

Progress in IS

Claudia Koschtial
Thomas Köhler
Carsten Felden *Editors*

e-Science

Open, Social and Virtual Technology for
Research Collaboration

OPEN ACCESS

 Springer

Progress in IS

“PROGRESS in IS” encompasses the various areas of Information Systems in theory and practice, presenting cutting-edge advances in the field. It is aimed especially at researchers, doctoral students, and advanced practitioners. The series features both research monographs that make substantial contributions to our state of knowledge and handbooks and other edited volumes, in which a team of experts is organized by one or more leading authorities to write individual chapters on various aspects of the topic. “PROGRESS in IS” is edited by a global team of leading IS experts. The editorial board expressly welcomes new members to this group. Individual volumes in this series are supported by a minimum of two members of the editorial board, and a code of conduct mandatory for all members of the board ensures the quality and cutting-edge nature of the titles published under this series.

More information about this series at <http://www.springer.com/series/10440>

Claudia Koschtial · Thomas Köhler · Carsten Felden
Editors

e-Science

Open, Social and Virtual Technology for
Research Collaboration



Springer

Editors

Claudia Koschtial
TU Bergakademie Freiberg
Freiberg, Germany

Thomas Köhler
Media Center
TU Dresden
Dresden, Germany

Carsten Felden
TU Bergakademie Freiberg
Freiberg, Germany



ISSN 2196-8705

Progress in IS

ISBN 978-3-030-66261-5

<https://doi.org/10.1007/978-3-030-66262-2>

ISSN 2196-8713 (electronic)

ISBN 978-3-030-66262-2 (eBook)

© The Editor(s) (if applicable) and The Author(s) 2021. This book is an open access publication.

Open Access This book is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this book are included in the book's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the book's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

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

Introduction

This publication, *e-science: The enhanced science*, is a collection of conference papers, reviewed and selected in a double-blind review process by a distinguished reviewer committee. From the very beginning when John Taylor introduced the term, e-science did not only comprise infrastructure as an enabler of scientific discovery, but also “global collaboration in key areas of science” (Taylor 1999). As computer technologies and digital tools pervade the academic world, it is time to ask what changes are implied when an “e” is added to science. What is primarily discussed in Germany and Great Britain under the term e-science corresponds in the USA to the concept of cyber infrastructures and in Australia to the concept of e-research.

More recently the discourse about e-science has been dealing with collaborative research that is based on a comprehensive digital infrastructure. This infrastructure both ultimately integrates all relevant resources for a research domain in a digital format and provides tools for processing such data. In computing-intensive research scenarios, e-science includes distribution of computing capacities, supporting collaborative processes of a rather inter-institutional character, such as (inter)national networks. The open innovation approach creates new platforms for developing and publishing research results. For example the MOVING platform (<http://moving-project.eu/moving-platform/> cf. Vagliano et al. 2018) supports new collaborative research practices and has become a resource for further research.

In this sense and in addition to the technological aspect (virtualization of hardware), e-science also has a social and politics-of-science aspect (cooperative research, reusability of data and interoperability of digital tools). Although there is the will to expand e-science methods into the wider economy and society, this development is occurring slowly. New skill sets are being acquired in the e-humanities, virtual engineering or visual analytics (Redecker and Punie 2017; Köhler 2018). Yet e-science also comprises open access, e-learning and grid computing; these changes are enabled by state funding and public interest. As a result, the concept of e-science continues to generate new concepts for particular disciplines such as e-geography, e-humanities, e-medicine or e-engineering.

The 2014 International Conference on Infrastructures and Cooperation in e-Science and e-Humanities reflected the broad ongoing discussion concerning the changes affecting research and teaching in universities nowadays. It addressed current

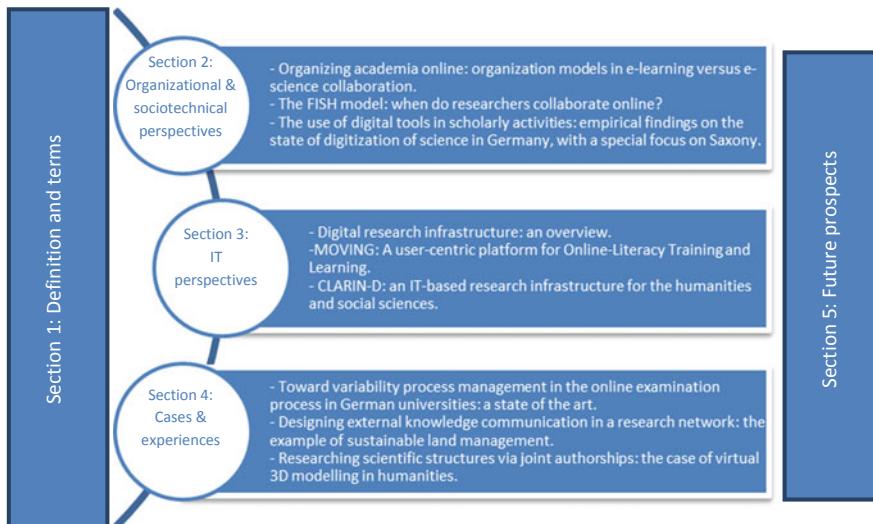


Fig. 1 Structure of “e-science: the enhanced science”

questions and solutions related to technologies or applications as well as their implications for the conduct of science. It investigated digitally enhanced academic initiatives from technological and socio-scientific perspectives.

This volume is subdivided into five sections representing different perspectives on e-science, as seen in the figure below. The first section introduces the book and reviews the literature concerning the definition of e-science. Section 2 provides organizational and socio-technical perspectives, especially the use of web 2.0 tools from an individual viewpoint and the successful implementation of such tools from an organizational viewpoint. As e-science of course relates to information technology, Section 3 covers IT perspectives, and Section 4 presents domain-specific cases and experiences. Finally, the proceedings close with future prospects (Fig. 1).

The introductory section of the proceedings *Digital research infrastructure: an overview* starts out with C. Koschtial’s contribution, an analysis of the terms covered by the field of digital research, that is, e-science itself, and related terms like cyber-science or science 2.0. As e-science is a socio-technical system, it can be approached from the perspective of the human user, the task or the technology, as identified by Heinrich (1993, pp. 8). The aim is to identify the dominant approach to e-science, to distinguish between the different terms and identify how the terms reflect changes in the prevailing research streams.

Section 2 deals with individual usage of tools and organizational enablement of this. The first paper of the second section, authored by T. Köhler, C. Lattemann and J. Neumann, is entitled *Organizing Academia Online: Organization models in e-learning Versus e-science Collaboration*, identifies forms of organizational governance enabling effective e-collaboration for scientists. Organizational governance

captures (social, output or behavioural) controls that are suitable for effective e-collaboration in scientific communities. Based on three case studies, the author identifies IT as a key factor in successful virtualization and concludes that there is a need for virtualized organization models which refer to processes and structure. The second contribution from in this section by B. Mohamed and T. Köhler investigates individual researchers and their will to use web 2.0 tools. In the third paper, focus on conceptualizing and validating digital research collaboration between novice researchers. Based on the FISH model, an online survey of 140 novice researchers was carried out and analysed using Partial Least Squares for the analysis of the data. One main result is that successful usage of online tools enhances the belief in web 2.0 as a useful instrument. The second main result is that benefits experienced by sharing enhance motivation for collaboration. Based on an online study comparing Germany as a whole with the federal state of Saxony, the final contribution of the second section authored by S. Albrecht, C. Minet, S. Herbst, D. Pscheida and T. Köhler presents research into the extent to which digital tools are adopted. One finding is that certain tools are now used by more than the half of the scientists in their daily professional life, but web 2.0 tools like microblogs and social networking sites are used far less often.

In Section 3, the focus is on digital tools or information infrastructures, which have not been considered yet. The first paper contributed by O. Schonefeld, M. Stührenberg and A. Witt in this section discusses important guidelines for research infrastructures, which are used to support teaching, research and young researchers. Regarding IT, research infrastructures should be maintained in collaboration between organizations. To reduce costs, energy efficient or green, technologies should be considered, and secure networks are needed enabling to minimize risks. Concerning the aspect of information infrastructure, the authors stress the relevance of data repositories and publication servers in a format that allows the stored documents or data to be used in the long term. Further important considerations regarding research infrastructures include copyright laws with specific national regulations and personal data protection. Accordingly, the authors identify a need for an IT strategy and corresponding roles such as that of data protection officer in organizations providing a research infrastructure.

The second paper authored by A. Apaolaza, T. Backes, S. Barthold, I. Bienia, T. Blume, C. Collyda, A. Fessl, S. Gottfried, P. Grunewald, F. Günther, T. Köhler, R. Lorenz, M. Heinz, S. Herbst, V. Mezaris, C. Nishioka, A. Pournaras, V. Sabol, A. Saleh, A. Scherp, U. Simic, A.M.J. Skulimowski, I. Vagliano, M. Vigo, M. Wiese and T. Zdolšek Draksler introduces *MOVING: A User-Centric Platform for Online Literacy Training and Learning*. The platform enables the usage of machine learning for searching, organizing and managing unstructured data sources. The data sources comprise but are not limited to publications, videos or social media. The contribution presents the web platform from a user-centred perspective in order to give an overview of the functionalities.

The final paper of Section 3 from G. Heyer and V. Boehlke presents a research infrastructure called CLARIN-D. This is a web-based platform for the e-humanities, used to collect and provide digital content, with the services needed to store the

content. One of the most important elements in search content is metadata, which is shown to be useful for finding data and algorithms.

Section 4 presents cases and experiences in the field of e-science. In the first paper, M. Heidari and O. Arnold show that fully digitalized scholarly activities such as online examinations can have a high variability, which presents a manageability challenge. The authors analyse the variability of legally analogue exam processes and prove the necessity for establishing management models. The authors of the second paper, *Designing External Knowledge Communication in a Research Network: The Case of Sustainable Land Management*, examine factors influencing the knowledge communication process. The aim is to find factors in successful communication between researchers and stakeholders as a representation of collaboration. The authors describe steps that need to be taken to enable successful communication: formulate the problem, analyse the situation, define communication objectives, identify target groups, formulate the message and develop a communication strategy and activities. S. Münster's paper, *Researching Scientific Structures Via Joint Authorships: The Case of Virtual 3D Modelling in Humanities* is the last in Section 4. This case study of scientific structures is an analysis of co-authoring for a defined set of conferences. The topics are interdisciplinarity, number of publications and co-authoring, and multipliers. The author identifies multipliers for knowledge in the field of 3D modelling.

Finally, in Section 5, A. Skulimowski presents a Delphi study trying to shed some light on future developments in e-science, especially in selected IT technologies. He focuses on two emerging systems, brain-computer interfaces and global expert systems that process databases, communication and unstructured formats like videos. These systems may lead to collective rather than collaborative research, as one researcher cannot manage the volume of information alone anymore. Another scenario based on the automated data analyses is that papers can be produced almost completely with minimal human intervention. In any case, Skulimowski paints an interesting picture of the future of science.

We hope that you will find this an interesting collection of a wide range of perspectives, which contributes to your ideas and visions of e-science.

Acknowledgements

First of all, the conference was part of the e-science Network of the Technische Universität Bergakademie Freiberg, Technische Universität Dresden and Leipzig University of Applied Sciences. This conference and the resulting publication have been enabled and financially supported by the European Social Fund ESF and the Saxon State Ministry of Science and Culture, whom we want to thank herewith. Additionally, we want to thank Dean Prof. Dr. Andreas Horsch for his financial support in order to make the book available as open access publication.

The editors especially want to thank all the authors whose contributions give this volume its special quality, and for their patient support throughout the process

of publication. Furthermore, we want to thank all reviewers for their helpful and progress enabling comments, enhancing the quality of all contributions. We want to thank Dominik Wuttke as well as Ilia Vershinin for their exact transfer of all the papers to LNCS. For the language correction, we want to thank Dr. Kate Sotejeff-Wilson for her support and quality assurance.

We wish you, the readers, inspiring reading!

Freiberg/Dresden, Germany
Spring 2020

Claudia Koschtial
Thomas Köhler
Carsten Felden

References

- Heinrich, L.J.: *Wirtschaftsinformatik*. Oldenbourg Verlag, München (1993)
- Köhler, T.: Research training for doctoral candidates in the field of education and technology. In: Drummer, J., Hakimov, G., Joldoshov, M., Köhler, T., Udarsteva, S. (eds.) *Vocational Teacher Education in Central Asia. Developing Skills and Facilitating Success*. Springer, Berlin (2018)
- Redecker, C., Punie, Y.: European framework for the digital competence of educators DigCompEdu. Publications Office of the European Union, Luxembourg (2017)
- Taylor, J.: e-Science. <http://www.e-science.clrc.ac.uk>, <https://web.archive.org/web/20010222231418/>, <http://www.e-science.clrc.ac.uk/> (1999). Last Accessed 18 Feb 2020
- Vagliano, I., Guenther, F., Heinz, M., Apaolaza, A., Bienia, I., Breitfuss, G., Blume, T., Collyda, C., Fessl, A., Gottfried, S., Hasitschka, P., Kellermann, J., Köhler, T., Maas, A., Mezaris, V., Saleh, A., Skulimowski, A.M.J., Thalmann, S., Vigo, M., Wertner, A., Wiese, M. & Scherp, A.: Open innovation in the big data era with the MOVING platform: An integrated working and training approach for data-savvy information professionals. *IEEE MultiMedia* **25**, 3, 8–21, July–Sept. 2018 (2018)

Contents

Understanding e-Science—What Is It About?	1
Claudia Koschtial	
Organising Academia Online	11
Thomas Köhler, Christoph Lattemann, and Jörg Neumann	
The Fish Model: When Do Researchers Collaborate Online?	29
Bahaeldin Mohamed and Thomas Köhler	
The Use of Digital Tools in Scholarly Activities. Empirical Findings on the State of Digitization of Science in Germany, Focusing on Saxony	49
Steffen Albrecht, Claudia Minet, Sabrina Herbst, Daniela Pscheida, and Thomas Köhler	
Digital Research Infrastructure	67
Maik Stührenberg, Oliver Schonefeld, and Andreas Witt	
MOVING: A User-Centric Platform for Online Literacy Training and Learning	77
Aitor Apaolaza, Tobias Backes, Sabine Barthold, Irina Bienia, Till Blume, Chrysa Collyda, Angela Fessl, Sebastian Gottfried, Paul Grunewald, Franziska Günther, Thomas Köhler, Robert Lorenz, Matthias Heinz, Sabrina Herbst, Vasileios Mezaris, Chifumi Nishioka, Alexandros Pournaras, Vedran Sabol, Ahmed Saleh, Ansgar Scherp, Ilija Simic, Andrzej M.J. Skulimowski, Iacopo Vagliano, Markel Vigo, Michael Wiese, and Tanja Zdolšek Draksler	
CLARIN-D: An IT-Based Research Infrastructure for the Humanities and Social Sciences	99
Gerhard Heyer and Volker Böhlke	
Toward Process Variability Management in Online Examination Process in German Universities: A State of the Art	111
Maryam Heidari and Oliver Arnold	

Designing External Knowledge Communication in a Research Network The Case of Sustainable Land Management	131
Thomas Köhler, Thomas Weith, Sabrina Herbst, and Nadin Gaasch	
Researching Scientific Structures via Joint Authorships—The Case of Virtual 3D Modelling in the Humanities	151
Sander Münster	
Visions of a Future Research Workplace Arising from Recent Foresight Exercises	169
Andrzej M. J. Skulimowski	

Understanding e-Science—What Is It About?



Claudia Koschtial

Abstract Our daily life has experienced significant changes in the Internet age. The emergence of e-science is regarded as a dramatic one for science. Wikis, blogs, virtual social networks, grid computing and open access are just a brief selection of related new technologies. In order to understand the changes, it is necessary to define these aspects of e-science precisely. Right now, no generally used term or common definition of e-science exists, which limits the understanding of the true potential of the concept. Based on a well-known approach to science in terms of three dimensions—human, task and technology—the author provides a framework for understanding the concept which enables a distinctive view of its development. The concept of e-science emerged in coherence with the technological development of web 2.0 and infrastructure and has reached maturity. This is impacting on the task and human dimensions as in this context, the letter “e” means more than just electronic.

Keywords e-Science · Open access · Grid computing · Science 2.0

1 Introduction

The “e” in combination with a number of well-known terms implies a transformation into online networks and the usage of information technologies, which has evolved in both private and professional life. Science, in its most general meaning as scholarship comprising all disciplines, has also been subject to this transformation. This development is being referred to as electronic/enhanced science, or e-science. The transformation may enable changes going beyond technology itself. According to Luskin, the big e means more than just electronic (Luskin 2012). Fausto et al. (2012) stated this more precisely: “Increasing public interest in science information in a digital and Science 2.0 era promotes a dramatically, rapid, and deep change in science itself”. The goal of this paper is to review research as work in progress.

C. Koschtial (✉)

Technische Universität Bergakademie Freiberg, Freiberg, Germany

e-mail: claudia.koschtial@web.de

The resulting literature analysis shows what and how science is changing due to the impact of using online networks and information technology.

The change in science can be traced back to the 1990s, when the concept of collaborative laboratories (collaboratories) evolved (Bly et al. 1997, p. 1). In 1996, the term cyberscience was sharpened by Nentwich (1999) who refers cyberscience to research activity which scientists were increasingly carrying out in the developing information and communication space. Taylor (1999) produced a definition close to this one: “e-science is about global collaboration in key areas of science, and the next generation of infrastructure that will enable it” and “e-science will change the dynamic of the way science is undertaken”. The definitions mark just the beginning of an ongoing transformation. Most recent aspects of e-science contain open access or science 2.0, referring to the usage of web 2.0 technologies like social networks, blogs or wikis. The cited definitions share some elements: activity of research, scientists, infrastructure, collaboration, information and communication. Nevertheless, a common definition does not yet exist, and more diverse terms have emerged since the first occurrence of this concept. Understanding the potential and extent of the change requires an analysis of the concept itself. The present research is an initial step towards this, which can be used as a basis for designing a comprehensive framework of the concept of e-science in order to support the work of scientists.

The remainder of the paper is as follows: the second section presents related work and the research gap. The third section explains how the research has been carried out and how the concept is going to be analysed in order to derive a definition. In Sect. 4, the results of the analyses are presented, leading to a discussion in Sect. 5.

2 Related Work

Science defines one possible way to make reality understandable. Leaving behind myth and religion, the ancient Greek philosophy represented an early systematic examination of the world. It dates from 2500 years ago, when the society transformed in the search for education and elucidation. Schools evolved, so science was (and still is) closely connected to teaching (Schülein and Reitze 2012, 31 p.)

Nowadays, there is no common perception or description of the change comprised by the term e-science (Yahyapour 2018, p. 369). The literature often deals with open access or particular problems related to data availability. Shneiderman (2012, p. 1349) stresses the potential for understanding and rethinking how a phenomenon is analysed. He promotes methodologies that move away from laboratory to real-world conditions, especially to analyse areas like “secure voting, global environmental protection, energy sustainability, and international development” (Shneiderman 2012, p. 1349). Eastman approaches the underlying process of e-science in terms of data analysis. He formulates an observational-inductive model in order to reflect on Knowledge Discovery in Databases and Data Sensor High-Performance Computing Models without a theoretical basis. His idea sounds promising, but he

provides few arguments for it (Eastman et al. 2005, 67 p.). Work and related organisational aspects of science like group learning and cooperative processes are addressed by Pennington (2011, 55 p.).

The mentioned literature is exemplary of a search in three literature databases (see Sect. 3.1). No general analysis of this area of discourse exists yet, so the usage and definitions of the terms have not been analysed before. Scientific understanding depends heavily on these papers, however. In order to sharpen the concept and identify discussed characteristics of e-science, the present authors performed the following literature analysis.

3 Research Approach

This section introduces the area of discourse and describes the applied methodology in Sect. 3.1. The applied research framework is then proposed in Sect. 3.2.

3.1 Research Field and Methodology

The research follows the method proposed by Fettke (2006, 257 p.). The research process itself demands that researchers have increasingly complex knowledge, which is usually beyond the borders of their own fields (Reinefeld 2005, p. 4). Two research challenges can be identified:

- The Internet can be used to search for and communicate information, but success in identifying information is not guaranteed.
- The vast amount of data is challenging to process in order to identify relevant content.

The mentioned challenges appear as well for the field of e-science. A couple of terms being used in e-science comprise some or all the elements mentioned above. The ones which have been mentioned so far are:

- e-science itself meaning electronic or digitally enhanced science (Hiller 2005, p.5);
- cyber infrastructure (Hey 2006);
- e-research (University of Technology Sydney 2013);
- cyberscience (Atkins 2005, 1 p.); and
- science 2.0 (Leibnitz 2012).

As these terms appear at different points in time, the meaning has to be reflected on and trends need to be considered in order to understand the circumstances in which they arose. Relevant literature was identified by searching the title, abstract and keywords for the terms “e-science”, “eScience”, “e-research”, “eResearch”, “science 2.0”, “cyberscience”, “cyberinfrastructure”, “grid computing” and “grid”

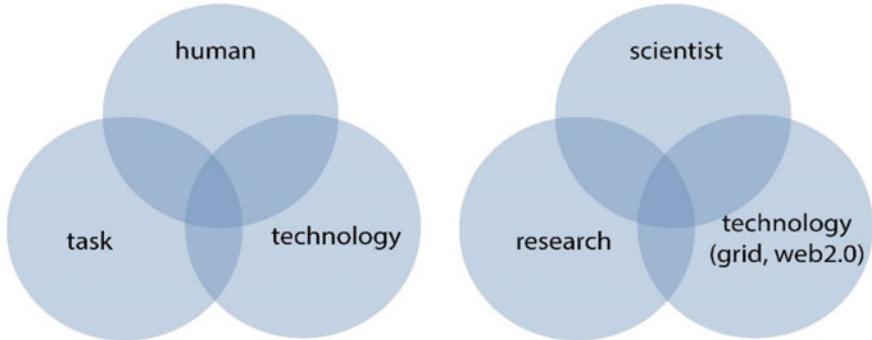


Fig. 1 Heirich's human—task—technology framework (Heinrich 1993, p. 8) and its adaption to the field of e-science

together with “e-science” in three databases: EBSCO Academic Search, ACM Digital Library and IEEE XPlore. To increase the amount of results, Google Scholar was also searched for titles in the period from 1994 to 2005. Digital humanities were excluded as it refers solely to e-science in the field of humanities.

3.2 *Research Framework*

A research framework is needed in order to identify the essence of the concept of e-science and differences between the terms being used.

Science 2.0 includes a range of topics. Shneiderman (2012, p. 1349) identified research on sociotechnical systems as the basis for an increasing collaboration. Heinrich (1993, p. 8) regards sociotechnical information systems as composed of human, task and technical dimensions; he sees such systems as open, complex and sophisticated. Figure 1 shows the general framework created by Heinrich (left-hand side) and its adaption to the context of e-science (right-hand side).

Regarding the given definitions, some initial characteristics can be extracted: scientists, information and communication, infrastructure, collaboration and research. In order to reflect all aspects of e-science, collaboration is added to the framework, as this was inherent in all definitions. Figure 2 shows the framework used.

4 Results

The literature search led to 148 definitions of the selected terms related to e-science. The most frequent definition was “e-science” (43%), followed by “grid”

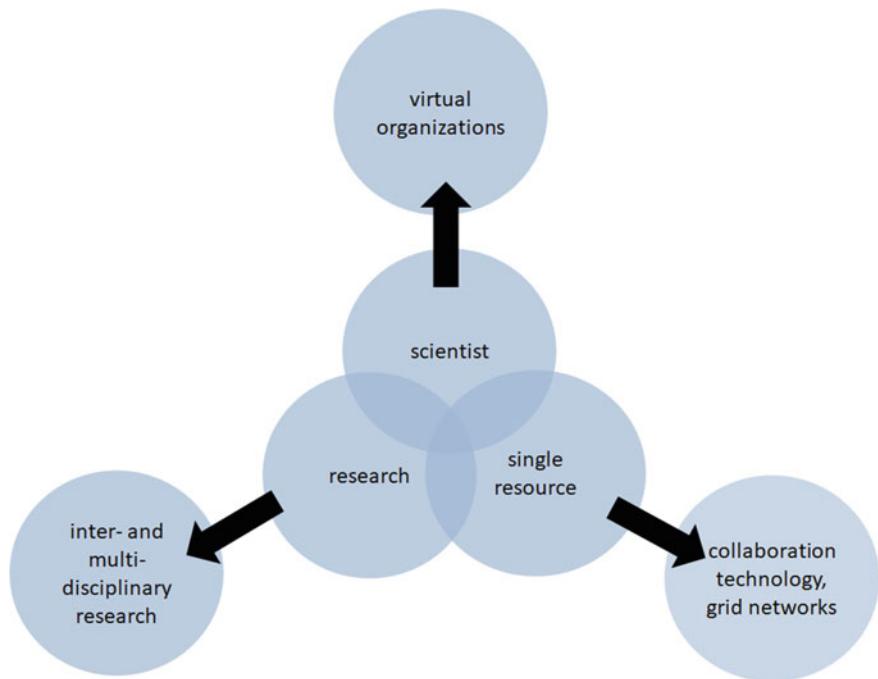


Fig. 2 E-science framework including collaboration

(32%), “science 2.0” (9%), “cyberinfrastructure” (8%), “e-research” (7%) and “cyberscience” (3%). Table 1 shows the number of definitions per year.

