achievement unlocked – bar & animations

Fig. 6. Achievement unlocked - bar & animations

5 Results and Discussion

5.1 Questionnaire 1 – Before the Application

The focus was on gathering data about the students' attitude and previous experiences regarding VR and gamification. Both groups had the same questions – some interesting outcomes are following:

Q: "How often did you have used a VR headset?"

53,9% (n = 55) had never used a VR headset
42,2% (n = 43) had one or two experiences and
just 3,9% (n = 4) had three or more experiences.

Q: "How often have you used learning games?"

36,3% (n = 37) had never used learning games
47,1% (n = 48) had one or two experiences
and 16,7% (n = 17) had three or more experiences

The following answers consists of numbers – 1 means "I fully agree", while 5 means
"I fully disagree"

Q: "I would welcome the use of VR in lessons."

1: 52,9% (n = 54)
2: 23,5% (n = 24)

3: 15,7% (n = 16)
4: 4,9% (n = 5)
5: 1,0% (n = 1)
2 participants did not answer (2,0%)

This outcome shows that the majority of students would welcome VR technology in lessons at their university. It also shows that most of the 102 participants had a positive attitude towards VR and its use in education.

Q: "A highscore within my cohort would motivate me to play the game more often."

1: 20,6% (n = 21)
2: 24,5% (n = 25)
3: 25,5% (n = 26)
4: 16,7% (n = 17)
5: 12,7% (n = 13)

Interestingly a slight tendency to the middle is noticed. Almost one third of all participants did not think that a highscore would motivate them before being confronted with the VR game itself.

Q: "Achievements would motivate me to play the game more often."

1: 17,6% (n = 18)
2: 21,6% (n = 22)
3: 25,5% (n = 26)
4: 23,5% (n = 24)
5: 11,8% (n = 12)

Even more participants did not fancy the idea of achievements as a motivator for themselves.

Q: "It would motivate me to answer under time pressure."

1: 17,6% (n = 18)
2: 25,5% (n = 26)
3: 28,4% (n = 29)
4: 15,7% (n = 16)
5: 12,7% (n = 13)

The majority of participants show a tendency to the middle, while 12,7% did not like the idea of time pressure 17,6% said they would be totally motivated by it.

5.2 Questionnaire 2 – After the Application

The following answers will be displayed in a boxplot. On the left is Group A and on the right is Group B. Starting from the bottom is 1 ("I fully agree") while at the top there is 5 ("I fully disagree").

Q: "I had the feeling of learning the contents while playing."

An obvious difference between the two groups can be observed in this question. The group with highscore & achievements did feel a lot less like learning the contents than the group without those elements but with time pressure instead. Group A has its median at 1 while Group B has its median at 3 which shows the evident difference between the two groups (Fig. 7).

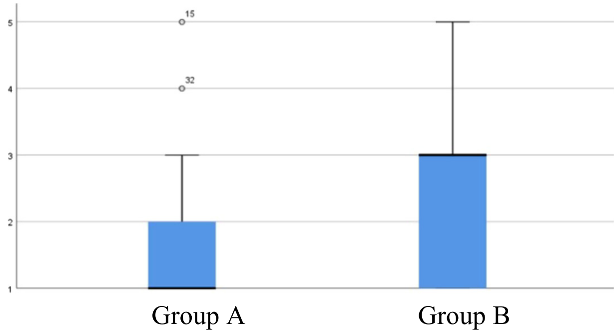


Fig. 7. Boxplot: Feeling of learning contents

Q: "I would like to play the learning game more often."

This question about motivation to play the game more often is not as obviously different as the question before, but still group A did evaluate this question more positive than group B. Group A has its median at 1 an just 3 participants gave a negative rating – Group B has its median at 2 and 4 participants gave a negative rating (Fig. 8).

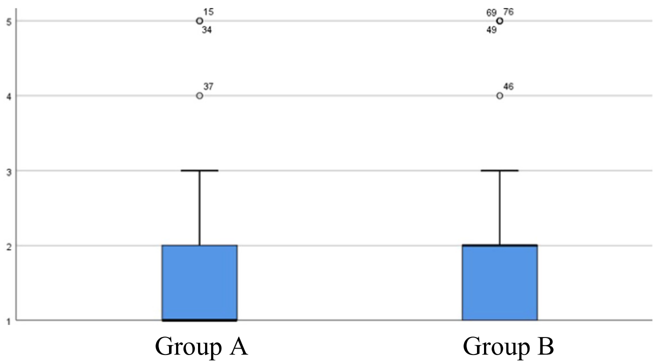


Fig. 8. Boxplot: Motivation

Q: "The interaction was intuitive and uncomplicated."