Figure 3 shows the occurrence of these terms over time.

In a second step, the authors analysed the development of the selected definitions over time and investigated whether the dimensions of the framework were mentioned in each definition. The following examples show key terms related to each dimension.

- Technical dimension:
 - Web 2.0 technologies as a single technology;
 - Networks and infrastructure as a collaboration technology.
- Task dimension:
 - Publishing, analysing or teaching as single tasks;
 - Collaborative projects which may have an interdisciplinary focus.

Table 1 Number of definitions per year

1998–2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
5	5	5	16	17	15	14	12	10	17	11	11	10

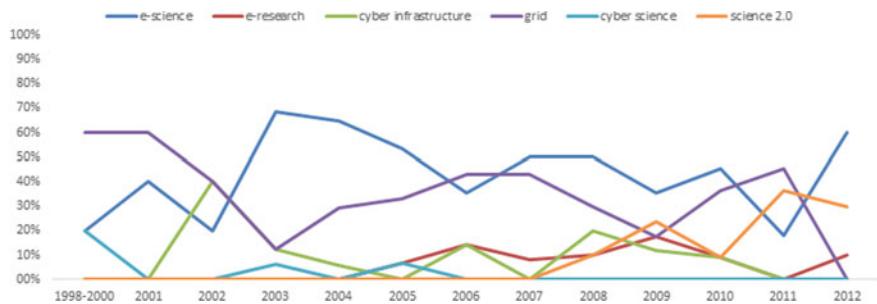


Fig. 3 Relative frequency of terms related to time

- Human dimension:
 - Researcher as human;
 - Virtual organisations like social networks.

The next step was to analyse the relations between the three dimensions, human, task and technology.

5 Discussion of Initial Results

Figure 3 shows that terms like cyberscience or cyberinfrastructure disappeared over time. The presence of the term e-science is relatively stable over the time, which can be seen as acceptance and establishment of this term. The frequency of the term grid is decreasing, which may hint that the technological side of the concept is already mature, established and needs no further development but that claim needs to be checked for the next years. Additionally, the funding period of the UK e-Science Core Programme stopped in 2006, resulting in a reduction of interest in the topic or at least resulting in a reduced amount of publications.

Figure 4 shows the content analysis of the definitions. The human dimension has an approximately stable occurrence over time. But technology is less often mentioned throughout the analysed period. Regarding technology, the number of definitions describing collaborative technology as a constitutive characteristic decreases over time. The term grid is also used less and less over time. Technology seems to be no longer a challenge, but an enabler. The single resource referring to web 2.0 technologies is stable over time. In the task dimension, collaborative/interdisciplinary research projects do not play a significant role. The intention of financial supporting institutions to encourage collaborative research may play an increasing role—but such a trend is not visible, yet. Research as task is an increasing part of the definitions, which might be a further hint that the technology itself is mature and the usage is becoming more important. This allows the concept to be used in more different fields.

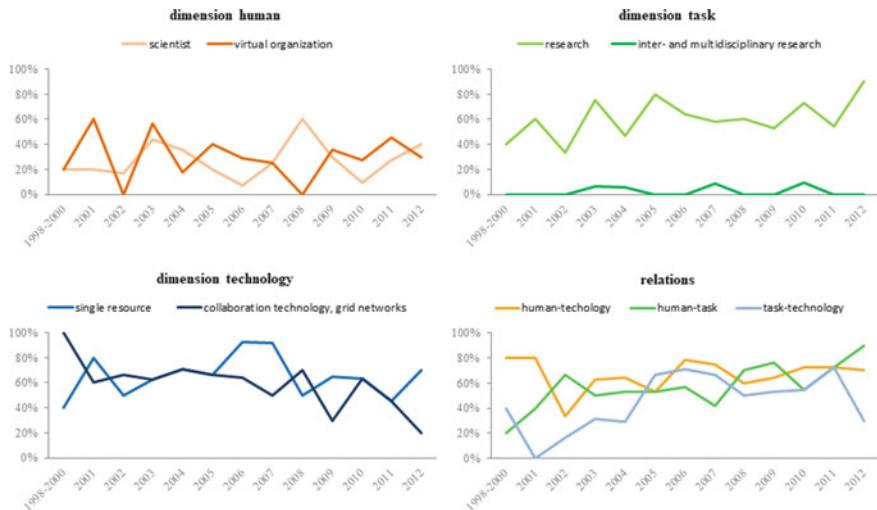


Fig. 4 Results of the analysis of the human, task and technology dimensions of e-science

Regarding the relations between the dimensions, an important link is emerging between task and technology. This may be understood as an indicator for increasing automation. Furthermore, the relation between human and task is the relation that is increasing most sharply.

The use of the selected terms varied by geographical location and in relation to public funding programmes in the respective area. The term e-science itself has been used by the UK e-science Core Programme from 1999 until 2006. Cyberinfrastructure comes from the USA, and e-infrastructure emerged in Europe. A further term appeared in 2005 on an initiative of the Australian Research Councils, which was entitled e-research. The focus here however is not on geographical differences and funding; this issue requires further investigation.

6 Conclusion

The aim of this paper was to show how the use of the term e-science is changing through a literature analysis. The initial results show that the concept of e-science changes over time. One aspect of the concept is technology, referring to infrastructure and single resources:

- Grid computing is “an important new field, distinguished from conventional distributed computing by its focus on large-scale resource sharing, innovative applications, and, in some cases, high-performance orientation” (Foster et al. 2001, p. 200).

- Web 2.0 technologies are an evolutionary stage in Internet use. Examples are virtual communities, blogs or wikis (Nentwich 2009).

Furthermore, e-science is oriented to tasks: processing vast amounts of data, searching for information or publishing content. The task of establishing collaborative projects is weakly represented in the analysed literature.

- Open access refers to “The first is a change in the publishing model to one more suited to the age of the Web; the second, a change in how scientists connect with society – their major funders through taxation” (e-science talk 2012).

Additionally, the scientist plays an important role in the concept of e-science in two ways:

- as a single researcher;
- as virtual communities, which exist only in the Internet. They form for a limited period in time as interdisciplinary groups of regional segregated elements (Mosch 2005). The key characteristic of such units is collaboration.

The changes related to e-science are apparent in all three of Heinrich's dimensions. Important concepts like open access or the grid have been attributed to the different dimensions. Therefore, the potential of e-science is not reduced to electronification, but expanded to include redesign of tasks, the emergence of virtual organisations and the rapidly increasing importance of collaboration. Right now, the technology dimension still dominates the concept, but it is maturing and this will form the basis for further changes.

It seems necessary to do further research to analyse related technologies and tasks behind the concept of e-science in more detail in order to provide a sufficient base for scientists to be able to learn about the potentials of e-science and to convert those potentials into realised benefits.

References

- Atkins, D.E.: Cyberinfrastructure and the next wave of collaboration. http://hydra-cog.fsl.noaa.gov/site_media/docs/atkins_2005_wave.pdf (2005). Accessed 24 Feb 2013
- Bly, S., Keith, K.M., Henline, P.A.: The work of scientists and the building of laboratories. <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.40.2678&rep=rep1&type=pdf> (1997). Accessed 30 Mar 2020
- Eastman, T.E., Borne, K.D., Green, J.L., Grayzeck, E.J., McGuire, R.E., Sawyer, D.M.: eScience and archiving for space science. <http://doi.org/10.2481/dsj.4.67> (2005). Accessed 29 Mar 2020
- e-Science talk: <http://www.e-scientetalk.org/> (2012). Accessed 30 Nov 2013
- Fausto, S., Machado, F.A., Fernando, L.: Research Blogging: Indexing and Registering the Change in Science 2.0. <https://doi.org/10.1371/journal.pone.0050109> (2012). Accessed 28 Mar 2020
- Fettke, P.: State-of-the-Art des State-of-the-Art. Eine Untersuchung der Forschungsmethode „Review“ innerhalb der Wirtschaftsinformatik. *Wirtschaftsinformatik* **48**, 257–266 (2006)
- Foster, I., Kesselman, C., Tuecke, S.: The anatomy of the grid. <https://doi.org/10.1177/109434200101500302> (2001). Accessed 29 Mar 2020
- Heinrich, L.J.: *Wirtschaftsinformatik*. Oldenbourg Verlag, München (1993)

- Hey, T.: A conversation with Microsoft's Tony Hey. <http://poynder.blogspot.de/2006/12/conversation-with-microsofts-tony-hey.html> (2006). Accessed 12 Jan 2013
- Hiller, W.: Aufgaben von D-Grid. *Wissenschaftsmanagement* **1**, 6–7 (2005)
- Leibnitz Forschungsverbund: Science 2.0. <http://wiki.zbw.eu/display/sci20/Forschungsverbund+Science+2.0> (2012). Accessed 6 Feb 2013
- Luskin, B.: Think “exciting”: e-learning and the big “E”. <http://www.educause.edu/ero/article/think-exciting-e-learning-and-big-e> (2012). Accessed 2 Dec 2013
- Mosch, K.: E-Science: Managementfragen. *Wissenschaftsmanagement* **1**, 2–3 (2005)
- Nentwich, M.: Cyberscience: Die Zukunft der Wissenschaft im Zeitalter der Informations- und Kommunikationstechnologien. <http://www.mpifg.de/pu/workpap/wp99-6/wp99-6.html> (1999). Accessed 30 Mar 2020
- Nentwich, M.: Cyberscience 2.0 oder 1.2? Das Web 2.0 und die Wissenschaft. http://epub.oeaw.ac.at/0xc1aa500d_0x0022a615.pdf (2009). Accessed 12 July 2013
- Pennington, D.D.: Collaborative, cross-disciplinary learning and co-emergent innovation in eScience teams. <https://doi.org/10.1007/s12145-011-0077-4> (2011). Accessed 29 Mar 2020
- Reinefeld, A.: Chancen eines internationalen Wissenschaftsverbundes. *Wissenschaftsmanagement* **1**, 4–5 (2005)
- Schülein, J.A., Reitze, S.: *Wissenschaftstheorie für Einsteiger*. 3. eds. UTB GmbH, Stuttgart (2012)
- Shneiderman, B.: Science 2.0. *Science* **319**(3), 1349–1350 (2012)
- Taylor, J.: e-Science; <http://www.e-science.clrc.ac.uk>; online retrieved from <https://web.archive.org/web/20010222231418/http://www.e-science.clrc.ac.uk/> (1999). Accessed 18 Feb 2020
- University of Technology Sydney: eResearch. <http://www.research.uts.edu.au/eresearch/index.html> (2013). Accessed 7 Dec 2013
- Yahyapour, R.: E-Science Infrastrukturen. *Informatik Spektrum (Informatik-Spektrum)* **41**, 369–375 (2018)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Organising Academia Online



Thomas Köhler, Christoph Lattemann, and Jörg Neumann

Abstract Research on organisational arrangements of scholarly networks in both e-learning and e-research is located at the intersection of different theoretical justifications and developmental contexts such as organisational theory, computer science, education science and media informatics. However, there is still a lack of research on the organisational context of e-learning arrangements and its impact on collaboration in academic communities. E-learning research shows that the integration of electronic media in scientific communities negatively impacts their effectiveness and causes conflicts within communities. Research networks however are far less investigated as there is not direct didactic focus on how to collaborate. Recent theories on organisational design, virtual organisations and governance provide concepts for organising e-collaboration more effectively. Managerial instruments such as direct control of results and behaviours need to be supplemented or even replaced by concepts of social control; typically trust and confidence become the central mechanisms for the new forms of inter- and intra-organisational coordination. This paper starts with concepts. Then, to exemplify the organisational coordination mechanisms in scholarly e-communities, the authors critically discuss and reflect on these organisational arrangements and managerial concepts for two higher education portals and one research network in Germany. The conclusion is that, just as previous research has confirmed for educational networks, governance within academic networks relies heavily on the functionality of social and communicative forms of control.

Keywords Research network · Education portal · Virtual organisation · Governance · Social control · Science collaboration · Scholarly collaboration · Online community

T. Köhler (✉)

Department of Education, Technische Universität Dresden, Dresden, Germany

e-mail: thomas.koehler@tu-dresden.de

C. Lattemann

Department of Business and Economics, Jacobs University, Bremen, Germany

e-mail: c.lattemann@jacobs-university.de

J. Neumann

Media Center, Technische Universität Dresden, Dresden, Germany

e-mail: joerg.neumann@tu-dresden.de

1 Introduction

The central aim of this article is to identify forms of organisational governance (social, output, or behavioural control) that are suitable for effective e-collaboration in scientific communities. Are “e-learning” and “e-science” fundamentally different things? Specifically, does e-learning concern teaching, and e-science, research? This is factually correct, but from an organisation theory perspective, not a sufficient criterion for differentiation. Above all, the clientele at issue here is the same: the teaching and research staff of universities. In addition, both activities are carried out within the same institution. In this respect, comparison is not only possible, it is mandatory.

Our evaluation is based on both a review of the relevant literature and empirical studies, some of which were conducted by the authors. Following the classification of virtual organisations, the main characteristics of organising academic activities are presented and validated through suitable institutional examples.

2 E-Learning Organisation: Media Integration as Organisational Development

2.1 *Online Technologies in Higher Education*

The integration of new media in educational settings has been intensively discussed in academic research and education for about 15 years. Various forms of online, distance, and blended learning have been implemented and tested. After a series of tentative, rather experimental tests to integrate new Internet technologies and electronic media in teaching processes, the management of students and eventually the teaching itself, we now see the results in the forms of web-based tutorials (WBT), virtual learning environments (VLE) and more recently in massive open online courses (MOOC).

With respect to developments in the online learning arena, in 1999 the German expert group on Higher Education Development by New Media predicted the higher education landscape would be as follows (cf. Köhler et al. 2010):

1. Global education providers and platforms offer worldwide accessible online courses.
2. Traditional universities are in competition with private online providers, in particular with corporate universities, and students use the opportunities of the global education market.
3. In order to survive in this competitive situation, many colleges have joined together in networks and offer common learning opportunities, while universities are jointly offering their academic programs together under the umbrella of a virtual university.

4. Student services are provided by facilitators and tutors, and less by classical university teachers, because more than 50% of students study online.

As of today, these predictions can only partly be confirmed. However, besides the established Open Universities like the British Open University or the German Fernuniversität Hagen, new global education providers such as the edX, Coursera or Udacity are emerging and become more relevant with the increasing need for lifelong learning and with growing numbers of students seeking for flexible online learning. Nevertheless, they still only play a niche role in higher education so far. But it is no surprise that the Centre for Higher Education Development (Hener and Buch 2006) concluded more than a decade ago that “[i]n academic education [...] uses of digital media in teaching and learning and integration of information technology-based administrative services have become widely established. Key questions of the future are seen especially in the interlinking of different services” (p. 2).

2.2 Virtualisation in Higher Education

Academic research has dealt with the use of Internet-based technology in teaching for many years (see, e.g., Lievrouw et al. 2000; Issing and Klimsa 2003, 2010). While initial claims were rather didactic (“classroom technology”), virtualised educational scenarios (VLEs, MOOCs, etc.) are of increasing interest nowadays. The concept of virtualisation is being used more and more often to describe the essential features and expectations of information and communication technologies (ICT) and multimedia, and to document the change. What exactly is behind it? Features of virtualisation described by Köhler et al. (2010) include the facts that students no longer meet their seminar leaders personally and that neither they nor the lecturers need to borrow books from the library. Researchers submit their conference abstracts, and expert opinions on other posts, via an Internet portal, while heads of research projects identify potential research partners in a database—without having ever met in person before. All in all, universities and virtual academies cooperate by uploading teaching content to a joint learning management system to be used by students from other institutions. In sum, such a far-reaching change in the educational landscape has established itself in less than 15 years and is on the verge of becoming the standard. However, acceptance by the teaching staff, especially at universities, is rather low; for example, professors in Toronto went on strike in 1997 and have managed to keep their teaching offline until today. Similarly, a study published by the Centro Nacional de Estadística, Geografía e Informática Mexico in 2004 (INEGI 2004) explained that 70% of professors in Mexico protested against the use of ICT in education. Their main reason was and perhaps still is the form of presentation of course content when using ICT in formats like PowerPoint and LaTeX. The distinctly reluctant behaviour of university staff is illustrated, for example, by the words of a professor from education sciences “you have to operate well didactically [...] and a part of this is the whole computer nonsense” (Misoch and Köhler 2005, p. 1). In the same

way, the dean of the engineering department at a leading German university stated in 2015 that “the nightmare is graduates who no longer draw without a computer, no more writing”.¹ The prevailing opinion is that this leads to a very impersonal design of seminar rooms and lecture halls, whereby students may lose their communication and personal contact with each other. Respondents continue to believe ICT should only be used in education to communicate data and not to communicate between people, nor do they see it as a new academic format or alternative for formation, though it may be used in addition to a classroom setting.

Hence, pivotal questions remain unanswered. What will the campus of 2025 look like? Which organisational models of e-learning and e-science collaboration will prevail? Despite the aforementioned reluctance in academia, other developments are observable. For example, online learning is proliferating in media-related disciplines; topics such as artificial intelligence, telemedicine and distance learning, MOOCs and open science are frequently and extensively discussed as powerful new opportunities for improving academic activity in general (Pscheida et al. 2014; Lattemann and Khaddage 2013).

Our first conclusion is that ICT has changed (academic) education. As the above examples illustrate, this change is not limited to education, academic teaching and learning. This raises the question of what exactly the virtualisation of education means. As early as 1999, Landfried, then President of the German Rectors' Conference, described unlimited access to stocks of knowledge independent of time and space; yet this knowledge is disconnected (separated) from physical institutions and, in particular, individuals (Landfried 2009). What is meant by this double separation? To answer this, it is important to analyse what is virtualised, which is more than the learning objects or knowledge content. In fact, relations (micro- and macro-social, but also those between learners and learning object) can be virtualised as well as knowledge, sometimes both at the same time.

3 Change of Organisational Theories and Paradigms

What has been known from both management and operational practice for a long time (cf. Frindte et al. 2000) now also appears to apply to education: ICT is becoming more important in managing organisational processes, and these infrastructures are becoming permanent. But these processes vary significantly, raising the question of the ideal configuration of technology and organisation. The first research to address this issue introduced new ICT to control operational processes in knowledge cooperation. Munkvold (2003) set up such a heuristic that can be transferred to the educational context almost directly. He divided the “implementation of collaboration technologies” into four sub-areas, the (1) organisational context, (2) implementation project,

¹This quotation was taken from an anonymised interview by the author.

(3) technological context and (4) implementation phase. Similarly, with explicit reference to the introduction of online learning in higher education when used as dimensions of change, Euler et al. (2004) proposed the following five dimensions: (1) economic dimension, (2) pedagogical-didactic (educational) dimension, (3) organisational/administrative dimension, (4) technical dimension and (5) sociocultural dimension.

Are these theories based on economics or technology? Neither. Organisation and organisational culture are central to change. With this assessment, the authors align with a strand in the German educational research tradition (Neumann and Schütte 2008) that is gaining ground but still rather new. This broadens the academic perspective on the use of media, which was previously dominated primarily by cognitive (psychology), teaching (pedagogy), education-oriented (educational science) or even technological (computer science, etc.) approaches. An organisational perspective adds a social and management science-based momentum, and macro-social perspectives. After 2005 more research programmes in Germany sought to meet the need for such an approach, including New Media in Education II or the later Digitisation Initiative (2014). In education and media studies, where approaches based on organisation studies, education science, or media economics are preferred, researchers are frequently challenged to take these approaches.

Just after 2010, based on the concept of openness—used when coining the terms of OER and MOOC—many became convinced that the technology used for university operations would be revolutionised. Within the next decade, it is expected that students will no longer attend lectures or work in a lab, but will join professors' research activities online, whenever and wherever they want. Academic knowledge will be tailored, or transferred from mass production to mass customisation. So what is the core of the “digitisation of teaching” or the “advent of information and communication technologies in the university”? Germany's former Minister of Science, Bulmahn (2004, p. 5), argued that “the new media in the combination of computer and Internet [will penetrate] all social and economic sectors [and will release] a fundamental structural change” combined with unprecedented speed of market globalisation. Ortner and Nickolmann (1999) stressed that the success of open universities will force conventional universities to adopt innovations in teaching organisation, such as distance learning, on-campus students as independent learners, modular course structures and the enrolment of mature part-time students. This goes along with changing forms of social micro-study, from online learning communities (Kahnwald and Köhler 2005) to more complex flexible online knowledge organisations (Köhler et al. 2003).

To speed up the new media restructuring of higher education, the Federal Ministry of Education and Research (BMBF) has targeted the existing New Media in Education Programme and the 2004 re-bid. The first phase of the programme from 2000 to 2004 aimed to develop high-quality e-learning content and concepts for mobile learning, and to put them into regular practice, particularly in undergraduate studies. These developments were intended to be available from 2005 and to be sustained and broadened by two conveyor lines. Conveyor line (A) was for projects in an interdisciplinary and university-specific context, called “e-learning integration”. This is

about developing organisational infrastructure and about changing management to develop utilisation of the opportunities provided by ICT innovation potential in the field of teaching, learning, and exams to universities systematically and sustainably. Conveyor line (B), for projects in a university-wide and primarily subject-specific context, referred to as “e-learning transfer”, was to lead to new organisational concepts and business models for services, related to the production and use of online learning primarily supporting professional and technical areas (cf. BMBF, 2004, all translations from German by the authors). By 2010, most of these projects were completed. What impact did the targeted re-organisation of online learning in German universities have?

3.1 The Research Framework: Virtual (Educational) Organisations

In view of the different organisational theories applicable to online teaching and learning in a university context, including its structural and procedural commonalities, the following issues should be noted. At the institutional level, online learning is integrated into the organisational structure of the university. This requires sufficient integration of external service providers. Figure 1 presents the value chain of e-learning from a university perspective, including the internal and external partners at the Technische Universität Dresden in 2008.

The e-learning value chain shows that teaching and learning in an electronically mediated environment is multifaceted and involves various stakeholders. Because of the various partners involved, the organisational concept shows many characteristics of a virtual organisation with loosely coupled partners (external content providers, platform providers, external and internal instructors and students, etc.). Hence, universities which provide online learning arrangements must also follow, or at least adopt, mechanisms of virtual organisations. They must change their structures from their traditional departmental separation towards more process-oriented, open and collaborative organisational settings.

These kinds of new virtual organisations are primarily shaped by their virtual character and are limited by their lack of “real” organisational boundaries. This applies to all organisational aspects: the location, bonds and stability of the organisation. Such a virtual organisation is “multisite, multi-organisational and dynamic” (Snow et al. 1999).

As shown by Köhler and Schilde (2003), virtual organisations can differ greatly in terms of size, durability or stability. Furthermore, various forms of virtual organisation and cooperation are described in theory and can be observed in practice, under an equally large number of names (network, cluster, virtual team, virtual organisation, etc.). In order to make these phenomena comparable and assign experimental findings, a further differentiation of the term is required. Okkonen (2002) proposed one

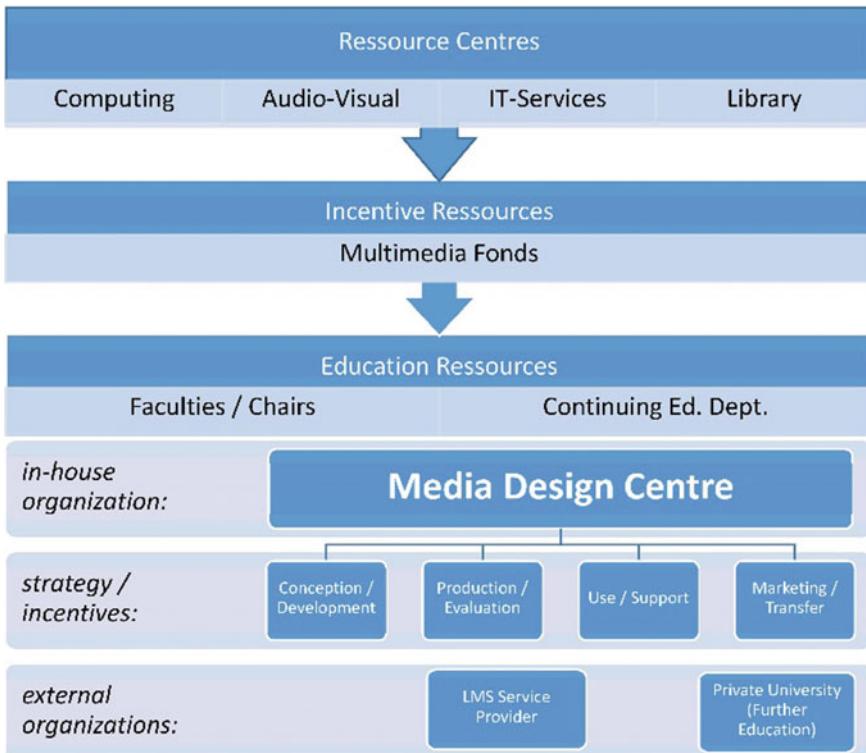


Fig. 1 Organisational framework of online learning using the example of the Technische Universität Dresden (own figure after Neumann and Schütte 2008)

way of doing this, presented by Köhler et al. (2003) as an advanced systematisation of virtualised organisational forms (see the following Table 1).

In the following, two case studies on online learning and one case study on online research are presented and critically discussed from the perspective of virtual organisations.

3.2 Research Methods

This paper follows an inductive research approach in order to identify relevant organisational mechanisms in an e-learning institution, based on three case studies. The case study method is selected as it is a common and comprehensive investigative tool for exploring individual, group, organisational or social phenomena (Yin 2013; Bryman and Bell 2011). In this instance, the weaknesses in corporate data security are investigated, in order to reveal potential causes, as discussed in the analysis section.

Table 1 Differentiated characteristics of virtualised organisational forms (own figure after Okkonen 2002; Köhler et al. 2003)

	Virtual Team	Virtual Project	Virtual Organisation (temporary)	Virtual Organisation (permanent)	Meta Network
Involvement	organisational element	function + overarching	overarching	overarching	overarching
Membership	small, local	medium sized	usually large	small, but scalable	large, vague
Goal	teams for specific tasks	representatives of >1 organisation on projects	multiple functions as reaction to market	all functions + full functionality of working organisations	multiple functions in reaction to market
Duration	membership varies / form permanent	temporary	temporary	permanent	permanent
Information Technology	Connectivity, knowledge remains separated	collection of shared data	shared infrastructure	replaces physical infrastructure	links the networks
Example from Education	Online learner team	Student group for term paper	R&D Project	Education Portal	ResearchGate

We have chosen two case studies because the authors of this paper are involved in the projects and they have deep insights. A triangulation approach was utilised as this is “the most desired pattern for dealing with case study data” (Yin 2011). Seminal articles on the case study topics were selected for analysis (Yin 2013).

For this particular example, differing sources have been consolidated to present a comprehensive case study summary, including scientific publications, research reports, and public descriptions on the websites of the chosen institutions. All material was either available publicly or from internal sources. Figures used come from self-descriptions of those projects—the layout was not changed, but translated.

Case I: Online learning in academic education through the education portal of Saxony (since 2001)

Since 2001, a university network has been supporting online teaching at public universities in the German federal state of Saxony. After an initial phase with the direct participation of the four universities which comprised this group since 2004, a system corporation, BPS Education Sachsen GmbH, was founded in 2006. In an evaluation of the state of development of online learning at Saxon universities for the Saxon Minister of Science and Art, the German National Centre for Higher Education Development (CHE), stated in 2006 that despite many years of funding by means of the country and the special commitment of many scientists concluded that online media is still used on a relatively small scale. Overall, however, acceptance is increasing among both university staff and students. But Hener and Buch (2006) noted a lack of liability for student usage, sustainability in higher education, and overall management of e-learning in higher education. This has been confirmed by further analyses (Köhler and Ihbe 2006) calling for a more systematic integration of online learning at Germany’s largest technical university, the Technische Universität Dresden. In 2007, control of the project passed to the newly established e-learning



Fig. 2 Model of the education portal of Saxony (cf. <https://bildungsportal.sachsen.de/>)

working group of the Rector's Conference Saxony. Since then, all public universities in Saxony and two private universities have joined the network. The following Fig. 2 shows the distribution of the educational portal in Saxony as of 2008:

Case II: Online-supported continuous learning in the education portal of Thuringian universities (2000–2013)

Based on analysis of the need for media-based academic training and organisational structures at and between the universities of Thuringia, and to support more sustainable development of such online training, the (online) education portal for Thuringia was constructed in 2001 (www.bildungsportal-thueringen.de). As a consequence of the above tests, this portal aimed to serve institutional training seekers or their staff, that is, employees who want to selectively add to their skills profile according to their academic or equivalent qualifications or needs. There was already significant potential demand for this when the portal opened. An expert (Stifterverband 2001) estimated that 20,000 of almost 60,000 students of the Distance University of Hagen alone are undergoing a hidden continuing professional development (CPD). The education portal of Thuringia competed with several private CPD providers. This fact should be mentioned because the expectations and attributions of training seekers were influenced by their experiences with these market leaders. Nevertheless, the participating universities have reconfigured themselves on the virtual organisation model, consisting of a core information broker and a network of partners meeting training needs, as in Fig 3.

The education portal of the Thuringian universities remained at the project stage until 2013 and was then closed by the responsible Ministry of Science.

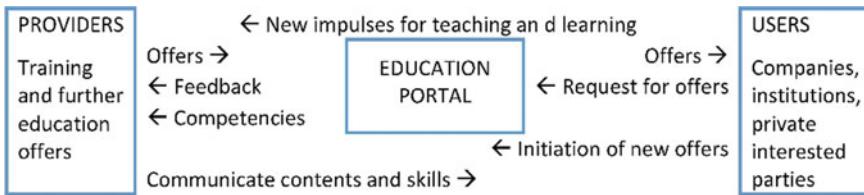


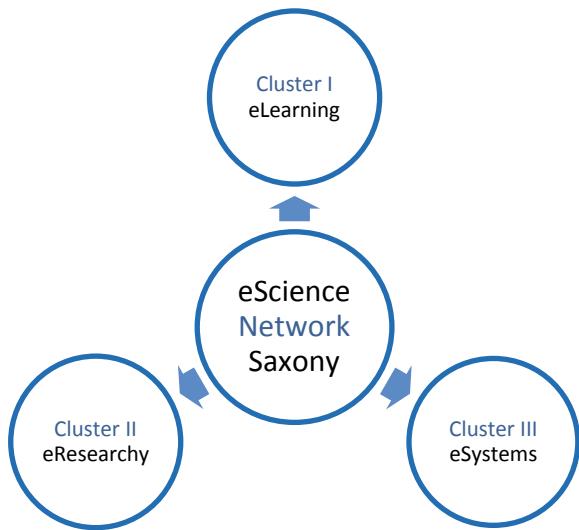
Fig. 3 Model of the education portal of Thuringia (own figure after Schmidt 2002)

Case III: The e-Science Saxony Research Network as a virtual science organisation (since 2011)

The e-Science Research Network project is a Saxony-wide comprehensive research network of all state universities created to explore approaches and methods in e-science (electronic science). The term e-science describes the different fields of scientific research and development related to the use of computer technologies. While this term is mainly used in Germany and the UK, comparable concepts include “cyber-infrastructure” in the United States or “e-research” in Australia. Currently, the slogan “Science 2.0” frames the discussion, in particular concerning cooperative digital scientific work (Weichselgartner 2010). The thematic range of infrastructures, application architectures, grid and cloud technologies extends to the educational technology known as e-learning. In addition, e-science systems support cooperative research between universities and with the private sector (cf. Ziegler and Diehl 2009). Research in e-science can be subdivided into disciplines such as e-humanities, e-medicine or e-engineering. In any case, it extends the scholarly process by integrating e-technologies and methods based thereon. The methodology was found to screen collaborative research activity, but knowledge organisation changed also dramatically and has been systematically underdeveloped by these e-disciplines. Even when research contexts are established or reused, it creates new paradigms, such as the concept of a “living lab”. This is user-centred research and open innovation practice, based on research work in multidisciplinary teams. One of the essential activities of these teams is co-creation, bringing together technological innovations and their applications through procedures such as crowdsourcing and crowdcasting. In these driven-by-research community practices, a variety of opinions, needs and knowledge exchanges can be used to brainstorm new scenarios, solutions and applications; yet these may be one-sided (Fig. 4).

Overall, starting with a steady drop in the “half-life of knowledge”, the changing demands of industry and the economy, and social changes in the knowledge society, the network partners have developed a new type of research and the accompanying scientific activities. New information and communication technologies can be used in this context, especially to provide, disseminate and use research information, such as laboratory data from simulations using complex aggregate social science information. Thus, media-based networking researchers are characterised by a high degree of flexibility and variability; usage may translate into new contexts through the restructuring of data and their usage. Through the coordinated action of the Saxon

Fig. 4 Clusters and organisational structure in the e-science Saxony research network



State Ministry for Science and Art and the Federal Republic of Germany, the Saxon universities have achieved an excellent level of “computational science”, especially in introducing e-learning support systems (Hener and Buch 2006). Summarised as e-sciences, the current project focusses on e-business, e-learning and e-systems, which are interwoven holistically at universities in the context of teaching and research.

4 Discussion and Conclusions

4.1 Theoretical Considerations About the Functioning of Virtual Organisations in the Academic Sector

Recent digitisation initiatives in academia demonstrate the pressing need of a serious discourse about its fundamental principles and practical meaning for the whole sector. In Germany since its launch in 2014, the Higher Education Forum on Digitisation has created an independent national platform to discuss the multiple facets of digitisation in higher education by consulting in six thematic groups on issues surrounding the digitisation of university teaching.²

Two decades ago, Malone and Davidow (1992) triggered the discussion about new organisation and management concepts in the economic sciences with their path-setting contribution “Virtual Corporation”. Until that moment, organisational change was marked by various headings such as “Computational Organisation”, “Learning

²<http://www.hochschulforumdigitalisierung.de/>, retrieved on 15 July 2015.

Organisation”, “Organisational Communication”, “Society and Internet Development”, “Trust Leadership and Decision Making” or “Augmented Reality” (cf. Köhler and Schilde 2003). All approaches share a similar basis: organisational units are reduced to their core competencies and have to cooperate in network-like structures. Complex tasks are realised by a number of independent organisational units or enterprises with complementary skills. This calls into question traditional organisational concepts, as published in governance research. Direct output and behaviour control, which are feasible in traditionally structured enterprises with divisional and functional organisation patterns, are supplemented or even replaced by concepts of social control. In the 1980s, psychological studies of cooperation and communication in virtual communities depicted computer-mediated communication as typically rather anomic in nature (Sproull and Kiesler 1986), less tolerant (Funkhouser and Shaw 1990) and lacking transferable behaviour (Köhler 2003). Postmes (1997) see this analysis as based on the less medium-socialised population of the “early years”. Therefore, these findings would be difficult to replicate. However, the cases presented here show that today’s changed environment creates completely new ways of medium-socialised collaboration. Once again, the majority are beginners in a new (mediated) organisational culture. Consequently, Lattemann and Köhler (2005) assumed that trust and security of contract would become key factors of cooperation in virtual organisations. This implies that social control becomes a strategic factor in competition among virtual organisations (Barney and Hansen 1994; Krysteck 1997) laying the foundation for new forms of cooperation. Their analysis based on literature review, and our own empirical studies, lead us to observe that the less output and behaviour can be assigned directly to specific individuals, the more important social control of the community becomes.

Our three case studies demonstrate that organisational development towards a networking, virtualised organisational structure can be found in both the academic education and research domains. For both domains, it is obvious that this development is going beyond existing organisational patterns; however, it is not necessarily sustainable, as the closure of the education portal of Thuringian universities after only ten years shows. Is this development merely the interface of a larger organisational change, or the beginning of a new era?

Networking organisations need to move beyond the purely project stage. In all cases, besides new organisational forms we found both close linkage to existing units, including several management instances like steering committees, information offices, and supervisory boards. Neither a classical hierarchy nor a clear linkage to all partners were found in these cases. Structures and opportunities for influencing the processes seem rather soft and depend on functioning communication.

In sum, virtual networks with flexibly aligned partners, who deliver different services and competencies, heavily rely on the coordination of and motivation for social control and trust. Appropriate instruments need to be strengthened. Long-established norms cannot be adopted because these are either insufficiently developed or simply not applicable—which led to the central question studied by the authors previously: Which governance concept is most efficient in the diverse forms of a virtual organisation? In their study, Lattemann and Köhler (2005) examined the extent

to which new governance concepts (i.e. social control) may be applied to forms of e-learning (i.e. virtual collaboration) and could propose a classification system for virtual organisations. Already before and after Köhler et al. (2003, 2010) studied the organisation of online learning. In a next step, the focus was directed on research networks as an organizational artefact, their functionality and technology. What can be concluded on how to steer the development and how to govern that functioning of those structures effectively?

4.2 Forms, Instruments and Mechanisms of Control in Virtual Organisations

Organisational theory examines traditional forms of governance (behavioural and output control) in detail, mostly uniformly. However, with the establishment of network-like organisational structures, the concept of social control has only recently been subjected to rigorous debate. Only the following forms of governance are considered here:

1. direct governance—inspection of behaviour (behavioural control), such as on the basis of standards won from experiences (Magretta 1998);
2. indirect governance—determination of output based on given goals (output control) (Thomson 1967; Magretta 1998);
3. social governance (social control)—comparison of conformity to certain moral and cultural rules (Ouchi 1979).

As Lattemann and Koehler (2005) argue, instruments of social control can be identified in relation to the level of objective and personnel management (Thomson 1967). Therefore, trust is not related to behavioural and output control mechanisms, as some authors postulate (see, e.g., Manchen and Grote 2000; Bradach and Eccles 1989), but rather supplementary to these (Das and Teng 1998; Ebner et al. 2003). In that sense, traditional control mechanisms and social control describe are different.

How can flexible and light organisational structures be designed and implemented? Based on the above discussion of the literature and cases, trust can be promoted by appropriate social standards and basic institutional conditions. A number of governance instruments can be applied to exercise social control, such as promoting common cultures among networking partners with homogeneous value creation processes, or reviewing and creating similar moral concepts through rituals or ceremonies. The observed networks apply different means, ranging from a project plan to an inter-institutional agreement. This method is particularly suitable for networking partners of a similar size, origin and organisational form (Ouchi 1979), that is, with almost no heterogeneity. Other effective means of social control include operational guidelines (Heck 1999), intensive use of modern and uniform ICT (Köhler 2003; Albers et al. 2002), promoters for public relations and conflict management (Hausschild 1997), job rotation or jointly offered training courses. In the three networks observed here, we found both inter-institutional agreements (such

as the integrated provision of academic master's programmes) and other measures (such as joint training) for using the platform.

Can the social control model (cf. Fig. 5) developed by Lattemann and Köhler (2005) for learning networks be transferred to research organisations with presumably less standardised activity?

The efficiency of the three governance forms discussed and the possible fields of their application depend upon the nature of the organisational arrangement. The more governance mechanisms are used; the more competencies are required in the process of cooperation. In contrast to traditional enterprises (Type 1 in Fig. 1), where mostly traditional forms of control (behaviour and output control) based on structural governance tools are used to promote coordination (information and communication) and motivation, virtual organisations may adopt concepts of social control with different degrees of intensity.

Virtual teams, virtual projects, temporary virtual organisations and meta-networks are characterised as maximally closed networks with unilateral dependency on the

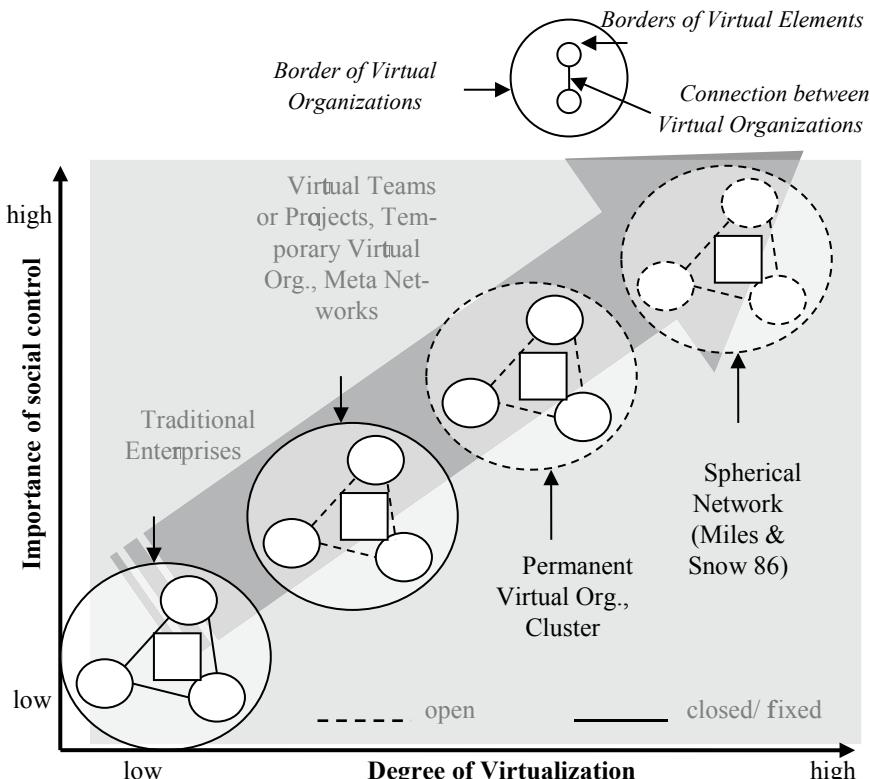


Fig. 5 Social control and organisational virtualisation (figure by authors, cf. Lattemann and Köhler 2005)

value creation process. The partners provide a wide spectrum of services and products. Such networks do not require a high degree of competency for cooperation. This reflects the fact that social governance tools were not applied intensively in these forms of virtual organisations. Business relations of this type are shaped by market-oriented or structural management instruments, such as a centralised coordinating body based on contractual arrangements (e.g. services or employment contracts). Virtual organisations like this frequently use ICT to collaborate and communicate. This is because both employees of the enterprise and long-term partners are often closely associated. Thus, ICT structures are implemented and do not need to be built up. Also—which is perhaps far more important—these structures do not need to be mediated between the partners, as they are obligatory in most temporary projects. Moreover, members of permanent virtual organisations and clusters need strong collaborative competencies due to their extremely intertwined mutual relations.

A maximum of informal relations is presupposed in spherical networks (Miles and Snow 1986). The roles of individual participants are distributed in a spherical network; resources and/or participants are boundlessly exchangeable. Such structures can be assumed in social networks; however, this article refers to profit-making, not non-profit, environments, so spherical networks are not the focus here. Even its proponents state that this structure cannot be observed in reality (Miles and Snow 1986).

In practice, the extent to which ICT is used to support coordination processes in virtual organisations varies greatly. However, in all virtual organisations, ICT plays a pivotal role; without it, virtual organisation is impossible. Research which was based on a set of unsystematic findings from case studies (Manchen and Grote 2000; Köhler and Schilde 2003; Köhler et al. 2003), recommended that the minimum required ICT support be identified first. The arrangement of information and communication processes determines the complexity of the ICT infrastructure (e.g. enterprise resource planning or e-mail). In less complex virtual organisations (e.g. virtual teams or projects), less sophisticated ICT solutions have been used in academic practice for approximately 20 years. However, in these research organisations, ICT-based groupware solutions were still rather exceptional (Köhler and Röther 2002; Köhler and Schilde 2003). More recently, it has been found out that only a small number of scientists are adopting social media technologies like Mahara, Mendeley or ResearchGate. For example, a Germany-wide survey conducted by Pscheida et al. (2015) found that social media applications such as social networks, microblogs and social bookmarking tools are used by a maximum of 8% of scientists in a research context. Only in 2020 the influence of the Corona pandemic will perhaps lead to a more massive adoption of such collaboration techniques, but not necessarily in a conscious use.

All in all, organisational models for academic institutions dealing with both education and research need to adapt to organisational models of virtual organisations. Universities and other research institutions have to change in both structure and process within their two main areas—education and research.

4.3 Limitations

Given the recent nature of this study, both the available literature and empirical access to the sectoral development were limited. Firstly, the empirical cases represent developments in German academia only. In the next stage, research must include data from other countries, to develop a more general understanding of organisational dynamics in the academic sector and avoid a national-only explanation.

Some sources, including website communications, were publicly available documents written by legal professionals or corporate representatives. Therefore, the case study may contain less reliable data than that supplied from exclusively academic sources.

Although the authors attempted to adopt a wide range of literature from several sub-disciplines in business, media and education studies, it is difficult to identify whether other researchers intentionally focussed on organisational development or whether this was a by-product of other considerations. Thus, the case made here is largely based on the previous work of the authors.

References

- Albers, S., Bisping, D., Teichmann, K., Wolf, J.: Management virtueller Unternehmen. Research Paper University of Kiel (2002)
- Barney, J., Hansen, M.: Trustworthiness as a Source of Competitive Advantages. *Strateg. Manag. J.* **15**, 175–190 (1994)
- Bradach, J., Eccles, R.: Price, authority, and trust: from ideal types to plural forms. *Ann. Rev. Sociol.* **15**, 97–118 (1989)
- Bryman, A., Bell, E.: Business Research Methods. Oxford University Press, Oxford (2011)
- Bulmahn, E.: Vorwort. In: BMBF: Förderprogramm Neue Medien in der Bildung Lehr- und Lernsoftware. Referat Öffentlichkeitsarbeit, Bonn (2004)
- Das T., Teng, B.S.: Between trust and control: Development confidence in partner cooperation in alliances. *Acad. Manag. Rev.* **23**, 491–512 (1998)
- Ebner, W., Leimeister, J.M., Krcmer, H.: Vertrauen in virtuellen Communities. In: Uhr, W., Esswein, W., Schoop, E. (eds.) Wirtschaftsinformatik 2003, Band II – Medien Märkte Mobilität, Physika: Wiesbaden, p. 619 (2003)
- Frindte, W., Köhler, T.: Computervermittelte Kommunikation und Unternehmen – Wirklichkeit und Virtualität?. In: Stahl, H. & Hejl, P.: Management und Wirklichkeit - Das Konstruieren von Unternehmen, Märkten und Zukünften; Carl Auer Verlag, Heidelberg (2000)
- Funkhouser, G.R., Shaw, E.F.: How synthetic experience shapes social reality. *J. Commun.* **40**(2), 75–87 (1990)
- Hausschildt, J.: Innovationsmanagement, 2nd edn. Vahlen, München (1997)
- Heck, A.: Strategische Allianzen: Erfolg durch professionelle Umsetzung. Springer, Berlin (1999)
- Hener, Y., Buch, F.: Evaluation des Bildungsportals Sachsen durch das CHE Centrum für Hochschulentwicklung. CHE, Gütersloh (2006)
- INEGI: Statistischer Report; Instituto Nacional de Estadística, Geografía e Informática México. INEGI, Mexico City (2004)
- Issing, J., Klimsa, P.: Information und Lernen mit Multimedia: Lehrbuch für Studium und Praxis; 3. vollst. überarb. Aufl. Beltz Psychologie-Verlags-Union, Weinheim (2003)

- Issing, L. J., Klimsa, P.: Online-Lernen. Ein Handbuch für das Lernen mit Internet; 2. Korrigierte Auflage. Oldenbourg Wissenschaftsverlag, München (2010)
- Kahnwald, N., Köhler, T.: Aufbau einer Online-Community für die nichtformale betriebliche Weiterbildung. In: Lattemann, C., Köhler, T. (eds.) Multimediale Bildungstechnologien I: Anwendungen und Implementation. Peter Lang Verlag, Frankfurt am Main (2005)
- Köhler, T.: Selbst im Netz? Die Konstruktion des Selbst unter den Bedingungen computervermittelter Kommunikation. Westdeutscher Verlag, Opladen (2003)
- Köhler, T., Ihbe, W.: Möglichkeiten und Stand der Nutzung neuer Medientechnologien für die akademische Lehre. Überlegungen zur aktuellen Situation an der TU Dresden; Wissenschaftliche Zeitschrift der Technischen Universität Dresden, 1-2 (2006)
- Köhler, T., Röther, U.: SAP-R/3 Einsatz durch einen externen Dienstleister im Konzernverbund. unpublished research report for SBI GmbH, Jena (2002)
- Köhler, T., Schilde, P.: From project teams to a virtual organization: The case of the education Portal Thuringia. *Front. E-Business Res.* (2) (2003)
- Köhler, T., Neumann, J., Saupe, V.: Organisation des Online-Lernens. In: Issing, L.J., Klimsa, P. (eds.) Online-Lernen. Ein Handbuch für das Lernen mit Internet (2. revised Version). Oldenbourg Wissenschaftsverlag, München (2010)
- Köhler, T., Wetzstein, A., Schilde, P.: Möglichkeiten der organisationalen Unterstützung räumlich verteilter Arbeits- und Lernprozesse. In: Enderlein, H., Müller, E. (eds.) Vernetzt planen und produzieren. Universitätsverlag, Chemnitz (2003)
- Krysteck, U.: Vertrauen als Basis erfolgreicher strategischer Unternehmungsführung. In: Hahn, D., Taylor, B. (eds.) Strategische Unternehmensplanung - Strategische Unternehmensführung. Stand und Entwicklungstendenzen 7. Aufl. Springer, Heidelberg (1997)
- Landfried, K.: Vortrag auf der Hochschultagung der Gewerkschaft Erziehung und Wissenschaft (GEW), Sylt, 3 Sept 2009 (1999/2009)
- Lattemann, C., Khaddage, F.: A review of the current status of mobile apps in education. In: Alon, I., Jones, V., McIntyre, J. (eds.) Innovation in Business Education in Emerging Countries. Macmillan-Palgrave, pp. 181–192 (2013)
- Lattemann, C., Köhler, T.: Trust or control—governance concepts for virtual organizations. *Front. e-Business Res.* 3 (2005)
- Lievrouw, L.A., Bucy, E., Frindte, W., Gershon, R., Haythornthwaite, C., Köhler, T., Metz, J., Sundar, S.S.: Current research in new media: an overview of communication and technology. In: Gudykunst, W. (ed.) Communication Yearbook 24. Lawrence Erlbaum Publishers, Mahwah (2000)
- Magretta, J.: The power of virtual integration: an interview with dell computer's Michael Dell. *Harvard Bus. Rev.* 76(2), 73 (1998)
- Malone, M.S., Davidow, W.H.: The Virtual Corporation: Structuring and Revitalizing the Corporation for the 21st Century. HarperCollins, New York (1992)
- Manchen, S., Grothe, G.: Vernetzte Zusammenarbeit von Unternehmen. Ergebnisse aus vier Fallstudien; Research Paper of Institut für Arbeitspsychologie der ETH Zürich (2000)
- Miles, R. E., Snow, C. C.: Organizations: New concepts for new forms, in: *California Management Review*, 28 (3), pp. 62 (1986)
- Misoch, S., Köhler, T.: „... man muss gute Didaktik betreiben [...] und da gehört dieser ganze Computerquatsch dazu ...“; Final report on new media usage in teacher training of Potsdam University (2005)
- Munkvold, B.E.: Implementation and use of collaboration technologies in a multinational engineering group: the case of Kvaerner. In: Munkvold, B.E. (ed.) Implementing Collaboration Technologies in Industry: Case Examples and Lessons Learned, pp. 109–128. Springer Verlag, Berlin (2003)
- Neumann, J., Schütte, T.: An organizational model for e-learning at universities. In: Hannula, M., Koiranen, M., Tommila, J. et. al. (eds.) Proceedings of EBRF 2007. University of Technology, Tampere (2008)
- Okkonen, J.: Performance in virtual organisations. *Front. e-Business Res.* 1, 267–279 (2002)