Group A did find the interaction more intuitive and uncomplicated than group B. While group A has its median at 1 with only one participant evaluating above 2 – group B has its median at 2 (Fig. 9).

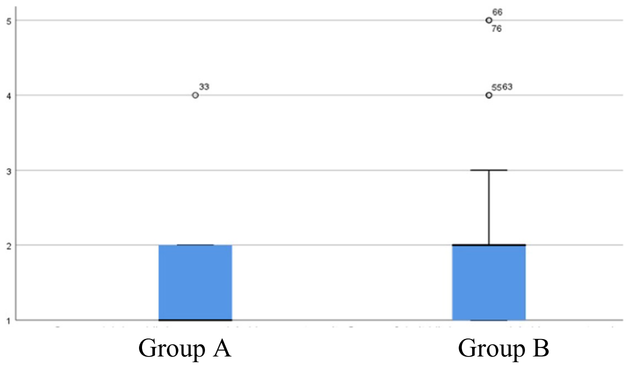


Fig. 9. Boxplot: Interaction

Q: "The game gave me the feeling of being fully immersed in a virtual world."

Against the trend group B did feel more immersed into the virtual world than group A. A possible explanation could be the additional highscore board and achievement trophies that were missing in the application of group A. Both groups have their median at 2, group B has its maximum at 3 & 4 outliers while group A has its maximum at 4 (Fig. 10).

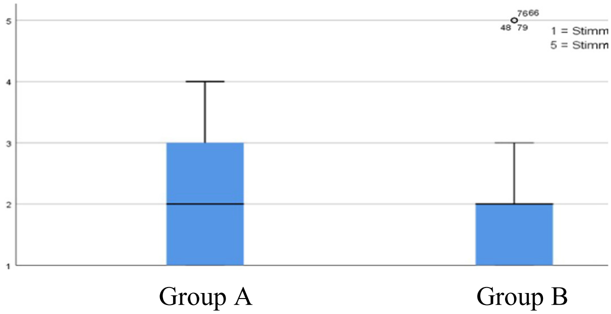


Fig. 10. Boxplot: Immersion

Q: "Learning games like this would support me while studying."

This is a subjective question which group A evaluated more positive than group B. Group A has its median at 1 and maximum without 2 outliers at 3 – Group B has its median at 2 and its maximum at 5. The group with highscore and achievements therefore had less often the feeling that the game would support them while studying (Fig. 11).

Q: "I would use learning games like this at home."

Once again group A did evaluate more positive than group B. This time the group without highscore and achievements but with time pressure instead would use games like this at home more often than the group with those elements but without time

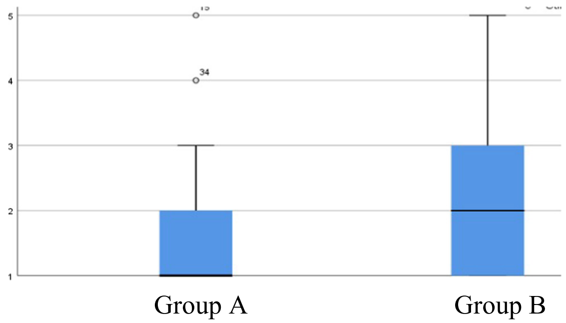


Fig. 11. Boxplot: Support while studying

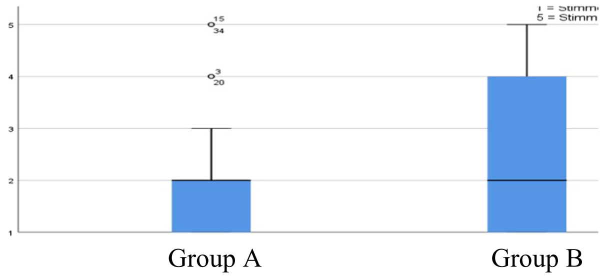


Fig. 12. Boxplot: Private Use

pressure. Group A has its median at 2 and its maximum without 4 outliers at 3, while group B has its median at 2 too, the maximum is at 5 with 50% giving a 3 or higher (Fig. 12).

The following questions were group-dependant, they were designed to get specific data about the respective methods used in the VR game. The answers could range between 1 (“I fully agree”) and 5 (“I fully disagree”).

5.3 Group A

Q: “The time pressure was disturbing.”

1: 31,1% (n = 14)
2: 20,0% (n = 9)
3: 17,8% (n = 8)
4: 11,1% (n = 5)
5: 20,0% (n = 9)

Even though group A did evaluate the VR experience more positive throughout the questionnaire – 51,1% did find the time pressure disturbing. This seemingly had no effect on this group when it comes to the application overall.

5.4 Group B

Q: "The highscore motivates me to play the game more often."