- Ortner, G., Nickolmann, F.: Socio-Economics of Virtual Universities. Beltz Deutscher Studien Verlag, Weinheim (1999)
- Ouchi, W.G.: A conceptual framework for the design of organizational control mechanisms. *Manag. Sci.* **25**(9) (1979)
- Postmes, T.: Social Influence in Computer-Mediated Groups. Print Partners Ipskamp, Enschede (1997)
- Pscheda, D., Minet, C., Herbst, S., Albrecht, S., Köhler, T.: Nutzung von Social Media und onlinebasierten Anwendungen in der Wissenschaft. Ergebnisse des Science 2.0-Survey und Datenreport Hochschulen 2014. <http://nbn-resolving.de/urn:nbn:de:bsz:14-qucosa-163135> (2015). Accessed 1 Nov 2020
- Pscheda, D., Albrecht, S., Herbst, S., Minet, C., Köhler, T.: Nutzung von Social Media und onlinebasierten Anwendungen in der Wissenschaft. Erste Ergebnisse des Science 2.0-Survey 2013 des Leibniz-Forschungsverbundes „Science 2.0“, TUD Press, Dresden (2014)
- Schmidt, K.: Sonderpublikation anlässlich der Eröffnung des Bildungspfads Thüringen. BP-T, Erfurt (2002)
- Snow, C., Lipnack, J., Stamps, J.: The virtual organization: promises and payoffs, large and small. In: Cooper, C.L., Rousseau, D.M. (eds.) *The Virtual Organization*. Wiley, Chichester (1999)
- Sproull, L., Kiesler, S.: Reducing social context cues: electronic mail in organizational communication. *Manag. Sci.* **32**(11), 1492–1512 (1986)
- Stifterverband für die deutsche Wissenschaft (Hrsg.): Campus online - Hochschulen, neue Medien und der globale Bildungsmarkt. Stifterverband für die deutsche Wissenschaft, Essen (2001)
- Thomson, J.: *Organizations in Action*. McGraw-Hill, New York (1967)
- Weichselgartner, E.: Digitales wissenschaftliches Arbeiten in der Psychologie. In: Petermann, F., Koglin, U. (eds.) *Abstractband zum 47. Kongress der Deutschen Gesellschaft für Psychologie*. Pabst, Lengerich (2010)
- Yin, R.K.: Qualitative Research from Start to Finish. Guilford, New York (2011)
- Yin, R.K.: Case Study Research: Design and Methods. Sage, Thousand Oaks, CA (2013)
- Ziegler, R., Diehl, M.: Computer-supported idea generation. In: Rickards, T., Runco, M.A., Moger, S. (eds.) *The Routledge Companion to Creativity*. Routledge, New York, NY (2009)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



The Fish Model: When Do Researchers Collaborate Online?



Bahaeldin Mohamed and Thomas Köhler

Abstract The questions of whether and how doctoral students are motivated for enhanced research collaboration deserve thorough consideration. Even though collaboration in general and its mediated forms, such as *computer-supported cooperative work and collaborative learning* (CSCW and CSCL), are prominent research topics, only a little is known about the methods necessary to design various activities to support research collaboration. With the upcoming generation of tools such as Mendeley, Conference Chair, ResearchGate, or Communote, scholars suspect that web 2.0 services play a decisive role in enabling and enhancing research collaboration. However, there is almost no data available on the extent to which researchers adopt these technologies, and how they do so. Therefore, the authors first present an overview of the current usage of web 2.0 among doctoral researchers in their daily academic routines, based on a survey ($n = 140$) conducted in the German Federal State of Saxony. It confirms a wide and often specified usage of web 2.0 services for research collaboration. For theoretical analysis, the authors propose a conceptual framework that reflects the requirements of scientific participation and scholarly collaboration within an average international doctoral programme adopting current digital technologies. The aim of this framework is to understand, support, and enhance research collaboration among doctoral researchers. Our fish model highlights the mutual relationship between the following dichotomous factors: (a) tasks/time factors; (b) beliefs/activities; (c) support/context; and (d) incentives/ethical issues. Our results indicate a significant relationship in terms of research collaboration. This relationship has particularly been identified between two dichotomous factors: beliefs/activities and incentives/ethics.

Keywords Research collaboration · e-science · Web 2.0 technology · Scholarly communication · Doctoral training

B. Mohamed (✉)
British Lincoln College, Riyadh, Saudi Arabia
e-mail: bahaa@bzoor.com

T. Köhler
Institute for Vocational Education, TU Dresden, Dresden, Germany
e-mail: thomas.koehler@tu-dresden.de

1 Introduction

Research collaboration is the foundation of research students' efforts in academia. Independently of disciplinary background, research is based on the social patterns of competition for the best explanation and joint evaluation of the quality of research. Therefore, research collaboration is a form of positive interaction between knowledge producers that have taken on management roles by using certain resources and tools to establish and pursue a scientific goal (Ynalvez et al. 2011). We define research collaboration as the current and future regulations, processes, and concepts which support interaction and cooperation between our doctoral candidates. Here, it is important to note that collaboration is not simply students and professors co-authoring a piece of research; instead, it requires establishing connections that might extend to communication which, over time, develops into sustainable collaboration among different researchers with similar interests. Accordingly, we may need to better understand the nature of scientific tasks and the time frame in which they should be completed, as well as how individual beliefs of using ICT and web 2.0 in a research context can help to define how online activities should be organised. In addition, the use of technology can be interpreted in relation to cultural contexts and disciplines. Finally, incentives act as the engine that encourages students to undertake collaborative research, and, in academia, this engine is covered and protected by research ethics. In this paper, we focus on collaboration of all PhD students in their first, second, or third year. This may take into consideration the form of any formal or informal social action and scientific activities that could increase the output and production of scholarly research, improve communication through the text, and encourage resource sharing and collaborative writing.

PhD students face new challenges in the age of digital research. In particular, this paper focuses on challenges such as dealing with digital material and resources, learning management systems, personal learning environments, social networks, and collaboration in research networks. Current PhD students, who are largely from the Generation Y demographic group (born between 1982 and 2000), are familiar with technology and are likely to encounter one or more web 2.0 technologies in their everyday life (Zaman 2010). In the academic context, web 2.0 technology shapes how PhD students learn, self-regulate, and communicate. Accordingly, universities have begun to use and provide these facilities of infrastructure to attract and connect students and develop—step by step—a better practice for research collaboration. However, as Zaman (2010) reports, current doctoral programmes struggle to follow up and meet these demands and requirements. Concerning social and scientific interaction and collaboration among our doctoral students, Mohamed et al. (2013) investigated PhD students' attitudes towards doctoral colloquium, online learning material

via Edu-tech,¹ and learning management systems via OPAL.² These scientific activities were used simply to provide an informative website for learning material and scientific events; PhD candidates usually found that the community of practice and the feeling of belonging were lacking.

We expect the digital form of research so-called e-research collaboration to comprise the attempt to enhance and develop not only scientific activities such as co-authorship or finding peers and peer reviewers, but what we refer to as “open-kitchen research”. This term refers to sharing research activities not only as a finished product, but also as processes. In fact, during the doctoral candidate education, they attempt to communicate and collaborate only in the context of theoretical curriculum. These learning formal courses are traditionally designed to provide students with only structured theoretical knowledge but no real practices. In most cases, we observed that part-time PhD students working in third-party projects at our laboratory give priority than ever before to the projects they are working in where there is more community support than working individually with their own dissertation.

The relevance of this study can be confirmed by the fact that doctoral education in Germany is rapidly growing in all academic disciplines, to a recent total number of 200,400 doctoral candidates being supervised at German universities (in the winter semester 2010/2011), while only half of this group ($n = 104,000$) was officially registered (Forschung & Lehre 2012; Wolters and Schmiedel 2010). How do those registered scholars participate in research activities? Do they follow their academic activities at the same pattern and do they regularly use the same research online tools? We can just guess that the new openness of social media and web 2.0 communication helps to provide similar conditions and borderless collaboration for all scholars depending on their access to the Internet. In the German Federal State of Saxony, where the data of this study was collected, the number of PhD degrees has increased more than tenfold, from $n = 111$ in 1993 to $n = 1,206$ in 2009 (Saxony State, Statistical Branch 2009).

In order to provide an adequate statement about how our novice researchers collaborate via using web 2.0 services, we explore which factors might shape this collaboration, particularly the collaborative opportunities offered by web 2.0, we begin by developing a theoretical framework for our investigation, and apply it to the current situation of PhD students in Germany.

¹This study focused on the European doctoral network “Education & Technology” (cp. <http://edutech.eu>).

²OPAL, an open-source Learning Management System, used by all universities of the Federal German State of Saxony (cp. <https://bildungsportal.sachsen.de/>).

2 The Fish Model: A Conceptual Framework for E-Research Collaboration

The authors conceptualised e-research collaboration as follows. Based on a meta-analysis, approximately 200 papers focussing on different aspects and approaches in e-science and e-humanities were recruited, organised, and analysed, in order to formulate a proposed conceptual framework, the fish model, previously published in Mohamed et al. (2013). The framework may be used to deepen our understanding of the daily scientific tasks, activities, technologies, and incentives that shape everyday academic practices for doctoral scholars, regardless of their disciplinary heritage. Databases consulted include Science Direct, Pro-quest, EBSCO, Scirus, and Mendeley. Inclusion criteria were limited to full-text papers concerning the use of web 2.0 in research communication and collaboration. Keywords used for collecting scientific articles directly from the mentioned databases included the following: researchers' digital habits, use of web 2.0 in research, e-research, social media in research, research collaboration, and scholarly communication. The following selection criteria were used for papers: (1) written in English, (2) situated only on the PhD and researcher levels, (3) either empirical or review articles only. In addition, a conceptual definition of collaboration factors from Patel et al. (2011) and the Folk Model of Intentionality (DeAndrea 2012) were used as guides to identify the fish model (Ringle et al. 2005). The first step in analysing the selected papers was to interpret online research behaviours and the academic activities associated with using web 2.0 technologies, in order to predict the future of research collaboration, using the Fish Model (Mohamed et al. 2013). As the model clarifies the factors and concepts behind the best practices associated with research collaboration using web 2.0 technologies, it was proposed to develop an understanding of daily scientific research tasks and activities.

As the authors suggested earlier, online research behaviour is controlled by some key factors and indicators, which was first framed in the Model of Collaborative e-Research (Reebs 2011). This model can be used to describe the factors that support online collaboration in e-science. The fish model (Mohamed et al. 2013), however, extends this research by giving evidence that individual factors (beliefs, self-regulation, etc.), in addition to group interaction organised by the institution, and time management, obviously influence the active production of research, communication among researchers, and subsequent collaboration. Using the fish model, the core factors in online research behaviours and the academic activities associated with using web 2.0 technologies all were investigated.

It is argued that a doctoral scholar would behave "like a fish living in a specific environment, taking part in a particular community, showing different individual behaviours to respond to an action, led by their own beliefs and framed by a certain culture" (Mohamed et al. 2013, p. 3275). Typical behaviours and activities are managed by incentives related to the qualification addressed and controlled by the scholar's role in the research ecology. The fish metaphor emerged when framing a body of collaboration patterns for the authors' previous study (Frewox 2010).

“Collaboration in research is managed by a dorsal fin to stabilise research against rolling and protect scientific environment from isolation and weakness. Inhalation through the mouth passes over the gills in fish to obtain fresh oxygen, communication is the oxygen of research project which is necessary for bringing activities and ideas to the project and achieve the tasks related. The backbone of our fish is web 2.0 technologies which connect and facilitate all functions of the whole body, these functions are divided concerning a dichotomous aspect (fish spine) – as we will describe it complementarily in the frame of this paper – in a task/time, activities/beliefs, support/context, and ethics/incentives division” (Tannen 2006, p. 3267 ff.).

Research collaboration is usually considered as a planned activity where knowledge can be produced and transferred. The authors predicted previously (Mohamed et al. 2013) that collaborative e-research (using web 2.0 technology to improve best research practices) will take place alongside dichotomies. Tannen (Wang 2010), in his book, The Argument Culture (1998), proposes the concept of perceived dichotomies, that is, binarisms between two connected concepts, while not distinguishing between them through the use of vocabulary such as “good” and “bad”. Building on Tannen’s work, the fish model proposes the integration of both factors. Research collaboration in this study can be interpreted as a relationship between eight concepts formed in pairs making up the total of four groups: (a) between scientific tasks or candidates’ needs and time available for implementing them; (b) between planned activities and individual research beliefs in dealing with these activities; (c) support from technology and understanding the uses of this technology within a certain context and culture of an institution; and (d) intentions/motivations for collaboration, which are directed by research ethics, as illustrated in Fig. 1.

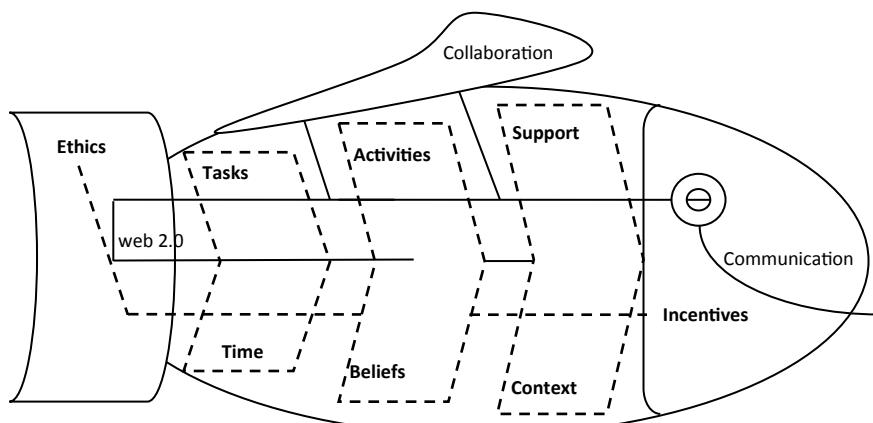


Fig. 1 Fish model: conceptual framework for developing e-research collaboration for PhD students and novice researchers (Mohamed et al. 2013)

2.1 The Reality of Managing Scientific Tasks in Terms of the Available Time

It can be expected that novice researchers are likely to collaborate and work with each other because they are more likely than experienced researchers to break their work down into various tasks, activities, and actions. Such individual behaviour is controlled by time management as short-/long-term academic tasks, primarily related to different actions such as information search, data analysis, reading, or possibly writing (Illeris 2004). Overall, the doctoral education system differs significantly from programmes at masters and bachelor level, as doctoral programmes prepare candidates for high-level careers in industry or provide long practical experience (Zaman 2010). In their previous studies (Mohamed et al. 2013; Mohamed et al. 2013), the authors identified two key tasks that doctoral students undertake in order to carry out their research. The first is marketing, that is, building a scientific competence profile in order to develop a scientific reputation. The second is doing research, that is activities in daily research practice, including mainly reading, writing, investigating, searching, and reviewing.

- H1:** Novice researchers are more likely to collaborate and work with each other when the work task (types, stages, and technologies) and timeframe are specified.
- H1-0:** An academic task to be done via web 2.0 is driven by a timeframe (when the task should be done/how much time is needed to do it).
- H1-1:** An appropriate timeframe for a task to be carried out via web 2.0 can lead to academic collaboration

2.2 Online Research Activities Led by Work-Based Beliefs

PhD students' daily research activities include specific online activities, as identified previously (Mohamed et al. 2013): accessing resources, information, and research funds; engagement in scientific discussions and being an active member in one or more academic communities of practice; communication in reviewing, sharing, and exchanging ideas; awareness of recently published scientific papers and events; presenting oneself online in social media and social networking in order to build up a profile and identification (Mohamed 2011; Lahenius 2010; Peggy and Borkowski 2007).

Typically, it is expected that PhD research work is completed through three main development phases (Terrell et al. 2009; Zaman 2010; Mohamed et al. 2013): (a) becoming a researcher by training, and reading activities for first-year PhD students; (b) becoming an expert in any required methods and the pressure to start publishing for second-year PhD students; and (c) becoming an author which includes participating in peer reviewing, co-authoring, and writing publications. Each of those phases requires a number of planned online activities. Additionally, gradual engagement

with the literature of one's own scientific discipline should be considered, because it leads to particular work beliefs. Three main explanations for scholars' success were identified (Patel et al. 2011): social culture, the culture of disciplines, and the individual beliefs (values, motivation, learning style, self-regulation, cognitive competence, confidence, and trust). Usually, beliefs are addressed by psycho-educational research, whereas the role of trust (versus control) as a governance concept has been addressed in earlier research on virtual organisations (Lattemann and Köhler 2005). Only the combination of these accepted beliefs defines a researcher's individual approach to scientific activities.

- H2:** Novice researchers are more likely to collaborate and work with each other when they believe in the work and participate in academic and research activities (online).
- H2-0:** Academic activities (online) affect a researcher's belief in using web 2.0 toward collaboration.
- H2-1:** Researchers' belief in using web 2.0 for research may increase their chances of collaboration

2.3 Support for Technology Use in Context

Even though web 2.0 is a rather young technology, multiple studies have investigated its benefits for learning, especially in the production and communication of scientific research, or e-science (Pscheidt et al. 2013; Kahnwald et al. 2015). A core aspect of ICT infrastructure (web 2.0) is its strong linkage to the sociocultural context and the disciplinary culture. While academic work triggers social interactions among PhD scholars, the cultural context drives and assists their use of web 2.0 technologies in order to interact. ICT and web 2.0 services in learning and research comprise all methods, techniques, online behaviours of scientists, tools used by researchers, knowledge sharing and transfer, acceptance/adoption, and building social networks via e-research identified by literature reviews (Meyer and McNeal 2011). A doctoral candidate's use of web 2.0 technologies is both supported by and understood through institutional context and discipline culture (Pscheidt et al. 2013). Those have a particular need for being involved in one or more academic communities on a national or international level in order to share and develop practice successfully, usually realised through web 2.0 services (Veletsianos and Kimmons 2012; Eyman et al. 2009; Illeris 2004; Gillet et al. 2009; Lam 2011).

- H3:** Researchers are more likely to collaborate when they have received technical support in their academic context.
- H3-0:** web 2.0 technology may enhance research communication, leading to future collaboration.
- H3-1:** Research context has a direct influence on collaboration

2.4 Incentives Protected by Research Ethics

PhD candidates need incentives to be strongly engaged in online collaboration (Pidd 2011); these incentives are intrinsic motivation, satisfaction, and reputation. Purely financial motivation is less important, but the motivation should be protected and controlled by research ethics related to the digital environment (Mutula 2010). The issue of trust should be considered by faculty involved in digital research processes (Jirotnka et al. 2006), as it has a special role in steering online networks (Lattemann and Köhler 2005). Young researchers need to develop e-strategies to use research portals to ensure and facilitate authentic human sources for knowledge transfer. While the majority of them have adopted web 2.0 tools already, their willingness to shift from offline to online digital research practices is crucial (Pscheidt et al. 2013; 2014) to build trust and protect scientific work in a virtual environment (Lam 2011).

- H4:** Novice researchers are more likely to collaborate and work with each other when they receive incentives (as external motivation) that are protected and combined with their trust and the value of their work (as internal motivation).
- H4-0:** Incentives as an external motivation can influence ethics as an internal motivation for enhancing research collaboration.
- H4-1:** Research ethics as an internal motivation is closely related to research collaboration

3 Method

For this paper, data was collected and analysed through the combination of two main methods: (a) description of a quantitative online survey conducted in the German Federal State of Saxony from 22 July 2012 until 22 October 2012, at the Technische Universität Dresden and (b) forming and testing the structured model. The main aim was to investigate novice researchers' intentionality to collaborate with each other through the use of web 2.0 and digital online technologies in academia. Our survey included two main parts: the first part reveals demographic data and the second part includes a 5-interval Likert scale with points ranging from 1 (strongly disagree) to 7 (strongly agree). The survey addressed doctoral students as novice researchers who are using web 2.0 technology to communicate and collaborate in research daily life. This 45-item measure was created for this study to assess participants' perceptions, profiling the nine main factors that shape the final structure of the fish model: task, time, activity, belief, support, context, incentive, ethics, and collaboration. The instrument was then tested by three independent experts in research collaboration before being given to respondents from the target audience. The authors received a total return of $n = 140$ doctoral students who completed the survey. The data was examined using factor analysis and our fish model was tested with the Partial Least Squares (PLS) technique. SmartPLS, Version 2.0 M3 software was used to test the model (Ringle et al. 2005, p. 1).

4 Results

The majority of respondents (57.71%) were male, 66.74% were not married and had no children, and 30.45% of respondents were from the School of Science, which includes the 13.41% which were PhD students from the Faculty of Mechanical Engineering. This can be considered typical for Saxony's higher education landscape, as it has a special focus on technical subjects.

4.1 *The Measurement Model*

PLS is “the second-generation structural equation modelling technique that assesses both the measurement and structural model in a single run” and was chosen for two reasons: it works well for smaller sample sizes and eliminates restrictions on data distribution such as normality (Serenko 2008, p. 465). Before analysing this model, its reliability was measured. Cronbach's alpha exceeded the required threshold of 0.7 for all items, implying high internal consistency of the scales (Serenko 2008).

In order to submit an accepted level of eligibility for the questionnaire, half of the items (24 of 45 items) were removed which do not have sufficient weight vis-à-vis their main factor (Table 5, see Appendix). Once these items were removed, the model was re-estimated. Reliability results are given in Table 1. The data shows that the measures are robust in terms of their internal composite reliability. The composite reliability of the different items ranges from 0.8 to 1.0, above the recommended starting value of 0.70 (Serenko 2008). In addition, consistent with the guidelines of Fornell and Larcker (Birnholtz 2005), the average variance extracted (AVE) for every component is above 0.50. Table 2 presents the results of measuring the discriminant validity for variable constructs. The matrix diagonal reports that the square roots of

Table 1 Assessment of the measurement model

Variable constructs	Composite reliability (internal consistency reliability)	Average variance extracted/explained (AVE)
Time	0.80	0.57
Task	0.80	0.57
Support/tech	0.88	0.66
Incentives	0.83	0.62
Ethics	1.00	1.00
Context	0.82	0.69
Collaboration	0.84	0.58
Beliefs	0.83	0.62
Activities	0.85	0.59

Table 2 Discriminant validity (inter-correlations) of variable constructs

Latent variables	1	2	3	4	5	6	7	8	9
1. Time	1.00								
2. Task	0.67	1.00							
3. Support/Tech	-0.43	-0.39	1.00						
4. Incentives	-0.23	-0.20	0.59	1.00					
5. Ethics	-0.32	-0.30	0.37	0.33	1.00				
6. Context	-0.027	-0.10	0.30	0.23	0.00	1.00			
7. Collaboration	-0.31	-0.30	0.50	0.50	0.46	0.06	1.00		
8. Beliefs	-0.44	-0.36	0.66	0.53	0.49	0.11	0.65	1.00	
9. Activities	-0.28	-0.22	0.35	0.37	0.34	0.25	0.41	0.54	1.00

the AVEs are greater in all cases than the off-diagonal element in their corresponding row and column, which supports the discriminant validity of the instrument.

The instrument was tested additionally through PLS-Graph and for convergent validity. Table 4 (see Appendix) shows the factor loading of all items to their respective latent constructs. All items loaded on their respective construct from a lower bound of 0.70 to the upper bound of 0.85. In addition, the *T*-test of outer model loading in the PLS-Graph output was highly significant ($p < 0.001$) for each factor's loading on its respective construct. The results confirm the convergent validity as demonstrating a distinct latent construct.

4.2 The Structured Model

Figure 2 presents the results of the structured model with interaction effect. In order to assess the structured model, a bootstrapping technique was applied. The examination of *t*-values was based on a 1-tail test with statistically significant levels of $p < 0.05$ (*), $p < 0.01$ (**), and $p < 0.001$ (***)¹. Dotted lines highlight the insignificant paths. Structured components were formulated by multiplying the corresponding indicators of the predictor and moderator construct.

For clarity purposes, the outcomes of the structural model in terms of direct effects, bootstrapping, and *t*-statistics confirmed the majority of the hypotheses, at various significance levels. However, the results show that only two factors in research collaboration are associated significantly (Fig. 2). Specifically, "Academic activities" is very significantly associated with "Researchers' beliefs" (H2-0 at $\beta = -0.67$, $p < 0.001$ level). In this first path, "Researchers' beliefs" has a significant relation with "Collaboration" (H2-1 at $\beta = 0.41$, $p < 0.001$ level). In the second path, "Incentives" and "Ethics" contribute significantly to "Collaboration". Accordingly, (H4-0) confirms a significant relation between "Incentives" and "Ethics" (H4-0 at $\beta = 0.71$,

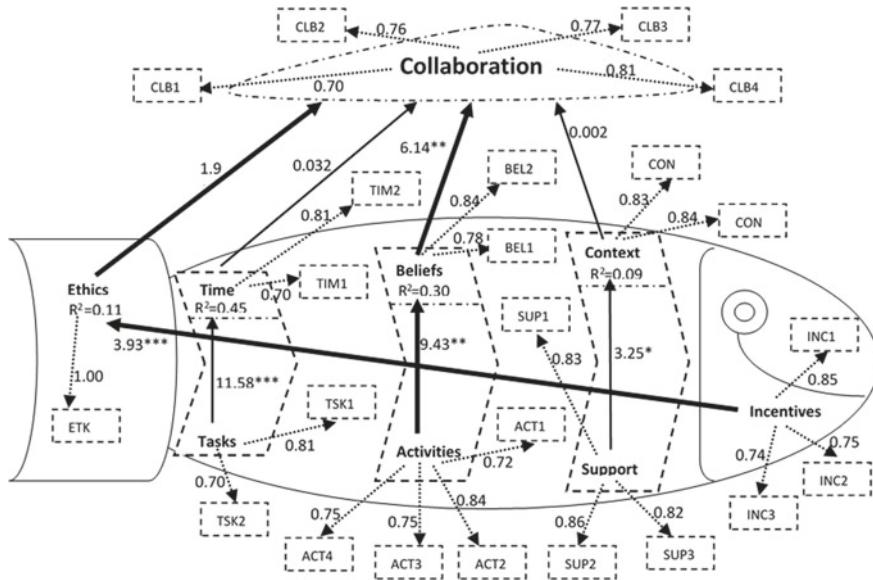


Fig. 2 Structure model (PLS bootstrapping “path coefficient”). *significant at 0.05 level (1.96); **significant at .01 level (2.58); ***significant at 0.001 level (3.29)

$p < 0.001$ level) along with the relationship (H4-1) between “Ethics” and shaping “Collaboration” ($\beta = 0.06, p < 0.05$).

The other two paths of predicting research collaboration are not significant. First, “Technology and support” has a significant relationship with “Context” (H3-0 at $\beta = 0.64, p < 0.05$), but, as a second path, the “Context” cannot predict research “Collaboration” (H3-1 at $\beta = 0.00$ not significant). Second, academic “Task” connected strongly with the factor “Time” (H1-0 $\beta = -0.70, p < 0.001$). On the other hand, the relationship between “Time” and shaping academic “Collaboration” (H1-1 $\beta = -0.00$, not significant) was unrelated in the context of shaping academic collaboration (Table 3).

5 Discussion: Conclusion and Limitations

5.1 Conclusions

The results of this study demonstrate the factors that might influence research collaboration among novice researchers in Germany. The study conceptualised and validated the fish model for understanding research collaboration in the digital age, highlighting where the model can be extended. A brief review of the findings raises the question of what drives researchers’ propensity to collaborate using web 2.0 services.

Table 3 Research hypotheses and conclusions

Hypothesis	β (path-coefficient)	t-value	p-value	Validation
H1-0: An academic task to be done via web 2.0 is driven by a timeframe (when the task should be done/how much time is needed to do it)	0.70	12.61	< 0.001	Supported
H1-1: An appropriate time frame for a task to be carried out via web 2.0 can lead to academic collaboration	-0.00	0.03	n.s.	Rejected
H1: Novice researchers are more likely to collaborate and work with each other when the work task (types, stages, and technologies) and time frame are specified				Rejected
H2-0: Academic and research activities (online) affect a researcher's belief in using web 2.0 for collaboration	0.67	9.71	< 0.001	Supported
H2-1: Researchers' belief in using web 2.0 to support research may increase their chances of collaboration	0.41	5.77	< 0.001	Supported
H2: Novice researchers are more likely to collaborate and work with each other when they believe in the work and participate in academic and research activities (online)				Supported
H3-0: web 2.0 technology may enhance research communication, leading to collaboration	0.64	3.74	< 0.05	Supported
H3-1: Research context has a direct influence on collaboration.	0.00	0.00	n.s.	Rejected
H3: Researchers are more likely to collaborate when they have received technical support in their academic context				Rejected
H4-0: Incentives as an external motivation can influence ethics as an internal motivation for enhancing research collaboration	0.71	3.79	< 0.001	Supported
H4-1: Research ethics as an internal motivation is closely related to research collaboration	0.06	1.74	< 0.05	Supported
H4: Novice researchers are more likely to collaborate and work with each other when they receive incentives (as external motivation) that are protected and combined with their trust and the value of their work (as internal motivation)				Supported

The first collaboration path showed that doing online doctoral research activities might shape beliefs in using web 2.0 technologies for academic purposes and, thus, enhance collaboration. An example is that using social media to connect with like-minded people eventually shapes one's belief about the importance of web 2.0. Researchers who believe in using such media are more likely to collaborate and more open to empathy.

Overall this study illustrates how the fish model can be applied to an online setting in order to understand how the interaction between academic activities and researchers' beliefs can influence research collaboration. The results are consistent with the previous mentioned literature as it was discussed by Terrell et al. (2009), being successful can shape a person's individual beliefs. Engaging in online research activities in order to communicate and collaborate reflects individual beliefs that control the actions that can enhance further collaboration offline, as has been observed in a professional context (Köhler 1997). Researchers' activities may reveal some of the individual beliefs that back and catalyse collaboration. When researchers engaged in online research activities, their belief in the use of web 2.0 in research increased. Use of web 2.0 services such as social media can also predict productive and conductive research collaboration (Pscheida et al. 2013).

The second collaboration path shows that in keeping a balance between internal "ethics" and external "incentives", motivation can confirm collaboration. An example is that researchers' trust in sharing their ideas via web 2.0 services only grows when they benefit from using such technology and, accordingly, it may lead to collaboration. These findings have important implications for the fish model. Internal and external motivations support future research collaboration. We argue that external and internal motivations are closely related; consequently, in academia both types of motivation help researchers become engaged in collaboration. Higher incentives predict higher levels of trust; researchers are more likely to collaborate when they trust the technologies they use. What motivates researchers to enhance collaboration into the web 2.0 sphere depends on the technologies they can trust and use to extend their professional networks. For collaboration among researchers, trust is synonymous with benefit, which is the catalyst for collaboration.

5.2 *Limitations*

In this study, research collaboration was defined as the use of web 2.0 technologies for communication and daily research routines (reading, searching, writing, etc.). The authors addressed a subset of the concept labelled e-science or science 2.0. They empirically observed doctoral scholars. These PhD students came mainly from the Faculty of Mathematics and School of Science at the Technische Universität Dresden in Germany. These aspects may limit the range and meaning of the findings presented.

Another limiting aspect is that the fish model reported only two significant paths that may predict research collaboration. It would, however, be more informative if measures of the other paths of the fish model (that appeared as non-significant in our study) were measured once again in a different research context with another sample.

Appendix

See Tables 4, 5, and 6.

Table 4 Factor loadings (bolded) and cross loadings

Factor	Time	Task	Support/Tech	Incentives	Ethics	Context	Collaboration	Beliefs	Activities
TIM1	0.70	0.46	-0.29	-0.14	-0.19	0.05	-0.08	-0.20	-0.17
TIM2	0.81	0.52	-0.49	-0.25	-0.32	-0.07	-0.33	-0.50	-0.33
TSK1	0.57	0.81	-0.26	-0.20	-0.27	-0.03	-0.29	-0.30	-0.15
TSK2	0.47	0.70	-0.24	-0.11	-0.27	-0.09	-0.17	-0.21	-0.31
SUP1	-0.42	-0.44	0.83	0.44	0.30	0.23	0.44	0.60	0.29
SUP2	-0.35	-0.28	0.82	0.45	0.35	0.27	0.39	0.49	0.29
SUP3	-0.31	-0.24	0.86	0.57	0.29	0.29	0.40	0.58	0.29
INC1	-0.21	-0.11	0.40	0.85	0.32	0.07	0.38	0.45	0.34
INC2	-0.22	-0.26	0.67	0.75	0.21	0.35	0.42	0.50	0.32
INC3	-0.10	-0.14	0.37	0.74	0.23	0.18	0.40	0.31	0.20
ETK	-0.32	-0.30	0.37	0.33	1.00	0.00	0.46	0.49	0.34
CON1	0.02	-0.05	0.25	0.24	0.05	0.84	0.08	0.09	0.21
CON2	-0.064	-0.129	0.26	0.14	-0.04	0.83	0.02	0.09	0.21
CLB1	-0.25	-0.29	0.40	0.48	0.25	0.18	0.70	0.50	0.22
CLB2	-0.25	-0.24	0.31	0.32	0.39	0.01	0.76	0.49	0.42
CLB3	-0.19	-0.17	0.38	0.37	0.32	0.05	0.77	0.45	0.30
CLB4	-0.24	-0.22	0.43	0.36	0.41	-0.03	0.81	0.53	0.25
BEL1	-0.22	-0.25	0.48	0.32	0.35	0.04	0.41	0.78	0.50
BEL2	-0.39	-0.29	0.54	0.53	0.50	0.09	0.63	0.84	0.39
ACT1	-0.14	-0.11	0.16	0.26	0.18	0.20	0.22	0.27	0.72
ACT2	-0.34	-0.30	0.39	0.35	0.33	0.15	0.42	0.56	0.84
ACT3	-0.26	-0.16	0.25	0.21	0.28	0.13	0.30	0.44	0.75

(continued)

Table 4 (continued)

Factor	Time	Task	Support/Tech	Incentives	Ethics	Context	Collaboration	Beliefs	Activities
ACT4	-0.00	-0.01	0.20	0.32	0.21	0.37	0.25	0.28	0.75

Table 5 Items removed

Factor	Item
TSK3	It is more effective to ask colleagues for a declaration about unclear work tasks by E-Mail or Skype than face-to-face
TSK4	The task of reading an online paper is more effective than a printed one
TSK5	Task of searching/sorting for literature review by using web 2.0 services more difficult than my traditional way
TSK5	When I need to effectively discuss something related to my research with colleagues, web 2.0 services are not the right solution
TIM3	It is necessary to invest a lot of time for communicating, researching, and working via web 2.0 services
TIM4	Web 2.0 services save time for organising and managing our teamwork and working in a scientific community
TIM5	Usage of synchronous web 2.0 services (in real time interaction) such as chat or video conferences is more useful for managing online discussions than asynchronous tools (e.g. online forum)
SUP4	The usage of web 2.0 services in a scientific research is difficult and I can't understand it
SUP5	I use web 2.0 services when others recommend something really interesting for me
SUP6	Peers and colleagues warn me to use web 2.0 services in research
CON3	My institute/faculty does not formally support using web 2.0 services among doctoral students
CON4	The best way to contact my supervisor is through e-mail
CON5	Collaboratively reading, writing, and reviewing a paper via web 2.0 services in our project/research group is not familiar yet
INC4	Receiving daily information about a recent paper, event, or colleagues' activity, is a big motivation for me to use web 2.0 services
INC5	Editing, commenting, reading, and reviewing dissertation tasks by using desktop word processing software are more familiar to me than using web 2.0 services
ETK2	Taking on more responsibility in scientific editing, reviewing, commenting via web 2.0 services among researchers is ambiguous and uncertain
ETK3	Web 2.0 services signify for me a place where there is a lower level of data security
ETK4	Data security for me is an important issue for participating in any scientific editing, reviewing, commenting, and reading via web 2.0 services
ETK5	My data can be stolen easily via web 2.0 services
ACT5	Giving online lectures is one of my usual online activities
BLF3	I believe that putting my data through cloud services is safe and enhances mobility
BLF4	Web 2.0 services may slow down my work load and research progress
BLF5	Managing time, procedures, reading, writing, reviewing, and daily events are effectively done without using web 2.0 services
CLB5	I intend to communicate only through e-mail in scientific research, due to the fact that research is an individual contribution

Table 6 Final measured items (items used)

Factor	Item
TSK1	Web 2.0 services may hinder my tasks in everyday research activities
TSK2	When I need to effectively discuss something related to my research with colleagues, web 2.0 services are not the right solution
TIM1	Web 2.0 services are not helpful services in situations when information is needed on the same day
TIM2	It is a waste of time to use web 2.0 services to establish communication or collaboration with other colleagues in the context of doing research
SUP1	The uses of web 2.0 services are useful for my research
SUP2	I enjoy using web 2.0 services in editing, commenting on, and reading a piece of research
S3	Using web 2.0 services may help a lot to inform me about important scientific events
CON1	My institute provides a proper knowledge management system/web 2.0 services (e.g. website) for improving communication and collaboration among doctoral students
CON2	Officially, Wiki is used as a platform for group activities and collaborative work reports in my research group
INC1	Creating a personal profile in web 2.0 services would enhance my reputation
INC2	Using web 2.0 services in research helps me to satisfy my interests in my scientific area
INC3	Web 2.0 services facilitate the presentation of myself and marketing my research
ETK1	I trust sharing my data through web 2.0 services
ACT1	I usually engage in one or more online scientific discussions
ACT2	Sharing files, links, videos, or photos with colleagues is one of my daily uses of web 2.0 services
ACT3	Peer review of scientific work via web 2.0 services is one of my usual online activities
ACT4	Commenting and writing in one or more scientific online forums, weblogs, or wikis is also one of my daily/weekly activities
BLF1	I believe that using web 2.0 services has become one of my everyday research routines
BLF2	I would say, to enhance academic collaboration, you should use web 2.0 services
CL1	I intend to engage and involve myself in a community of practice by using web 2.0 services
CLB2	I intend to share my reading, writing, review, and resources with other colleagues when it is mediated by web 2.0 services
CLB3	I intend to coordinate and work together more when this coordination is facilitated by web 2.0 services
CLB4	Willingness to communicate and collaborate in research with other disciplines could be enlarged by using web 2.0 services

References

- Birnholtz, J.P.: When do researchers collaborate? Toward a model of collaboration propensity in science and engineering research. Dissertation submitted at the University of Michigan (2005)
- DeAndrea, D.C.: Participatory social media and the evaluation of online behavior. *Human Commun. Res.* **38**(04), 510–528 (2012)
- Eyman, D., Sheffied, S., DeVoss, D.N.: Developing sustainable research networks in graduate education. *Comput. Compos.* **26**, 19–57 (2009)
- Forschung & Lehre: Promotion in Deutschland: Daten und Fakten, *Forschung & Lehre*, vol. 8 p. 631. http://www.forschung-und-lehre.de/wordpress/Archiv/2012/ful_08-2012.pdf (2012). Accessed 14 Apr 2015
- Frewox: Large moray eel and grouper hunting together in the Red Sea. <http://www.youtube.com/watch?v=ukiw4wYwqPg> (2010). Accessed 27 Oct 2012
- Gillet, D., Helou, S., Joubert, M., Sutherland, R.: Science 2.0: supporting a doctoral community of practice in technology enhanced learning using social software. *Science 2.0 for TEL Workshop* (2009)
- Illeris, K.: Learning in working life. Roskilde University Press, Frederiksberg (2004)
- Jirocka, M., Procter, R., Rodden, T., Bowker, G.C.: Special issue: collaboration in e-research. *Comput. Support. Coop. Work* **15**, 251–255 (2006)
- Kahnwald, N., Albrecht, S., Herbst, S., Köhler, T.: Informelles Lernen Studierender mit Social Software unterstützen. *Strategische Empfehlungen für Hochschulen; Reihe: Medien in der Wissenschaft*. Waxmann, Münster (2015)
- Köhler, T.: Sozialpsychologie der Kommunikation im Internet. Erkenntnisstand und Forschungsbedarf; Gardez! Verlag. [CD-ROM], Mainz (1997)
- Lahenius, K.: Communities of practice supporting doctoral studies. *Int. J. Manag. Educ.* **10**, 29–38 (2010)
- Lam, A.: What motivates academic scientists to engage in research commercialization: ‘Gold’, ‘ribbon’ or ‘puzzle’? *Res. Policy* **40**(10), 1354–1368 (2011)
- Lattemann, C., Köhler, T.: Trust or control-governance concepts for virtual organizations. In: *Frontiers of e-Business Research*, vol. 3 (2005)
- Meyer, A.M., McNeal, L.: Academics online: their interests and foibles. *Internet High. Educ.* **14**, 113–120 (2011)
- Mohamed, B.: Definition of a system for the management of project-based E-learning; Doctoral Dissertation, Technische Universität, Dresden (2011)
- Mohamed, B., Köhler, T., Mabed, M.: Designing e-research: A framework for researcher’s social online knowledge; In: Köhler, T., Kahnwald, N.: *Online Communities: Enterprise Networks, Open Education and Global Communication*. Proceedings of the GeNeMe 2013. TUDPress, Dresden (2013)
- Mohamed, B., Pscheidla, D., Köhler, T.: The Fish Model: a conceptual framework for enhancing e-research collaboration. In: Martí, D., Gómez, L., López, A. (eds.) *EDULEARN13 Proceedings. 5th International Conference on Education and New Learning Technologies*. IATED. Barcelona (2013)
- Mutula, S.M.: Ethics and trust in digital scholarship. *Electron. Libr.* **29**(2), 261–276 (2010)
- Patel, H., Pettitt, M., Wilson, J.R.: Factors of collaborative working: a framework for a collaboration model. *Applied Ergonomica* **43**, 1–26 (2011)
- Peggy, L.M., Borkowski, N.A.: The assessment of doctoral education: emerging criteria and new models for improving outcomes, library and information science research **29** (2007)
- Pidd, H.: German defence minister resigns in PhD plagiarism row. *The Guardian*. <http://www.guardian.co.uk/world/2011/mar/01/german-defence-minister-resigns-plagiarism> (2011). Accessed 1 Mar 2011
- Pscheidla, D., Albrecht, S., Herbst, S., Minet, C., Köhler, T.: Nutzung von Social Media und onlinebasierten Anwendungen in der Wissenschaft. Erste Ergebnisse des Science 2.0-Survey 2013 des Leibniz-Forschungsverbunds “Science 2.0”. TUD Press, Dresden (2014)

- Pschaida, D., Köhler, T., Mohamed, B.: What's your favorite online research tool? Use of and attitude towards Web 2.0 applications among scientists in different academic disciplines. In: Marsden, C., Tassiulas, L. (eds.) Proceedings of the 1st International Conference on Internet Science. Sigma Orionis, Brussels (2013)
- Reebs, S.-G.: Cooperation in Fishes. Moncton, Canada. <http://www.howfishbehave.ca/pdf/cooperation.pdf> (2011). Accessed 26 Oct 2012
- Ringle, C.M., Wende, S., Will, A.: SmartPLS. Release 2.0 (beta). University of Hamburg, Hamburg (2005)
- Saxony State, Statistical Branch: Abschlussprüfungen an den Hochschulen im Freistaat Sachsen. http://www.statistik.sachsen.de/download/100_Berichte-B/B_III_3_j09.pdf (2009). Accessed 24 Oct 2012
- Serenko, A.: A model of user adoption of interface agents for email notification. *Interact. Comput.* **20**, 461–472 (2008)
- Tannen, D.: The Argument Culture. <http://www.petracaci.com/argpp1.pdf> (2006). Accessed 24 June 2012
- Terrell, S.R., Snyder, M.M., Dringus, L.P.: The development, validation, and application of the doctoral student connectedness scale. *Internet High. Educ.* **12**, 112–116 (2009)
- Veletsianos, G., Kimmons, R.: Scholars and faculty members' lived experiences in online social networks. *Internet Hig. Educ.* **16**, 43–50 (2012)
- Wang, M.-L.: Scholarly journal use and reading behavior of social scientists in Taiwan. *Int. Inf. Libr. Rev.* **42**, 269–281 (2010)
- Wolters, M., Schmiedel, S.: Promovierende in Deutschland 2010. Statistisches Bundesamt, Wiesbaden retrieved from: <https://www.destatis.de/DE/Publikationen/Thematische/BildungForschungKultur/Hochschulen/Promovierende5213104109004.html> (2012). Accessed 14 Apr 2015
- Ynalvez, M.A., Shrum, W.M.: W.M. Professional networks, scientific collaboration, and publication productivity in resource-constrained research institutions in a developing country. *Res. Policy* **40**, 204–216 (2011)
- Zaman, M.: Doctoral programs in the age of research2.0; In: Anandarajan, M.: e-Research Collaboration: Theory, Techniques and challenges. Springer-Verlag GmbH & Co., Berlin and Heidelberg (2010)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



The Use of Digital Tools in Scholarly Activities. Empirical Findings on the State of Digitization of Science in Germany, Focusing on Saxony



Steffen Albrecht, Claudia Minet, Sabrina Herbst, Daniela Pscheida, and Thomas Köhler

Abstract Scholars are only beginning to understand what digitization means for their work, that is, the conduct of science. Taking a broad perspective on e-science, this paper provides empirical insights about two important aspects of the digitization of science, namely the use of digital tools in scholarly activities and scholars' perceptions of the change such use entails. The results of a German-wide survey of scholars and supplementary qualitative interviews in the years 2012 and 2013 show that the majority of scholars have adopted digital tools and that scholarly practice is affected profoundly by the use of such tools. This does not apply to web 2.0 tools, which remain a niche medium for some scholars. Small but significant differences exist between disciplines, and decisions about individual tool use are utilitarian. Further research is needed to assess the changes from a longitudinal perspective.

Keywords E-science · Digitization · Scholarly practice · Survey results

1 E-Science, Cyberscience, Science 2.0: The Digitization of Science Is on the Move

Ever since Galileo's successful use of the telescope, scientists have relied on new tools in their scholarly practice (Hankins and Silverman 1995). The advent of computer

C. Minet

Hochschule Mittweida—University of Applied Sciences, Mittweida, Germany
e-mail: Minet@hs-mittweida.de

S. Herbst · D. Pscheida · T. Köhler

Media Center, Technische Universität Dresden, Dresden, Germany
e-mail: Sabrina.herbst@mailbox.tu-dresden.de

T. Köhler

e-mail: Thomas.koehler@tu-dresden.de

S. Albrecht (✉)

Berlin, Germany

e-mail: Steffen.albrecht@berlin.de

technology and digital networks was no exception, impacting not only the communication of research, but also the production of new knowledge. The World Wide Web with its network of hypertexts and its gradual change to web 2.0 with the additional manifestation of online social networks reinforce this impact and generate more potential. Although this cooperative process is not new (Bijker and Law 1992; Mayntz 1993), we are still only beginning to understand the changes that digitization entails for science. This paper furthers our understanding by providing empirical insights from an online survey of German scholars related to two specific aspects of the digitization of science, namely the use of digital tools in research, teaching and other scholarly activities, and scholars' perceptions of the change that such use brings.

Several terms have been proposed to apprehend how science is influenced by networked computer technologies. In 1999, the term e-science was introduced by John Taylor, General Director of the Research Council at the Office of Science and Technology of the UK. Taylor realized that new technological infrastructures were needed to foster global cooperation and data-intensive research in science. In other words, “e-science is not a new scientific discipline; rather, the e-science infrastructure developed [...] should allow scientists to do faster, better or different research” (Hey and Trefethen 2005: 818). Jim Gray specified what such “different research” could look like. He identified a “fourth paradigm” of scientific inquiry, “data-intensive science” that is characterized by the use of massive amounts of data to generate new theoretical models (Gray 2009: xix).

Michael Nentwich (2003) took a more holistic view of “cyberscience,” including how academic work is organized, how it functions, and what its products are. While emphasizing its novelty compared to “traditional science,” he technically assesses the digitization of science. In a recent update with René König, Nentwich used the term “cyberscience 2.0” to acknowledge the emergence of web 2.0 and its relevance to scholarly communication (Nentwich and König 2012). “science 2.0” is another term that emphasizes the importance of web 2.0 in facilitating openness and collaboration, focusing on online communication tools such as weblogs or wikis that open up science communication to external audiences (Waldrop 2008).¹

Despite their nuances, all these terms are more similar than different. A broad notion is best suited to address the diverse issues involved in the digitization of science. Here and in the e-Science Research Network Saxony (www.escience-sachsen.de), we use the term e-science to comprise science, social science, and humanities disciplines, not only in research and collaboration, but also in teaching and science communication. In terms of technology, e-science comprises digital tools used in scholarly work that go beyond the individual computer and represent digital media or online-based, networked software systems.