1: 35,1% (n = 20)
2: 17,5% (n = 10)
3: 22,8% (n = 13)
4: 12,3% (n = 7)
5: 12,3% (n = 7)

52,6% did either agree or fully agree to this question, but there seems to be less effect of the highscore when it comes to motivation and the feeling of learning the contents overall.

Q: "The achievements motivate me to play the game more often."

1: 22,8% (n = 13)
2: 19,3% (n = 11)
3: 21,1% (n = 12)
4: 17,5% (n = 10)
5: 19,3% (n = 11)

This outcome shows that achievements were less appealing and motivating to group B than the highscore.

5.5 Significant Correlations Within the Dataset

The interpretation of this correlation will be using the Cohen's [27] conventions – the correlations itself will be displayed in the Pearson convention.

Q: "I would like to play the learning game more often."

- ,837 correlation with "The game was fun."
- ,600 correlation with "I had the feeling of learning the contents while playing."
- ,418 correlation with "The game gave me the feeling of being fully immersed in a virtual world."
- ,640 correlation with "The highscore motivates me to play the game more often."
- ,577 correlation with "The achievements motivate me to play the game more often."

A very significant positive correlation (,837) when it comes to motivation is with fun. Those who had the feeling of learning the contents also seem to be more motivated to play the game more often – and vice versa. (,600) Other correlations could be witnessed with immersion, the highscore and achievements.

Q: "The game was fun."

- ,559 correlation with "I had the feeling of learning the contents while playing."
- ,436 correlation with "The game gave me the feeling of being fully immersed in a virtual world."
- ,482 correlation with "The interaction was intuitive and uncomplicated."
- ,640 correlation with "I would use learning games like this at home."

The biggest positive correlation with fun was with using learning games at home (,640). Also, those who found the interaction intuitive and uncomplicated had more fun and vice versa. The same goes for immersion and the feeling of learning the contents.

Q: "I had the feeling of learning the contents while playing."

- ,386 correlation with "The interaction was intuitive and uncomplicated."
- ,283 correlation with "The game gave me the feeling of being fully immersed in a virtual world."

A slight correlation between the feeling of learning contents and immersion have been found (,283) and a moderate correlation with interaction.

6 Conclusion

This study was carried out in a controlled environment and with two picked cohorts. Both groups therefore were not picked to represent all students at the university, they only included of students' courses in the health department of IMC university of applied sciences Krems. During the experiment it became evident, that it would have been better to show the highscore and achievement count to the students that were not playing the game and that one person at a time in the labor leading to more intensive instructions on how to adjust the VR glasses could have prevented some minor issues with the adjustment of the VR glasses. The results showed that both groups had a very positive attitude towards the use of VR in their lessons which is a take-away for other teachers and lecturers.

The outcome showed that the initial hypothesis that a VR game for higher education should include a highscore and achievements, but no time pressure has been falsified. The group with those elements but without time pressure were less motivated to play the game more often and did feel less supported by a game like this. Group A did feel to learn the contents more often than group B which suggests that a highscore and achievements can lead to more competition and evaluation. Those two factors are known to have negative impact on intrinsic motivation. It was also shown that time pressure did not have a negative impact on the participants when it comes to motivation, even though the majority did state they found it disturbing. A factor that has initially been not paid close attention to was fun, as the outcome suggests that the biggest correlation with motivation in both groups was fun. While also interaction and immersion correlated heavily with the motivation of the participants. The results suggest that the group that found the interaction more intuitive and uncomplicated also were more motivated and had more fun while playing.

There are still a lot of questions to answer, like evaluating those outcomes against traditional learning methods and against other gamification elements like storytelling and more social components with multiplayer modes.

References

1. Deterding, S., Khaled, R., Nacke, L.E., Dixon, D.: Gamification: toward a definition. In: Mindtrek Proceedings, Tampere. ACM Press (2011)
2. Vogel, J.J., Vogel, D.S., Cannon-Bowers, J., Bower, C.A., Muse, K., Wright, M.: Computer gaming and interactive simulations for learning: a meta-analysis. J. Educ. Comput. Res. **34** (3), 229–243 (2006)
3. Ke, F.: A qualitative meta-analysis of computer games as learning tools. In: Effecitve Electronic Gaming in Education, no. 1, pp. 1–32 (2009)
4. Li, W.I., Grossman, T., Fitzmaurice, G.: Gamicad: a gamified tutorial system for first time autocad users. Cambridge: Paper Presented at the 25th Annual ACM Symposium on User Interface Software and Technology (2012)
5. McDaniel, R., Lindgren, R., Friskics, J.: Using badges for shaping interactions in online learning environments. Orlando: Paper Presented at the IEEE International Professional Communication Conference (2012)
6. Dominguez, A., Saenz-de-Navarrete, J., de-Marcos, L., Fernández-Sanz, L., Pagés, C., Martínez-Herráiz, J.-J.: Gamifying learning experiences: practical implications and outcomes. Comput. Educ. **63**(2), 380–392 (2013)
7. Hakulinen, L., Auvinen, T., Korhonen, A.: Empirical study on the effect of achievement badges in TRAKLA2 online learning environment. Macau: Paper Presented at the Learning and Teaching in Computing and Engineering (2013)
8. Banfield, J., Wilkerson, B.: Increasing student intrinsic motivation and self-efficacy through gamification pedagogy. Contemp. Issues Educ. Res. **7**(4), 291–298 (2014)
9. Malone, T.: Towards a theory of intrinsically motivating instruction. Cogn. Sci. **5**(4), 333–369 (1981)
10. Lepper, M.R.: Motivational considerations in the study of instruction. Cogn. Instr. **5**(4), 289–309 (1988)
11. Csikszentmihalyi, M.: Play and intrinsic rewards. J. Humanist. Psychol. **15**(3), 41–63 (1975)
12. Deci, E.L., Ryan, R.M.: Die selbstbestimmungstheorie der motivation und ihre bedeutung für die Pädagogik. Zeitschrift für Pädagogik **39**(2), 223–238 (1993)
13. Deci, E.L., Ryan, R.M.: The what and why of goal pursuits: human needs and the self-determination of behavior. Psychol. Inq. **11**(4), 227–268 (2000)
14. Ryan, R.M., Deci, E.L.: Overview of self-determination theory: an organismic dialectical perspective. In: Ryan, R.M., Deci, E.L. (eds.) Handbook of Self-Determination Research, Herausgeber, pp. 3–33. University of Rochester Press, Rochester (2002)
15. Vansteenkiste, M., Ryan, R.M.: On psychological growth and vulnerability: basic psychological need satisfaction and need frustration as a unifiying principle. J. Psychother. Integr. **23**(3), 263–280 (2013)
16. Rigby, C.S., Ryan, R.M.: Glued to Games: How Video Games Draw us in and Hold us Spellbound. Praeger, Santa Barbara (2011)
17. Vansteenkiste, M., Niemic, C.P., Soenens, B.: The development of the five mini-theories of self-determination theory: an historical overview, emerging trends and future directions. In: Urdan, T.C., Karabenick, S.A. (eds.) The Decade Ahead: Theoretical Perspectives on Motivation and Achievement, Herausgeber, pp. 105–165. Emerald Group Publishing, London (2010)
18. Preece, J., Rogers, Y., Sharp, H.: Interaction Design: Beyond Human-Computer Interaction. Wiley, New York (2007)
19. Bartle, R.: Hearts, Clubs, Diamonds, Spaces: Players Who Suit MUDS. Muse Ltd., Colchester (1996)

20. Jensen, L., Konradsen, F.: A review of the use of virtual reality head-mounted displays in education and training. Educ. Inf. Technol. **23**(4), 1515–1529 (2017). https://doi.org/10.1007/s10639-017-9676-0
21. Janssen, D., Tummel, C., Richert, A., Isenhardt, I.: Towards measuring user experience, activation and task performance in immersive virtual learning environments for students. S. Graz University of Technology & M. Computer, vol. 621, pp. 45–58 (2016)
22. Alhalabi, W.S.: Virtual reality systems enhance students' achievements in engineering education. Beahv. Inf. Technol. **35**(11), 919–925 (2016)
23. Fernandes, L.M.A., Matos, G.C., Azevedo, D., Nunes, R.R., Paredes, H., Morgado, L.: Exploring educational immersive videogames: an empirical study with a 3D multimodal interaction prototype. Behav. Inf. Technol. **35**(11), 907–918 (2016)
24. Ray, A.B., Deb, S.: Smartphone based virtual reality systems in classroom teaching - a study on the effects of learning outcome. Mumbai: Paper Presented at the 8th IEEE International Conference on Technology for Education (2016)
25. Makransky, G., Terkildsen, T.S., Mayer, R.E.: Adding immersive virtual reality to a science lab simulation causes more presence but less learning. Learn. Instr. **60**, 225–236 (2017)
26. Stepan, K., et al.: Immersive Virtual Reality as a Teaching Tool for Neuroanatomy: Immersive VR as a Neuroanatomy Teaching Tool. International Forum of Allergy & Rhinology, San Diego (2017)
27. Cohen, J.: Statistical Power Analysis for the Behavioral Sciences, 2nd edn. L. Erlbaum Associates, Hillsdale (1988)

Author Index