Taking a broad perspective on digitization means normatively assessing this process without bias. This involves considering not only the changes in technology, but also changes induced by the social environment. Soon before e-science got onto

¹Other notions of the digitization of science, include “digital scholarship” (Weller 2011) and “digital science” (European Commission 2013).

the agenda, several authors recognized fundamental changes in how science was generally understood. “In response to the challenges of policy issues of risk and the environment, a new type of science – ‘post normal’ – is emerging,” wrote Silvio Funtowicz and Jerome Ravetz (1993: 739). In a similar vein, Michael Gibbons and colleagues (1994) observed the emergence of what they called “mode 2” science, which was transdisciplinary and involved stakeholders from outside the scientific community. Both diagnoses overlap in noting an increasing external influence on science from politics, the economy, and civil society. The interrelations between technological and social change are not the main focus here, but developments within science such as the role of web 2.0 in opening up the research process might be part of a broader social change, in which technology plays only a moderating role.

The aim of this paper is to provide empirical observations of changes in scientific practice in relation to technological change, focusing on media use in science. Our approach is based on two fundamental assumptions. First, we find that considerable attention is devoted to the potential and affordances of digital technologies, but much less notice is taken of what scholars actually do with these technologies in their day-to-day practices, including potential non-adoption and refusal to use them (cf. Barassi and Trer 2012: 1282). Second, while there are a number of empirical case studies of scholars’ use of technology in specific fields, we think a broad view on all aspects of scholarly practice is necessary to identify the changes, before the nature of such change in specific areas can be analyzed.

2 The Empirical Question: Is Digitization Really on the Move?

The empirical perspective of this paper goes beyond the rhetoric and euphoric expectations of some e-science discourse. We ask three questions. To what extent do scholars use digital media and online tools in their day-to-day academic activities? What kind of new practices emerge from such use? How do such changes in the conduct of science contribute to the bigger structural changes?

Technological innovation always takes place in form of co-evolution of engineering and social domains (Köhler 1998). Adoption theorists have pointed out that the adoption process is not just a matter of time, but also of individual differences, system characteristics, social influence, and facilitating conditions (Venkatesh and Bala 2008). Against this background, we can assume that the adoption of digital tools is not as straightforward a process as is depicted by some of the theoretical accounts discussed above: it is ongoing and has to be observed empirically to determine its state and direction. Our paper adds to the small, but growing empirical literature about the impact of using digital tools in science.

Previous research has shown that investigating scholars’ use of digital tools poses methodological problems. There are a number of different approaches, all with specific merits and pitfalls. We can broadly distinguish a qualitative orientation with

a focus on in-depth analysis of a limited number of cases, often based on stakeholder interviews or case studies (see, e.g., Currier 2011; RIN/NESTA 2010; Bullinger et al. 2010), and a quantitative orientation with a focus on assessing the whole field, often based on standardized surveys. As our aim is to provide a holistic and realistic assessment of the state of adoption, we mainly review previous quantitative research.

Lattemann et al. (2010), ZBW (2011), Donk (2012), and Pscheid and Köhler (2013) all address a limited target group (principal investigators in funded research projects, economics researchers and students, researchers at one specific university, and scholars at universities in Saxony, respectively). Thus, none of this research is particularly helpful in terms of either methodology or results. In an early study of 1477 UK researchers, Procter et al. (2010) found that 60% used a web 2.0 tool (blogging, commenting, sharing resources, or contributing to wikis) in their scholarly activities, but only 13% did so frequently. The authors consider this figure “rather low” and observe that frequent users are most likely to be computer scientists or mathematicians, engineers, or scholars in the arts and humanities. Ponte and Simon (2011) also focus on web 2.0 use, but based on a self-selected sample of 345 persons from across Europe. They report the use of wikis and blogs by about 40% of respondents, academic social networking sites by 35%, and microblogging by 18% of researchers. Results for specific groups are not presented.

Bader et al. (2012) analyzed 1053 responses to an online survey of scholars at German universities. They found that communication tools such as e-mail (94%), mailing lists (24%), and Skype (21%) were widely adopted, web 2.0 tools like blogs or research portals were much less used (6% use wikis, 5% use research portals or social networking sites, 4% use academic blogs, and 2% use Twitter). The tools used varied greatly by discipline: wikis were mostly used in science and engineering, whereas mailing lists and blogs were more popular in humanities and social sciences, especially in law, and research portals were favored by social scientists. In general, the authors consider German researchers to be at an “early stage” of adopting digital communication tools.

Despite the small number of studies of sufficient scope and methodological quality, these results raise doubts about the predicted impacts of digital tools on science. In the best case, adoption is too early to have had a significant impact. In the worst case, apart from some small groups, scholars are not tempted to actually use digital tools in their work. The existing research shows that digital tool adoption in scholarly activities is low (apart from very popular tools such as search engines), that web 2.0 tools are much less likely to be used than more conventional ones, and that disciplines seem to play a role in the choice of tool.

3 Hypotheses, Data, and Methods

To remedy the obvious lack of comprehensive, quantitative research, our paper seeks to empirically assess the state of adoption of digital tools by scholars and their impacts on the basis of new data from the Science 2.0 Survey 2013 (Pscheid et al. 2014).

Based on our above analysis of previous research, we assume that to have an impact on scientific activity, digital tools have to be used for scholarly purposes in the first place. This leads us to the following hypotheses and research questions.

Hypotheses. Our first two hypotheses concern the extent to which scholars use digital tools in their activities.

H1: The adoption of digital tools in scholarly activities is still in an early phase, with diffusion levels (Rogers 1995) below 50% of the population (“early majority”).

H2: Web 2.0 tools like weblogs, wikis or microblogging are used by a minority of scholars for professional activities, with diffusion levels below 16% (“early adoption”).

Our third hypothesis concerns the differences between disciplines and is formulated as research question, since previous research is inconclusive.

RQ1: Do scholars in different disciplines use digital tools differently?

Finally, we are interested in how digital tool use impacts on the conduct of science, leading to another research question.

RQ2: What changes in the conduct of science as a result of digital tool use do scholars perceive?

Data. We collected data in two related steps. First, an online survey of 778 scholars at German universities was conducted in autumn 2013, addressing questions such as scholars’ use of 17 different tools and services, their academic and sociodemographic background, their motives for and attitudes to using digital tools (see Pscheida et al. 2014).² Although quota sampling of universities was applied in the recruitment procedure, the sample shows some deviations from the population with regard to gender (women are slightly overrepresented), professional status (professors are overrepresented, research assistants or “WHK” underrepresented), discipline (with medicine strongly underrepresented, humanities, mathematics, and natural sciences slightly overrepresented), location, and type of university. While the latter two could be adjusted by weighting, the other deviations should be kept in mind in interpreting the results. In addition, all scholars at universities in the German federal state of Saxony were invited to participate in the same survey, with 442 questionnaires being submitted. The Saxony sample shows similar patterns of deviations from the population, except that with regard to disciplines, engineers and mathematicians/natural scientists are strongly overrepresented, whereas medicine and the fine arts are underrepresented.

The quantitative survey was supplemented in the first half of 2013 by 19 interviews with scholars in Saxony, chosen to map the various disciplines and status groups. The semi-structured interviews focused on the scholars’ perception of the use of digital tools and of the changes this entails. Due to the variety of scholarly practices and lack of knowledge about the precise impact of digital tools on them, qualitative interviews were chosen to address our second research question.

Methods. The hypotheses and research questions were statistically analyzed, comparing the Saxony and German-wide samples. Adoption was measured by asking scholars “to what extent do you use the following?” followed by a list of 17 different

²The data set of the Science 2.0 Survey 2013 is open access: see www.escience-sachsen.de.

online tools. Answers were categorized by frequency of use. In the analysis, only uses for scholarly work (research, teaching, research administration, and science communication) were taken into account. Based on Schmidt's (2007) definition, social networking sites, wikis, video/photo community portals, weblogs, microblogs, and social bookmarking services are regarded as "social software" or "web 2.0 tools," as they constitute social or hypertextual relationships of (at least partially) public character (Schmidt 2007: 32). The disciplines were categorized based on the definition of the German Federal Bureau of Statistics (2012) into arts and humanities; (natural) sciences (including mathematics); engineering and social sciences (including law and economics). Finally, the qualitative interviews were transcribed and anonymized, and qualitative content analysis methods were applied to the responses.

4 Results

4.1 General Level of Adoption of Digital Tools in Scholarly Activities

Scholarly activity at German universities in 2013 was affected considerably by the use of digital tools. Of all 17 tools the survey asked about, ten were used by more than 50% of all respondents in a professional context (with two others by 49%, see Fig. 1). Only general-purpose social networking sites, online editors like Etherpads

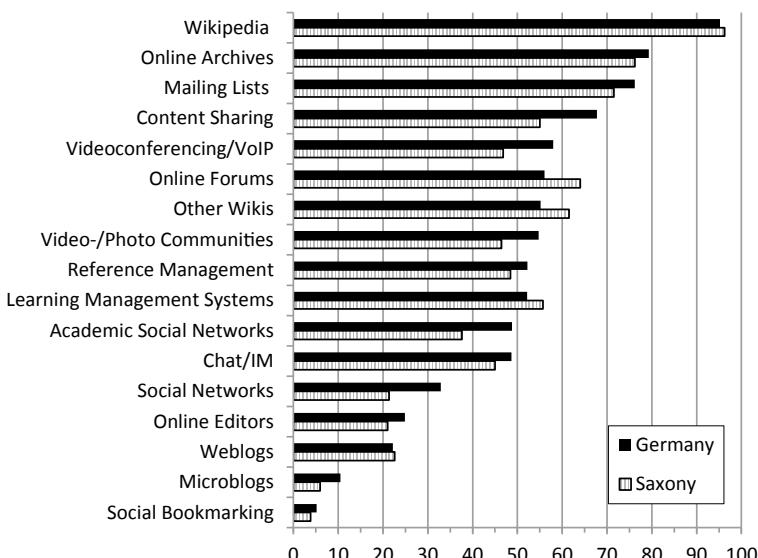


Fig. 1 Level of adoption (in %) of digital tools in scholarly use in Germany and Saxony

or Google Docs, weblogs, microblogs like Twitter, and social bookmarking services are used by less than about half of all respondents. Wikipedia is the tool with the broadest diffusion in academia, with 95% of respondents reporting to have used Wikipedia in their scholarly work. Comparably extensive more than three quarters of the respondents use online archives like Arxiv.org and mailing lists.

The pattern for scholars in Saxony is quite similar to the national one. Wikipedia, online archives, and mailing lists are the tools with the highest level of adoption in the context of scholarly work. Social networking sites, online editors, weblogs, microblogs, and social bookmarking services are used by less than half of the respondents, including professional social networking sites like Xing or Academia.edu. The general level of adoption of digital tools in Saxony is a bit lower than in Germany as a whole. The reverse is true for online forums (64% in Saxony, 56% in Germany) and wikis other than Wikipedia (62% in Saxony, 55% in Germany). Of course, such seems still to be a contradictory observation as many scientists recommend their students not to use digital tools like Wikipedia due to the “non-scientific” nature. So is especially valuable exploring in more detail practices of adoption among scientists.

In some cases, digital tools may be used more for general than work-related purposes (see Table 1). Content sharing and cloud services, video conferencing and VoIP (“online telephone”) services, online forums, video/photo communities,

Table 1 General and work-related use (in %) of digital tools in Germany and Saxony. Note the high level of general use

Digital tool	Germany		Saxony	
	General	Work	General	Work
Social networking sites	57.9	32.9	51.1	21.0
Academic networking sites	52.7	48.8	41.0	37.6
VoIP	71.7	58.0	64.3	46.2
Microblogs	15.1	10.5	5.9	1.6
Weblogs	29.2	22.2	28.3	22.6
Wikipedia	98.9	95.2	100	96.2
Other wikis	56.5	55.1	63.6	61.5
Content sharing/cloud server	73.5	67.7	64.3	54.5
Online editors	26.7	24.9	24.4	21.0
Online forums	65.3	56.0	74.0	64.0
Mailing lists	77.4	76.2	74.0	71.5
Chat/IM	69.1	48.7	69.2	44.3
Online archives	79.8	79.3	76.7	76.2
Reference management	52.2	52.2	49.1	48.4
Social bookmarking services	5.9	5.2	4.3	3.8
Video/photo communities	80.4	54.8	78.3	45.7
Learning management systems	52.3	52.2	56.1	55.7

chat/instant messaging (IM), social networking sites, and microblogs all show a significantly higher level of personal than professional use. Where data for comparison exists, the use of digital tools is more widespread among scholars than in the German population in general (cf. data from the ARD/ZDF Online Study 2013, van Eimeren and Frees 2013).

With regard to our first hypothesis, we can thus infer that the adoption of digital tools in scholarly activities has left the phase of early adoption and has reached a more mature state with more widespread use. Although a considerable number of scholars do not use certain digital tools, and some tools have not reached broad adoption, the majority of tools our survey asked about are used by more than 50% of respondents.

4.2 Use of Web 2.0 Tools Among Scholars

The situation is different for web 2.0 tools such as wikis, blogs, social networking sites, social bookmarking services, and video/photo communities. While about half of respondents use wikis, video/photo communities and academic social networking sites in work-related contexts, only between 5 and 32% of scholars use general-purpose social networking sites, weblogs, and microblogs as well as social bookmarking services for work. Considering the broad adoption of digital tools in general and the length of time that web 2.0 tools have been in use, the latter have to be considered a niche product with regard to scholarly use. The figures for Saxony are comparable, but generally lower than for the national level (except for wikis, see above).

However, with regard to our second hypothesis, web 2.0 tools have a higher adoption level than the 16% “early adoption” rate, at both the national and the Saxony level, with the exception of microblogs and social bookmarking services. At the same time, only a minority of scholars use web 2.0 tools that have not been designed specifically for academics. Given that these tools have been in use for a long time and are well known among scholars (except for social bookmarking services, which about 50% of respondents said they didn’t know about), we have to conclude that web 2.0 tools have only reached specific groups of scholars (cf. Pscheida et al. 2014: 18). It seems to be difficult for most scholars to find useful applications for these tools.

From the results of the survey, we can more generally infer that tools are adopted when a specific use is found for them. Most of the tools with high levels of adoption are specialized for one or more areas of scholarly work. Most respondents said that the digital tools they use are practical or make their work easier and faster (Pscheida et al. 2014: 24f.), indicating the prevalence of utilitarian motivation. This was not equally the case for all tools. General-purpose social networking sites and microblogs show a different pattern of use: both tools are used twice as often in a personal than a professional capacity, that is, utilitarian motivation was less important. The two main reasons for not using general-purpose social networking sites in a scholarly context are disagreement with the terms of use (indicated by 24% of those researchers who

do not use them for work) and personal use (indicated by 18%). For microblogs, the most salient obstacle is the lack of additional benefit (indicated by 56% of those researchers who don't use microblogs for work).

Besides the prevalence of pragmatic reasons, researchers who use web 2.0 tools (like academic network sites, weblogs, or microblogs) also mentioned an interest in new technologies or that these tools help to boost their reputation. This shows awareness of the social and hedonistic (as one could say) affordances of web 2.0 tools.

4.3 Disciplinary Differences

Our third hypothesis was about the differences between disciplines in scholars' adoption of digital tools. The results of the survey indicate a more nuanced picture than previous studies have drawn. Cross-tabulation and computation of Cramer's V (a bivariate measure of association) for the German study show small but significant differences in professional use of Wikipedia, wikis, online editors, mailing lists, online archives, reference management systems, social bookmarking services, and video/photo communities (see Table 2). In all other cases, no difference between disciplines in the use of digital tools for scholarly purposes is found.

Digital tools are most highly adopted in the (natural) sciences and in arts and humanities, whereas engineering and social sciences show lesser degrees of adoption. However, engineering scientists use wikis and video/photo communities quite heavily, and social scientists use mailing lists and online archives to a similar degree as the (natural) scientists.

For Saxony, significant differences between the disciplines are more frequent and related to other tools than in the German-wide study. Social networking sites, academic networking sites, VoIP, microblogs, weblogs, content sharing services, chat/instant messaging services, reference management systems, and learning management systems all show small but significant differences between disciplines (see Table 3). Social scientists use tools most across all categories, followed by scholars in arts and humanities. The only tools which are used more extensively by natural scientists and engineers are Wikipedia and other wikis, but these findings are not statistically significant.

The differences between the Saxony and German-wide results are striking. They might be explained by the special disciplinary structure of universities in Saxony, which have a strong emphasis on natural sciences and engineering. For Germany as a whole, such differences might exist, but are leveled off due to the mix of academic cultures and institutional structures across the various federal states. However, the differences are generally small, with low values of Cramer's V, so we can conclude with regard to our first research question that there are only small differences between the disciplines, highly dependent on the disciplinary context in which each tool is used.

Table 2 Professional use (in %) of online tools by scholars at German universities in 2013 across the most relevant disciplines: arts and humanities; (natural) sciences; engineering and social sciences

Digital tool	General n=778	Humanities n=225	Natural sciences n=243	Engineering n=116	Social sciences n=133 ^a	Cramer's V ^b
Social networking sites	32.9	37.6	32.2	24.1	35.1	—
Academic networking sites	48.8	48.7	53.5	44.0	53.8	—
VoIP	58.0	57.8	60.7	57.8	56.4	—
Microblogs	10.5	12.0	10.7	8.6	11.9	—
Weblogs	22.2	25.8	23.0	16.2	23.3	—
Wikipedia	95.1	93.8	98.8	94.9	90.2	.112
Other wikis	55.1	52.2	68.7	56.0	44.4	.114
Content sharing/cloud server	67.7	69.9	70.8	59.5	69.9	—
Online editors	24.8	28.0	26.7	22.2	21.2	.092
Online forums	56.0	58.0	59.3	65.5	45.1	—
Mailing lists	76.2	81.0	76.9	63.8	75.2	.093
Chat/IM	48.7	51.8	54.5	40.5	49.6	—
Online archives	79.3	86.3	78.2	67.5	78.9	.110
Reference management	52.2	53.3	63.6	33.3	43.6	.158
Social bookmarking services	5.3	4.4	9.1	2.6	3.8	.100
Video/photo communities	54.8	64.2	48.6	57.8	43.3	.116
Learning management systems	52.2	56.9	44.7	51.3	59.0	—

^aincluding law and economics^bsignificance $\alpha < .05$

— indicates that no significant correlation is observed

If discipline does not explain differences in the use of digital tools, how else might we explain them? Some indications can be found in the 19 semi-structured interviews that were conducted to supplement the quantitative investigation. Content analysis methods according to Mayring (2000) were used to analyze these. In a first step, categories of analysis were generated based on the interview guideline. These categories were then tested against the empirical material, and continuously

Table 3 Professional use (in %) of online tools by scholars at universities in Saxony in 2013 across the most relevant disciplines: arts and humanities; (natural) sciences; engineering and social sciences

Digital tool	General n=442	Humanities n=49	Natural sciences n=124	Engineering n=191	Social sciences n=47 ^a	Cramer's V ^b
Social networking sites	21.3	28.6	21.8	13.6	34.0	.165
Academic networking sites	37.6	32.7	37.9	31.9	66.0	.165
VoIP	46.8	42.9	53.2	39.3	63.8	.145
Microblogs	5.9	6.1	4.8	1.6	21.3	.197
Weblogs	24.0	30.6	25.0	16.2	38.3	.133
Wikipedia	96.2	91.8	98.4	96.3	93.6	—
Other wikis	61.5	59.2	64.5	61.8	61.7	—
Content sharing/cloud server	55.0	71.4	50.8	47.1	72.3	.146
Online editors	27.8	20.4	25.8	15.7	31.9	—
Online forums	64.0	49.0	67.7	64.9	68.1	—
Mailing lists	71.5	77.6	75.8	64.4	85.1	—
Chat/IM	45.0	42.9	54.8	36.1	55.3	.140
Online archives	76.2	71.4	77.4	74.9	85.1	—
Reference management	48.4	51.0	51.6	40.8	66.0	.138
Social bookmarking services	3.8	4.1	5.6	3.7	2.1	—
Video/photo communities	46.4	53.1	45.2	41.9	59.6	—
Learning management systems	55.7	77.6	53.2	49.2	80.9	.175

^aincluding law and economics

^bsignificance $\alpha < .05$

— indicates that no significant correlation is observed

revised and amended with sub-categories during analysis. In a third step, relations and causalities between categories and sub-categories of analysis were carved out.

The results do not point to disciplinary differences, but instead to the influence of the tangible working practices in which scholars are involved. These can be collaborations, such as in projects with many partners (interview 8), working groups (interviews 1, 7) or institutions (interview 7), but also institutional contexts, such as when “interdepartmental” wikis are created (interview 8) or the institutional website is used for information sharing, because “those who are concerned and who will look at the information [are] mainly limited to the institute” (interview 19).

Another important determinant for the use of a digital tool is its quality and suitability for specific working contexts. Digital tools are expected to make working processes more efficient: “the biggest obstacle and a huge inefficiency is how we communicate data. We copy, we process data again and again” (interview 3). Wikis, for example, are used to facilitate collaborative work: “to manage data, I do not have to send e-mails around where nobody knows what the current state is” (interview 8). Similar motivations underlie the use of cloud services like Dropbox (interview 5) or instant messaging clients and VoIP services such as Skype and ICQ (interview 8). The use of e-mail for collaborative work is considered rather inefficient (interviews 3, 5, 8).

Wikis can be used as “encyclopedias,” to provide information in a structured and clearly arranged way, “where you can collect things that you maybe will have to look up in future” (interview 19). Wikis are repositories “where all kinds of information are collected” (interview 19); “just to preserve collected knowledge that you can look up again” (interview 7); “where knowledge for all is provided” (interview 1); “to upload files in the current state, where it can then be downloaded” (interview 8). This also refers to the exchange of administrative information, such as in managing technical infrastructure (interview 9) or as an organizational manual (interviews 9, 15).

However, once data protection becomes an issue, web-based tools and services are not used despite their efficiency savings: “I would have proposed Dropbox, but because of data protection requirements we cannot use it” (interview 11). Instead, local network servers are used to exchange data, especially in cases where the cooperation is limited to partners in the same institution: “to some extent we have this in our working group internally, using the university file system. There we have our account and there is our stuff, i.e. the programs, and everybody in our working groups who wants can use it” (interview 1). Yet, web-based applications are important “especially if you collaborate with external partners” (interview 8). From a qualitative perspective, too, the conclusion is that the requirements of collaborative work and the affordances of the technology have a stronger influence on scholars’ choice of digital tools than professional affiliations.

4.4 *Changing Scholarly Practices*

The above analysis sheds some light on an important prerequisite for any changes in the conduct of science induced by digitization, namely the actual use of digital

tools in scholarly activity. But this is just a necessary, not a sufficient, condition for change. The semi-structured interviews indicate what kind of changes scholars in Saxony perceive. Of course, such perceptions might not give an accurate account of the situation, but until better data is available (e.g., from long-term observations of scholarly practices), qualitative interpretation of perceptions from actors within the field with a variety of perspectives gives a good approximation.

Our interview partners indicated a number of changes in their work practices which they did not see as related to technology use, but rather to a changing social environment in science. The importance of collaborative work is seen as growing. Short-term research projects can require the use of certain digital tools (interview 8) or new competencies, such as writing research proposals (interview 7). Work biographies are seen as becoming more flexible. Temporary retirement affects the way of working and the adoption of new technologies: “so I did not slowly get used to [web-based tools], I knew work without them and when I was re-entering, I had to use and become acquainted with the different tools” (interview 1). This flexibility also includes geographical mobility and increasing independence from local contexts: the use of Skype “has actually naturalized through stays abroad, only to stay in touch” (interview 14).

Interviewees mentioned several changes in their personal way of working that they related to technological conditions. The entire process of scientific enquiry, including literature research, has tremendously accelerated: “you simply find something immediately rather than writing letters to ask: ‘what did you do there actually?’” (interview 19). Before using digital archives, “we had to [...] ask for interlibrary loan literature and had to wait” (interview 16). Technology is explicitly mentioned as an attractive agent of change: “so I still remember card indexes in libraries and of course when the online catalog was there, then you liked to use it” (interview 14). This also relates to the management of literature: “I think the trigger was that the university offered Refworks and that we got an account for the group” (interview 1).

Communication processes seem to be particularly affected: “if I have a meeting with someone today, I just search for her or him on the Internet before and look up who it is. If I’m lucky, I have a small CV or at least I see what she or he does” (interview 13). Communication is increasingly shifting into virtual spaces (interview 17): “in times of my diploma one rather met personally, [...] so if calling on the telephone did not work, then one rather met personally somehow” (interview 7). What is more, the way information is stored and made available has changed: “I have scarcely printed or written documents, [...] all of my documents are digitized. Either as a PDF or HTML page or in another format, like video or other scripts or programs” (interview 2). This in turn influences the access to information: “in the past, I can still remember that I used a usual lexicon from the bookshelf, which I not so long ago just sold because I have not been using it anymore and it stood around useless” (interview 9). Again, the affordances of new technologies are described as attractive: “I notice that I still prefer printed paper, but this changes step by step and in ever more cases, I do not print the reports I read, but rather read them on the screen somewhere and if I have the opportunity also highlight sections as it is possible with various apps on the iPad,

then for me this actually replaces printed and nicely annotated reports because I thus have the same opportunities to work" (interview 18).

Advancements in data infrastructure, including the availability of faster and more efficient Internet connections or better computing capacities, primarily affect information sharing and data analysis, "just because it was somehow difficult to load ten megabytes from the Internet with the first emerging DSL connections" (interview 15), and "if one then changed from modem to ISDN and DSL, then you increasingly used it, it went faster" (interview 14). The ubiquity of computing and network power eases the work process "because you don't have to rack your brain, should I resize photos in the dataset or not, instead you just send it" (interview 12). Parts of data analysis are replaced by automated processes: "30 years ago or maybe more, you went with stacks of punch cards to the computing center and tried to compute a t-test or something similar, and now you just have to push the button" (interview 5, cf. interview 19). The availability of portable devices, "that we have just passed on to equip all staff with laptops, [i.e.] no location-bounded work on the computer we sit in front of anymore" (interview 15), supports highly flexible working practices.

Finally, the attitude of researchers toward new technologies, their openness and curiosity to try something new also affect their working practices: "then I had a telephone bill of about 80 marks which was very high for a student, just because I intensively explored the Internet" (interview 16). Or, as another interviewee said: "whenever a new technology emerges, I deal with it and watch to see if it makes sense to use it" (interview 3). Last but not least, cost-benefit considerations also play an important role: "If I have the feeling that there is something that helps me on [...] then I try it" (interview 19, cf. interview 16).

Coming back to our second research question about perceived changes in the conduct of science, the qualitative interviews confirm the result from the survey that in science, digital tools are widely adopted. Scholars are not only using digital tools for work, but also perceive their work as being changed by these tools, partly even dramatically. The change is described as making research more efficient and faster, and this acceleration also affects communication and collaboration.

The precise nature of this change requires more thorough analysis. From the results presented here, it is clear that technology is just one of the driving forces underlying the change, and that the increasing collaboration and mobility of scholars is another important factor interacting with the use of technology. With regard to the motivations for decisions about technology use, both the qualitative and the quantitative analysis underscore the importance of a pragmatic, utilitarian orientation. The affordances of digital technologies and the institutional contexts appear less perceptible, but also relevant factors in determining which technologies are used and to what extent they affect scholarly work. Based on our study, more detailed research into the interplay of these factors can be designed and carried out.

5 Summary and Discussion

Our study empirically observes the digitization of science and its effects on scholarly practices. Starting from the individual use of digital tools by scholars as the most important element in the digitization process, we have measured the adoption of digital tools in scholarly work in Germany, focusing on Saxony. By critically extending previous work, our results show that the majority of scholars adopted such tools and that scholarly practice is affected profoundly by their use. We have also shown that this does not apply to all kinds of tools. Web 2.0 and its affordances for scholars might stimulate much debate, but a minority of scholars only uses tools such as weblogs and social networking sites. Presumably, a neglect of epistemological and technological sociological analyses of scientists' activities can be identified, which has to be overcome. This goes hand in hand with the need for a review with regard to digital science technologies, which has so far been reflected neither in the curricula for training nor in the self-image of the scientists.

Our survey indicates that there are small but significant differences in disciplinary adoption of digital tools. The arts and humanities show higher levels of adoption than engineering and the social sciences. However, the degree of use greatly depends on the tool in question. Similarly, the change which the tools induce varies greatly by scholarly activity. Our analysis of the qualitative interviews has confirmed that tools are chosen based on utilitarian motives, and given rise to new hypothesis about the interrelation between individual, technological, and systemic factors of change in the digitization of science.

In comparison with the discourse on e-science, cyberscience, and science 2.0, but also to the results of previous empirical studies, our results show that the digitization of science is indeed on the move in Germany. The level of adoption is higher than in previous studies, with many digital tools reaching broad professional diffusion. The full potential of e-science has yet to be exploited. Our interviews indicate that institutional cultures and the affordances of the technologies do not fit well enough to let these online applications evolve into widely used professional scholarly tools. Still there is need for further consideration, including individual competency development.

Our results certainly do not provide definite answers to the questions raised at the beginning of this paper. As well, one may observe different and perhaps contradictory patterns of adopting digital tools in science. The scope of our analysis is too limited to assess the digitization of science broadly. Thus, we deliberately chose to analyze tool use first, to gain as precise a measure of adoption as possible. More detailed analysis of the specific kind of tool use would merit attention, taking into account the institutional conditions of science or the affordances of digital tools. Moreover, the digitization of science is an ongoing process, which calls for a longitudinal perspective toward understanding the character of digitization. As stated in the introduction, there are still very few empirical studies on the digitization of science. Our aim was to contribute to a growing body of (empirical) research and we hope to have laid the foundations for future, longitudinal studies.

References

- Bader, A., Fritz, G., Gloning, T.: Digitale Wissenschaftskommunikation 2010–2011. Eine Online-Befragung. Gießener Elektronische Bibliothek, Gießen. <https://www.yumpu.com/de/document/read/6774710/digitale-wissenschaftskommunikation-2010-2011-eine-online> (2012). Accessed 11 March 2020
- Barassi, V., Trer, E.: Does Web 3.0 Come After Web 2.0? Deconstructing theoretical assumptions through practice. *New Media & Soc.* **14**(8), 1269–1285 (2012)
- Bijker, W.E., Law, J. (eds.): *Shaping Technology / Building Society: Studies in Sociotechnical Change*. MIT Press, Cambridge (1992)
- Bullinger, A.C., Hallerstede, S.H., Renken, U., Söldner, J.-H., Mösllein, K.M.: Towards Research Collaboration—a Taxonomy of Social Research Network Sites. *AMCIS 2010*. Lima, Peru (2010)
- Currier, S.: Open Science Project: Final Report. Centre for Research Communications. University of Nottingham, Nottingham (2011)
- Donk, A.: Ambivalenzen der Digitalisierung. Neue Kommunikations- und Medien-technologien in der Wissenschaft. MV Wissenschaft, Münster (2012)
- European Commission: Digital science in Horizon 2020. Concept Paper by DG Conect. <http://ec.europa.eu/digital-agenda/en/digital-science> (2013). Accessed 21 July 2013
- Funtowicz, S.O., Ravetz, J.R.: Science for the Post-normal Age. *Futures* **25**(7), 739–755 (1993)
- German Federal Bureau of Statistics: Bildung und Kultur. Studierende an Hochschulen – Fächer-systematik 2013. Fachserie 11, Reihe 4.1, WS 2012/2013. Federal Bureau of Statistics, Wiesbaden. https://www.destatis.de/DE/Methoden/Klassifikationen/BildungKultur/StudenterPruefungsstatistik.pdf?__blob=publicationFile (2012). Accessed 8 Mar 2020
- Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., Trow, M.: *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*. Sage, London (1994)
- Gray, J.: Jim Gray on eScience: A Transformed Scientific Method. Based on a presentation by Jim Gray at the NRC-CSTB Meeting in Mountain View, 11 January 2007, In: Hey, T., Tansley, S., Tolle, K. (eds.) *The Fourth Paradigm. Data-Intensive Scientific Discovery*, pp. xvii–xxxi. Microsoft Research, Redmond (2009). <http://research.microsoft.com/en-us/collaboration/fourth-paradigm/>. Accessed 8 Mar 2011
- Hankins, T.L., Silverman, R.L.: *Instruments and the Imagination*. Princeton University Press, Princeton, NJ (1995)
- Hey, T., Trefethen, A.E.: Cyberinfrastructure and E-science. *Science* **308**, 818–821 (2005)
- Köhler, T.: Technikgenese als sozialer Prozeß: Das Beispiel Telefon und die Konsequenzen für neue Informations- und Kommunikationstechnologien. In: Coenen, O., Anke, P. (eds.) *Der technologische Wandel am Ende des 20. Jahrhunderts*, pp. 61–74. Shaker-Verlag, Aachen (1998)
- Lattemann, C., Kallischnigg, M., Stieglitz, S.: Einsatz und Nutzung von Social Software in kooperativen Forschungsprojekten. In: Schumann, M.; Kolbe, L.M.; Breitner, M.H., Frerichs, A. (eds.): *Multikonferenz Wirtschaftsinformatik 2010*, pp. 1783–1793. Universitätsverlag Göttingen, Göttingen (2010). http://webdoc.sub.gwdg.de/univerlag/2010/mkwi/03_anwendungen/iktgestuetzte_unternehmenskommunikation/05_einsatz_und_nutzung_von_social_software_in_kooperativen_forschungsprojekten.pdf. Accessed 11 Mar 2020
- Mayntz, R.: Große technische Systeme und ihre gesellschaftstheoretische Bedeutung. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* **45**(1), 97–108 (1993)
- Mayring, P.: Qualitative content analysis. *Forum qualitative social research* **1**(2), Art. 20. <http://nbn-resolving.de/urn:nbn:de:0114-fqs0002204> (2000)
- Nentwich, M.: Cyberscience. Research in the age of the Internet. Austrian Academy of Sciences Press, Wien (2003)
- Nentwich, M., König, R.: Cyberscience 2.0. Research in the age of digital social networks. Campus, Frankfurt am Main and New York (2012)
- Ponte, D., Simon, J.: Scholarly communication 2.0: exploring researchers' opinions on Web 2.0 for scientific knowledge creation, evaluation and dissemination. *Serials Rev.* **37**(3), 149–156 (2011)

- Procter, R., Williams, R., Stewart, J., Poschen, M., Snee, H., Voss, A., Asgari-Targhi, M.: Adoption and use of Web 2.0 in scholarly communications. *Philos. Trans. R. Soc. A* **368**, 4039–4056 (2010)
- Pscheida, D., Köhler, T.: Wissenschaftsbezogene Nutzung von Web 2.0 und Online-Werkzeugen in Sachsen 2012. Studie des „eScience – Forschungsnetzwerks Sachsen“. Medienzentrum der Technischen Universität Dresden, Dresden (2013)
- Pscheida, D., Albrecht, S., Herbst, S., Minet, C., Köhler, T.: Nutzung von Social Media und onlinebasierten Anwendungen in der Wissenschaft. Erste Ergebnisse des Science 2.0-Survey 2013 des Leibniz-Forschungsverbunds „Science 2.0“. Medienzentrum der Technischen Universität Dresden, Dresden (2014)
- Research Information Network and NESTA: Open to all? case studies of openness in research: a joint RIN / NESTA report. http://www.rin.ac.uk/system/files/attachments/NESTA-RIN_Open_Science_V01_0.pdf (2010). Accessed 11 Mar 2020
- Rogers, E.M.: Diffusion of Innovations. 5th ed. Free Press, New York (1995)
- Schmidt, J.H.: Social Software: facilitating information-, identity- and relationship management. In: Burg, T.N., Schmidt, J. (eds.): BlogTalks reloaded, pp. 31–49. Books on Demand, Norderstedt (2007)
- van Eimeren, B., Frees, B.: Ergebnisse der ARD/ZDF-Onlinestudie 2013. Media Perspektiven 7–8/2013, pp. 358–406. http://www.ard-zdf-onlinestudie.de/files/2013/Eimeren_Frees.pdf (2013). Accessed 11 Mar 2020
- Venkatesh, V., Bala, H.: Technology acceptance model 3 and a research agenda on interventions. *Decis. Sci.* **39**(2), 273–315 (2008)
- Waldrop, M. Mitchell: Science 2.0. *Sci. Am.* **298**(5), 68–73 (2008)
- Weller, M.: The digital scholar: how technology is transforming scholarly practice. Bloomsbury Academic. London, New York (2011). <https://www.bloomsburycollections.com/book/the-digital-scholar-how-technology-is-transforming-scholarly-practice>. Accessed 11 Mar 2020
- ZBW: World Wide Wissenschaft – Wie professionell Forschende im Internet arbeiten. Kiel: ZBW Leibniz-Informationszentrum Wirtschaft (2011)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Digital Research Infrastructure



Maik Stührenberg, Oliver Schonefeld, and Andreas Witt

Abstract Digital research infrastructures can be divided into four categories: large equipment, IT infrastructure, social infrastructure, and information infrastructure. Modern research institutions often employ both IT infrastructure and information infrastructure, such as databases or large-scale research data. In addition, information infrastructure depends to some extent on IT infrastructure. In this paper, we discuss the IT, information, and legal infrastructure issues that research institutions face.

Keywords Digital research infrastructure · IT infrastructure · Information infrastructure

1 Introduction

This paper was originally submitted late 2014 and the final publication was delayed until 2019. The authors are well aware that the view and state of the art for digital research infrastructures have evolved in the last 5 years.

A research infrastructure can be defined as a public or private institution that has been established mainly for research, teaching, and the support of young researchers. Research infrastructures can be divided into four main categories (Wissenschaftsrat 2011b, 17f.)¹:

- large equipment, including research platforms such as scientific research vessels, planes, or satellites;

¹Combinations of more than one category are possible as well.

M. Stührenberg (✉)
Universität Bielefeld, Bielefeld, Germany
e-mail: maik.stuehrenberg@uni-bielefeld.de

O. Schonefeld · A. Witt
Institut Für Deutsche Sprache, Mannheim, Germany
e-mail: schonefeld@ids-mannheim.de

A. Witt
e-mail: witt@ids-mannheim.de

- IT infrastructure, such as computer hardware and software;
- social infrastructure, that is, research institutions that offer scholars a place to exchange ideas and collaborate (Wissenschaftsrat 2011a, 20f.), for example, the Leibniz Center in Dagstuhl Castle, Germany;
- information infrastructure, that is, research institutions that collect and curate primary data and make them accessible to a larger group of scholars.

While large technical equipment is only seldom used in digital humanities disciplines, and social infrastructure is beyond the scope of this paper, combinations of IT infrastructure and information infrastructure are quite common. Therefore, the purpose of this paper is to give insight into various aspects of modern research infrastructures with an emphasis on both the latter categories. In addition, we have conducted a qualitative analysis by interviewing twelve German research institutions (Fiedler et al. 2012). The institutions were interviewed and asked to participate in a survey. The 74 survey questions were structured into different topic areas, such as organizational aspects, data management, hardware and software, environmental aspects, and legal issues. We will reflect on some of these topics in the respective sections of this article.

2 IT Infrastructure

Digital humanities research institutions working with huge amounts of data (e.g., language corpora) have special needs regarding IT infrastructure, such as a growing demand for storage space, computing capacity (for querying and analyzing linked data), and durability (including distributed access over large-scale networks such as the Internet for a huge number of potential users). This results in significant amounts of money spent on hardware and software. In addition, operating costs (divided into maintenance and personnel costs) have to be taken into account, including IT staff, hardware maintenance, software updates, and licensing. Especially energy costs should not be underestimated, as the price of electricity is increasing over time. A green-IT strategy can help an institution to reduce some of these costs. A key way of doing this is buying new equipment and replacing old (less energy-efficient) hardware. However, green IT consists of more aspects, such as efficient cooling (like separation of warm and cold aisles in the data center or using free cooling techniques), institutional policies (e.g., obliging employees to turn their computers off before leaving the workplace), or using supplies made of recycled material (like recycled paper). Implementing a green-IT strategy is generally a project of its own for a research institution and is currently a low priority for the institutions that we analyzed.

Therefore, one of the issues modern-day research institutions have to deal with is to optimize these costs, usually by undertaking the following steps. Firstly, a transparent accounting system, including every single asset for salaries, maintenance costs, and so forth, has to be established, allowing for a more accurate estimation of current

and future demands for IT infrastructure. Replacing proprietary software with open-source software may only slightly decrease licensing costs, but may be cheaper in the long run since the latter can be adopted to the institution's needs and usually has better support of open formats (see Sect. 3.2). However, two points have to be considered regarding this assumption:

1. Additional costs for user training may be necessary if the open-source application differs from the formerly used product;
2. In-house IT expertise is necessary to adapt open-source software, which may result in even higher personnel costs.

For these reasons, it is advisable, especially for smaller research institutions, to collaborate in the field of IT infrastructure to reduce costs. Examples of such cooperation include a shared Internet connection, server housing, or archival storage. A majority of the interviewed research institutions already collaborate with other external facilities to lower IT costs and to distribute archival and backup storage. Since research institutions are nowadays connected to the Internet, storage of and access to the information infrastructure involves special security requirements. Two main issues have to be considered:

1. preventing unauthorized access to systems, processes, or data (including information infrastructure);
2. ensuring that hardware and software continue to function.

Although there is no such thing as a completely secure network, the first step to prevent unauthorized access is a complete risk analysis for the relevant computer systems, including estimating possible losses and limitations on daily work (e.g., due to vandalism or sabotage). The outcome of this analysis should be a prioritized list of data and systems to be protected.

The concrete security measures (the security policy) are defined by the IT security officer and the data protection officer and are mandatory for the whole staff of the research institution (ISO/IEC 27002:2013 2013; BSI 2014). Important points for a security policy are:

- prioritization of data according to their value for the research institution;
- identification of possible risks (including computer viruses and network infrastructure attacks);
- backup strategy;
- data encryption.

While a backup strategy for research data is considered crucial (nine out of twelve interviewees have a central backup strategy and the remaining institution plans to implement one), only a third of the institutions surveyed have a central in-house IT security policy.

3 Information Infrastructure

Research data, especially primary data (e.g., recordings, measurements, and curated corpora), are among the most valuable assets for a research institution. Research institutions that can be categorized as information infrastructures (such as libraries, archives, collections, and smaller non-academic research institutions) that collect and curate primary data, scientific and non-scientific knowledge, and databases, and provide access to researchers [34], who may use this data for research projects on their own. To ensure access to the information infrastructure, various technical aspects have to be taken into account.

3.1 *Repositories and Publication Server*

Repositories have already been used in large-scale collaborative projects, often international ones, such as CLARIN.² The CLARIN centers provide repositories storing academic research data (such as curated corpora) accessible via the Internet. Retrieval of a desired information item is highly dependent on metadata. Following on from existing metadata standards such as Dublin Core (ISO 15836:2009 [2009](#); DCMI [2012](#)), IMDI (ISLE Metadata Initiative [2003](#); Broeder and Wittenburg [2006](#); ISLE Metadata Initiative [2009](#)), or OLAC (Simons et al. [2008](#); Bird and Simons [2009](#)), the Component Metadata Structure (CMDI) (Broeder et al. [2011](#), [2012](#); Trippel et al. [2012](#)) has been created to facilitate documenting research information and querying it over the distributed repositories. In our survey, five out of the twelve interviewed institutes already run a repository on their own, while four are in the process of building one.

Another aspect of information infrastructure is the archiving and accessibility of publications. Establishing and maintaining an in-house publication server can be a way for a research institution to retain both copyright (see Sect. 4.1) and access control over information that has been produced by its academic staff. Open-source implementations, such as ePrints³ or eSciDoc,⁴ often combine the functionalities of publication servers and primary data repositories. For all these tasks, staff working on IT and information infrastructure need to collaborate closely. In particular, research institutions having their own libraries can benefit from the expertise of IT and information departments regarding archives, metadata, and retrieval. Seven of the interviewees already run a publication server.

²See <http://www.clarin.eu> for further details.

³See <http://www.eprints.org/> for further details.

⁴See <https://www.escidoc.org/> for further details.

3.2 *Data Formats*

Although the creation of research data is often quite expensive, a large portion of this information gets lost shortly after the end of the project in which it was gathered. Apart from the hardware failures or insufficient metadata discussed above, another possible reason can be a proprietary storage format, for which the corresponding application is not available any more.

Data formats usually exist for two reasons: (1) as serialization of a specification, or (2) as the import and export format of an application. A format as such may be open or proprietary, which may be important for processing and archiving the information encoded in it. An example of a proprietary de facto standard format is the ubiquitous.doc format, produced by Microsoft Word.⁵ Since it is a binary format, it is not possible to extract information with arbitrary text editors; instead, one has to use specific programs, and applications other than MS Word may not be able to successfully render the document as it was intended by the author.

For research data which are curated by an information infrastructure, open text-based formats should be preferred. Formats based on the open meta language XML (Bray et al. 2008) are quite common in academic research and can be defined by document grammar formalisms such as XML DTD (part of the aforementioned specification), XML Schema (Gao et al. 2012; Peterson et al. 2012), or RELAX NG (ISO/IEC 19757-2:2008 2008), allowing for on-the-fly validation during the creation of instances. Examples of open XML-based annotation formats in the digital humanities are the TEI Guidelines (Burnard and Bauman 2014) or DocBook (Walsh 2010) for technical documentation. Information encoded in those formats is not only readable with common text editors, but separates content from formatting, since the rendering is usually controlled by separate XSLT (Kay 2007, 2014) or CSS (Bos et al. 2011) stylesheets. This not only prevents vendor lock-in, but significantly eases the process of archiving. The attitude to open standards and open-source software compared with proprietary in-house development is mixed; however, there is a tendency to use standardized APIs and formats, or at least consider open-source applications. Seven surveyed institutes keep data in proprietary formats, while four aim to use standard formats and one is still determining its strategy. Often, institutes lack the human resources to convert data into standard formats.

4 Legal Issues

Research institutions are confronted with a number of legal issues, the most important of which are: (1) copyright and (2) personal data protection and privacy.

⁵Note that we are talking about the binary .doc, not the XML-based .docx format used by Office 2004 onwards and that is standardized as ISO/IEC 29500-1:2011 (2011). However, even the latter format uses a number of features that cannot easily be interpreted by application programs without further knowledge.

4.1 Copyright Issues

Research data is often based on material contributed by third parties. The primary data of text corpora, for example, often originate from newspaper articles or similar non-academic sources. German copyright law protects literary, artistic, and scientific works (including software) that are the author's own intellectual creation. Copyright-protected works may only be modified (and, arguably, annotated) with the authorization of the copyright holder. Copyright expires 70 years after the death of the original author. In Germany (unlike in most other jurisdictions), copyright cannot be transferred and is reserved by the author until his death (and 70 years after it), but it can be licensed. In practice, authors often license their rights out to publishers.

Although the German copyright law (UrhG) does not contain the American concept of “fair use”, there are copyright limitations (§§ 44a–63a UrhG) that apply to certain specific uses of copyright-protected works (e.g., citations, personal use, scientific use) (Mönch 2006). However, in order to be covered by a copyright limitation of § 52a UrhG, scientific use has to be restricted to “small groups of researchers” (Hoeren 2014, 157). This is especially important if a research institution wants to publish annotated corpora—in that case, the primary data has to be licensed beforehand.

Research data to which a research institution holds the copyright (e.g., primary data produced in-house) should be made available to others under a liberal license, e.g., an open-access license such as Creative Commons.⁶ Creative Commons (CC) is a free license (similar to the software license, BSD,⁷ or the General Public License, GNU⁸) that was originally developed for creative work and that consists of several building blocks, such as Attribution (BY: minimal requirement), NoDerivatives (ND), NonCommercial (NC),⁹ and ShareAlike (SA). The current version (4.0) also addresses specific database rights that exist in EU Member States.

Apart from human-readable CC license deeds, laundry symbols (similar to those established in the CLARIN research group (Oksanen et al. 2010) for its own specific licenses) provide a quick overview of the license requirements.¹⁰ For a detailed discussion about legal implications of institutional repositories see Bargheer et al. (2006).

Regarding publications, a research institution's staff may agree to publish their works on the institution's publication server under an open-access license (Degkwitz 2007). Open-access publications have steadily gained ground in countries such as the US, Denmark, or Japan, while there is still an ongoing discussion about them in Germany, especially in the digital humanities disciplines¹¹—although the Berlin

⁶See <http://creativecommons.org> for further details.

⁷See <http://opensource.org/licenses/bsd-license.php> for further details.

⁸See <http://www.gnu.org/licenses/#GPL> for further details.

⁹Especially NC may have undesired side effects, see Klimpel (2012) for a discussion.

¹⁰The categories have recently been extended by Kupietz and Lüngen (2014).

¹¹See Görl et al. (2011) for a discussion about the impacts of information infrastructure in universities of North Rhine-Westphalia.

Declaration on Open Access to Knowledge in the Sciences and Humanities¹² has boosted their reputation. While open-access journals are still sometimes seen as less reputable than traditional journals (although both publication types monitor quality through peer review), they often have higher citation numbers.¹³ Research institutions can play an active role in the process of building the reputation of open access by publishing in this format. It is therefore pleasant to see that an open-access strategy is already present in five of the institutions interviewed, while three of them plan on implementing one.

4.2 Personal Data Protection

Personal data protection issues may arise when living persons are involved in the process of creating research data, such as voice or video recordings. Publication of personal data is only allowed if the persons recorded have given their (written) consent. For every collection of personal data, a register of processing operations has to be created (according to §4 g, §§18 and 4e of the German data protection law, BDSG). The type of personal information, how it is processed, and the data protection measures, are recorded in this register.

Despite the variety of legal issues that may arise for research institutions, most of the interviewees rely either on their own (general) legal department or on cooperation with external law firms. Licensed (IT law) attorneys are seldom employed. However, since German research institutions are required to employ a data protection officer if they deal with personal data, they already have at least some existing in-house expertise. This expert should be involved in any data collection activities as soon as possible.

5 Conclusion

We have discussed a number of information infrastructure issues that modern research institutions need to consider. Most of the technical issues can be addressed by implementing a sustainable long-term IT strategy that reflects both costs and demands. Additional technical aspects such as security, open storage formats, and metadata can be addressed in such an IT strategy. Legal issues cannot be underrated, especially for service-oriented research institutions. Therefore, a data protection officer should be involved in the early stages of research projects that plan to create personal data.

¹² See the text of the declaration at http://openaccess.mpg.de/3515/Berliner_Erklaerung.

¹³ See Stempfhuber (2009, 119) and <http://opcit.eprints.org/oacitation-biblio.html> for a number of studies about open-access impact factors.

References

- Bargheer, M., Bellem, S., Schmidt, B.: Open Access und Institutional Repositories—Rechtliche Rahmenbedingungen. In: Spindler, G. (ed.) Rechtliche Rahmenbedingungen von Open Access-Publikationen, pp. 1–20. No. 2 in Göttinger Schriften zur Internetforschung, Universitätsverlag Göttingen (2006)
- Bird, S., Simons, G.F.: OLAC: Accessing the world's language resources. In: Proceedings of the 1st International Conference on Language Documentation and Conservation. Hawaii (2009)
- Bos, B., Celik, T., Hickson, I., Lie, H.W.: Cascading style sheets level 2 revision 1 (css 2.1) specification. W3C Recommendation, World Wide Web Consortium (W3C), <http://www.w3.org/TR/2011/REC-CSS2-20110607> (2011)
- Bray, T., Paoli, J., Sperberg-McQueen, C.M., Maler, E., Yergeau, F.: Extensible Markup Language (XML) 1.0 (Fifth Edition). W3C Recommendation, World Wide Web Consortium (W3C), <http://www.w3.org/TR/2008/REC-xml-20081126/> (2008)
- Broeder, D., Schonefeld, O., Trippel, T., van Uytvanck, D., Witt, A.: A pragmatic approach to XML interoperability—the Component Metadata Infrastructure (CMDI). In: Proceedings of Balisage: The Markup Conference. Balisage Series on Markup Technologies, vol. 7. Montréal (2011)
- Broeder, D., Windhouwer, M., van Uytvanck, D., Trippel, T., Goosen, T.: CMDI: a component metadata infrastructure. In: Arranz, V., Broeder, D., Gaiffe, B., Gavrili-dou, M., Monachini, M., Trippel, T. (eds.) Proceedings of the LREC 2012 Workshop on Describing LRs with Metadata: Towards Flexibility and Interoperability in the Documentation of LR, pp. 1–4. (2012)
- Broeder, D., Wittenburg, P.: The IMDI metadata framework, its current application and future direction. Int. J. Metadata Semant. Ontol. **1**(2), 119–132 (2006)
- Bundesamt für Sicherheit in der Informationstechnik: IT-Grundschutz- Vorgehensweise. Version 2.0. BSI-Standard 100–2, BSI, https://www.bsi.bund.de/SharedDocs/Doloads/DE/BSI/Publikationen/ITGrundschatzstandards/BSI-Standard_1002.pdf?__blob=publicationFile (2014)
- Burnard, L., Bauman, S. (eds.): TEI P5: Guidelines for electronic text encoding and interchange. Text Encoding Initiative Consortium, Charlottesville, Virginia, version 2.6.0. Last updated on 20th January 2014, revision 12802 (2014)
- DCMI Usage Board: Dublin Core Metadata Element Set, Version 1.1. DCMI Recommendation, Dublin Core Metadata Initiative, <http://dublincore.org/documents/2012/06/14/dces/> (2012)
- Degkwitz, A.: Open access und die Novellierung des deutschen Urheberrechts. Zeitschrift für Bibliothekswesen und Bibliographie **54**(4/5), 243–245 (2007)
- Fiedler, N., Werthmann, A., Stührenberg, M., Bingel, J., Witt, A.: Forschungsinfrastrukturen in außeruniversitären Forschungseinrichtungen. Forschungsbericht, Institut für Deutsche Sprache. <https://nbn-resolving.org/urn:nbn:de:bsz:mh39-28325> (2012)
- Gao, S.S., Sperberg-McQueen, C.M., Thompson, H.S.: W3C XML Schema Definition Language (XSD) 1.1 Part 1: Structures. W3C Recommendation, World Wide Web Consortium (W3C), <http://www.w3.org/TR/2012/REC-xmlschema11-1-20120405/> (2012)
- Görl, S., Puhl, J., Thaller, M.: Empfehlungen für die weitere Entwicklung der Wissenschaftlichen Informationsversorgung des Landes NRW. epubli (2011)
- Hoeren, T.: Internettrecht. Universität Münster. <https://www.itm.nrw/wp-content/uploads/Skript-Internetrecht-April-2014.pdf> (2014)
- ISLE Metadata Initiative: Metadata Elements for Session Descriptions. version 3.0.4. Reference Document, MPI, Nijmegen, https://tla.mpi.nl/wp-content/uploads/2012/06/IMDI_MetaData_3.0.4.pdf (2003)
- ISLE Metadata Initiative: Metadata Elements for Catalogue Descriptions. version 3.0.13. Reference Document, MPI, Nijmegen, https://tla.mpi.nl/wp-content/uploads/2012/06/IMDI_Catalogue_3.0.0.pdf (2009)
- ISO 15836:2009: The Dublin Core Metadata Element Set, Version 1.1. International Standard, International Organization for Standardization, Geneva (2009)

- ISO/IEC 19757-2:2008: Information technology—Document Schema Definition Language (DSDL) – Part 2: Regular-grammar-based validation—RELAX NG. International Standard, International Organization for Standardization, Geneva (2008)
- ISO/IEC 27002:2013: Information technology—Security techniques—Code of practice for information security controls. International Standard, International Organization for Standardization/International Electrotechnical Commission (2013)
- ISO/IEC 29500-1:2011: Information technology—Document description and processing languages—Office Open XML File Formats—Part 1: Fundamentals and Markup Language Reference. International standard, International Organization for Standardization, Geneva (2011)
- Kay, M.: XSL Transformations (XSLT) Version 2.0. W3C Recommendation, World Wide Web Consortium (W3C). <http://www.w3.org/TR/2007/REC-xslt20-20070123/> (2007)
- Kay, M.: XSL Transformations (XSLT) Version 3.0. W3C Last Call Working Draft, World Wide Web Consortium (W3C). <http://www.w3.org/TR/2014/WD-xslt-30-20141002/> (2014)
- Klimpel, P.: Freies Wissen dank Creative-Commons-Lizenzen. Folgen, Risiken und Nebenwirkungen der Bedingung 'nicht-kommerziell—NC'. Wikimedia Deutschland and iRights.info and Creative Commons Deutschland. http://irights.info/userfiles/CC-NC_Leitfaden_web.pdf (2012)
- Kupietz, M., Lüngen, H.: Recent developments in DeReKo. In: Calzolari, N., Choukri, K., Declerck, T., Loftsson, H., Maegaard, B., Mariani, J., Moreno, A., Odijk, J., Piperidis, S. (eds.) Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14). European Language Resources Association (ELRA), Reykjavik (2014)
- Mönch, M., Nödler, J.M.: Hochschulen und Urheberrecht – Schutz wissenschaftlicher Werke. In: Spindler, G. (ed.) Rechtliche Rahmenbedingungen von Open Access- Publikationen, pp. 21–54. No. 2 in Göttinger Schriften zur Internetforschung, Universitätsverlag Göttingen (2006)
- Oksanen, V., Lindén, K., Westerlund, H.: Laundry symbols and license management—practical considerations for the distribution of LRs based on experiences from CLARIN. In: Arranz, V., van Eerten, L. (eds.) Language Resources: From Storyboard to Sustainability and LR Lifecycle Management, Workshop held at the seventh conference on International Language Resources and Evaluation (LREC 2010), pp. 10–13. Valletta (2010)
- Peterson, D., Gao, S.S., Malhotra, A., Sperberg-McQueen, C.M., Thompson, H.S.: W3C XML Schema Definition Language (XSD) 1.1 Part 2: Datatypes. W3C Recommendation, World Wide Web Consortium (W3C). <http://www.w3.org/TR/2012/REC-xmlschema11-2-20120405/> (2012)
- Simons, G.F., Bird, S.: OLAC Metadata. Olac standard, Open Language Archives Community. <http://www.language-archives.org/OLAC/metadata-20080531.html> (2008)
- Stempfhuber, M.: Die Rolle von "open access" im Rahmen des wissenschaftlichen Publizierens. In: Publikationsverhalten in unterschiedlichen wissenschaftlichen Disziplinen. Beiträge zur Beurteilung von Forschungsleistungen, pp. 116–131. Diskussionspapier der Alexander von Humboldt-Stiftung. 2 edn. Alexander von Humboldt Stiftung, Bonn (2009)
- Trippel, T., Hoppermann, C., Depoorter, G.: The component metadata infrastructure (cmdi) in a project on sustainable linguistic resources. In: Arranz, V., Broeder, D., Gaiffe, B., Gavrilidou, M., Monachini, M., Trippel, T. (eds.) Proceedings of the LREC 2012 Workshop on Describing LRs with Metadata: Towards Flexibility and Interoperability in the Documentation of LR. pp. 29–36. (2012)
- Walsh, N.: DocBook 5: The Definitive Guide. O'Reilly Media, Sebastopol (2010)
- Wissenschaftsrat: Empfehlungen zu Forschungsinfrastrukturen in den Geistes- und Sozialwissenschaften. Berlin. <http://www.wissenschaftsrat.de/download/archiv/10465-11.pdf> (2011a)
- Wissenschaftsrat: Übergreifende Empfehlungen zu Informationsinfrastrukturen. Berlin. <http://www.wissenschaftsrat.de/download/archiv/10466-11.pdf> (2011b)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



MOVING: A User-Centric Platform for Online Literacy Training and Learning



Aitor Apaolaza, Tobias Backes, Sabine Barthold, Irina Bienia, Till Blume, Chrysa Collyda, Angela Fessl, Sebastian Gottfried, Paul Grunewald, Franziska Günther, Thomas Köhler, Robert Lorenz, Matthias Heinz, Sabrina Herbst, Vasileios Mezaris, Chifumi Nishioka, Alexandros Pournaras, Vedran Sabol, Ahmed Saleh, Ansgar Scherp, Ilija Simic, Andrzej M.J. Skulimowski, Iacopo Vagliano, Markel Vigo, Michael Wiese, and Tanja Zdolšek Draksler

Abstract In this paper, we present an overview of the MOVING platform, a user-driven approach that enables young researchers, decision makers, and public administrators to use machine learning and data mining tools to search, organize, and manage large-scale information sources on the web such as scientific publications, videos of research talks, and social media. In order to provide a concise overview of the platform, we focus on its front end, which is the MOVING web application.

A. Apaolaza · M. Vigo
University of Manchester, Manchester, UK
e-mail: aitor.apaolaza@manchester.ac.uk

M. Vigo
e-mail: markel.vigo@manchester.ac.uk

T. Backes
GESIS, Cologne, Germany
e-mail: tobias.backes@gesis.org

S. Barthold · P. Grunewald · F. Günther · T. Köhler (✉) · R. Lorenz · M. Heinz · S. Herbst
Media Centre, TU Dresden, Dresden, Germany
e-mail: thomas.koehler@tu-dresden.de

S. Barthold
e-mail: sabine.barthold@tu-dresden.de

P. Grunewald
e-mail: paul.grunewald@tu-dresden.de

F. Günther
e-mail: franziska.guenther1@tu-dresden.de

R. Lorenz
e-mail: robert.lorenz@tu-dresden.de

M. Heinz
e-mail: matthias.heinz@tu-dresden.de

S. Herbst
e-mail: sabrina.herbst@tu-dresden.de

By presenting the main components of the web application, we illustrate what functionalities and capabilities the platform offer its end-users, rather than delving into the data analysis and machine learning technologies that make these functionalities possible.

Keywords MOVING platform · MOVING web application · Recommender system · Adaptive training support

1 Introduction

Scholars and professionals in various sectors of the economy, including public administrators, corporate compliance officers, and auditors, deal with an ever-increasing flow of information (new scientific publications, business documents and multimedia files, laws, etc.). They need sophisticated tools to evaluate all this information fast and accurately and to visualize the analysis results. Specifically this means that, on the one hand, they need tools that enable state-of-the-art search and semantic analysis of large digital contents, by providing: (i) access to an extensive source inventory, (ii)

I. Bienia · M. Wiese
Ernst & Young, Essen, Germany
e-mail: irina.bienia@de.ey.com

M. Wiese
e-mail: michael.wiese@de.ey.com

T. Blume
University of Kiel, Kiel, Germany
e-mail: tbl@informatik.uni-kiel.de

C. Collyda · V. Mezaris · A. Pournaras
CERTH-ITI, Thessaloniki, Greece
e-mail: ckol@iti.gr

V. Mezaris
e-mail: bmezaris@iti.gr

A. Pournaras
e-mail: apournaras@iti.gr

A. Fessl · V. Sabol · I. Simic
Know-Center, Graz, Austria
e-mail: afessl@know-center.at

V. Sabol
e-mail: vsabol@know-center.at

I. Simic
e-mail: isimic@know-center.at

S. Gottfried
DLR Institute of Software Methods for Product Virtualization, Dresden, Germany
e-mail: sebastian.gottfried@dlr.de

advanced search and visualization methods, and (iii) functionalities for generating new knowledge from these digital assets. On the other hand, these tools need to be reasonably easy for their users to understand and support them through: (i) a detailed and scientifically proven help system (tutorials, guidance), individually configurable training programmes (learning modules, videos), and a lively community of people that have similar interests or problems to be solved. To face these challenges, the interdisciplinary trans-European project called MOVING (“TraininG towards a society of data-saVvy inforMation prOfessionals to enable open leadership INnovation”) (Vagliano et al. 2018) has built an innovative training platform that enables users from various societal sectors to fundamentally improve their information literacy by training in how to choose, use, and evaluate data mining methods in their daily research and business tasks, and to become data-savvy information professionals.

2 Digitized Science

Initiatives by the European Union (which has long been pursuing a digital agenda) to support research in the field of digitized science illustrate the need to investigate related change processes (European Commission 2016). Obviously, empirical and theoretical justification is needed to develop the practice of science. The innovative approach dealt with here was developed in the MOVING project, which offers an innovative training platform to support scientists and other users from all areas of society to fundamentally improve their information literacy in research-oriented

C. Nishioka

University of Kyoto, Kyoto, Japan

e-mail: nishioka.chifumi.2c@kyoto-u.ac.jp

A. Saleh

ZBW, Kiel, Germany

e-mail: a.saleh@zbw.eu

A. Scherp

Ulm University, Ulm, Germany

e-mail: ansgar.scherp@uni-ulm.de

A. M.J. Skulimowski

Progress Business Foundation, Krakow, Poland

e-mail: ams@agh.edu.pl

T. Zdolšek Draksler

Jožef Stefan Institute, Ljubljana, Slovenia

e-mail: tanja.zdolsek@ijs.si

I. Vagliano

University Medical Center, Amsterdam, The Netherlands

e-mail: i.vagliano@amsterdamumc.nl

contexts.¹ The project is about training users to select, apply, and evaluate technologies and data mining methods, so that the relevant research staff can develop into ‘data-savvy’ information professionals in their daily research routines (Scherp et al. 2016; Köhler et al. 2016a, b).

In terms of content, the research methodological changes in scientific action cannot easily be explained as domain-specific activities. This requires analyses of both current technological developments and the changes in how scientists use these technologies (or methods). The eScience Saxony research network provides statements on both perspectives (see, e.g., [Pscheida et al. 2013, 2014]). The network has observed the following:

- there is great potential for the use of new digital tools in research;
- preferred topics for development are scientist collaboration and the visualization of (often large or new) databases;
- transitions between the subject areas of research and teaching can also be observed in technology development;
- almost all scientists do most of their work using computer-based technologies and have access to appropriate infrastructures;
- scientists sometimes find it difficult to adopt new media technologies in research and teaching (e.g. social media), although there are also subject-specific differences;
- there is still uncertainty regarding the requirements, possibilities, and assumed risks of open-access publishing;
- research methodology has not been fully systematically discussed and is often inadequately implemented;
- there are no clear standards for high-quality research technology and no recognizable institutionalization to support open-access trends in science, so these still need to be worked out together;
- digital change in science is comparatively rapid from an individual (scientist) perspective, the outcome is not known, especially regarding location-determining infrastructures.

Indeed the listing matches to a larger proportion with the demands of these cases addressed by the MOVING project. Nevertheless MOVING did set focus on two more main characteristics. First there was a serious interest to address research activity not only in academia but as well in public administration and industry. Second, when developing the approach the project consortium decided to include as well a direct focus on the related skill development, i.e. include a serious effort on innovation in the educational dimension (the Online Literacy Training and Learning) that needs to go along with any new technology in every sector.

¹Platform.moving-project.eu, last accessed 7 May 2020.

3 Overview of the MOVING Platform

An overview of the MOVING platform architecture is illustrated in Fig. 1, which shows the most important components and their relationships. The main component

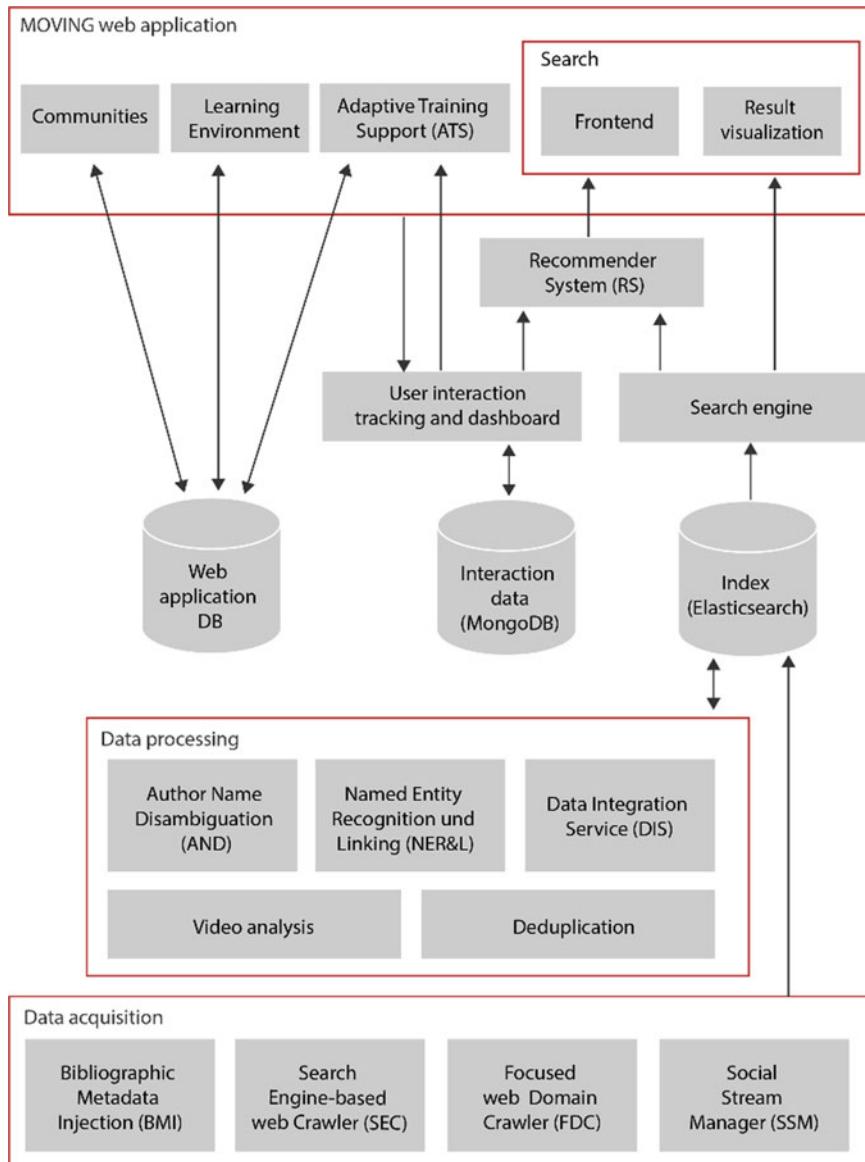


Fig. 1 MOVING platform architecture

blocks are (i) data acquisition, (ii) data processing, (iii) back-end data storage, user tracking, search and recommendation, and (iv) the MOVING web application that includes the front-end search. In this section, we briefly describe the overall platform.

The MOVING web application is the core of the platform and the interface to the user. The main entry points to the web application are the community section, the learning environment, and the search interface. The search interface offers different visual representations of search results. These visualizations allow the user to explore the search results in various ways. For this purpose, four visualizations have been added to the MOVING platform, namely: (i) the Concept Graph, which displays the search results as an interactive network, (ii) uRank, a dynamic document ranking view, (iii) Top Properties, a bar chart visualization that aggregates the results based on their properties, and (iv) a Tag Cloud, showing the most frequently occurring keywords. Moreover, the Adaptive Training Support (ATS) widget supports users learning how to search and provides material suited to their needs (Fessl et al. 2018) and the Recommender System (RS) widget (bridging the front and back ends of the platform) points users to potentially relevant documents by evaluating their last search queries. Thanks to its responsive design, all the views adapt to different screen sizes, automatically changing the layout according to the capabilities of the device.

Private user data and public documents are stored in three separate databases: The web application database holds the data for the communities, the learning environment, and the ATS. The index holds the public documents and generated meta-data information such as topics, authors, and extracted entities. The user-interaction tracking captures user interactions with the web application and stores them securely in a third database. User tracking provides additional data for both the ATS and the RS, which form the basis for user support by these two widgets.

The index used by the search interface is populated by various data acquisition components (e.g. web crawlers and a Bibliographic Metadata Injection service), to increase the amount of data accessible through the MOVING platform. To date, it hosts over 22 million documents and metadata records. These records include books, scientific articles, laws and regulations, documents about funding opportunities, videos (e.g. of lectures and tutorials), and social media posts. Data processing components have been incorporated into and applied to these records, to improve the quality of data and make it easier to search. Additional features, the Data Integration Service, Author Name Disambiguation, Deduplication, Named Entity Recognition and Linking, and Video Analysis, all refine and enrich the documents stored in the index.

Author name disambiguation addresses the problem that many author names belong to different real-world authors. To deal with this problem, a novel method (Backes 2018a, b) has been developed which applies, for a given author name, agglomerative clustering on features extracted from documents containing the author mention in question, such as affiliation, co-authors, referenced authors, email addresses, keywords, and publication years. The disambiguation procedure calculates the probability with which author mentions with the same name belong to the same person. Name mentions having a high probability to belong to the same author are

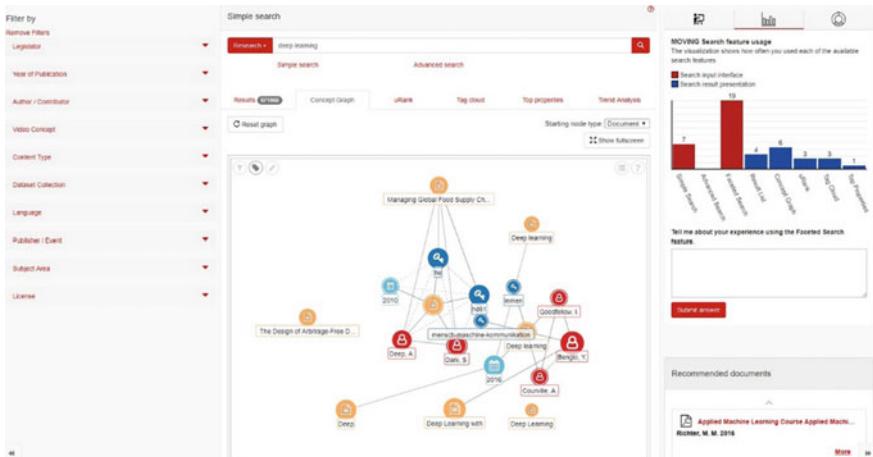


Fig. 2 MOVING search and results page

assigned a unique internal authorID. By this, authors with the same name are distinguished if they refer to different real-world persons. As a result, users who click on the name of an author of a document in the result list of a search will only see documents from authors who have the same author ID as the selected author (instead of showing all documents authored by any person with that name). A modified version of this method has been applied for document deduplication.

In the following, we present the front end of the MOVING platform in detail, in order to provide a concise summary of what a user can do with it. For details on how individual data processing, data acquisition, and other back-end components work, the interested reader is referred to the relevant publications, such as (Nishioka and Scherp 2016; Galanopoulos and Mezaris 2019; Tzelepis et al. 2018), as well as the documentation available on the MOVING project web site.²

4 The MOVING Web Application

4.1 Search

Search is a key functionality in the MOVING web application. At the back end, the MOVING search engine is based on Elasticsearch,³ given appropriate parameters, and fine-tuned to efficiently index dozens of millions of documents. At the front end, the user sees a search page (Fig. 2), with various search options and filters on the left, visualizations of the results in the centre of the window, and training functionalities

²www.moving-project.eu, last accessed 7 May 2020.

³www.elastic.co, last accessed 7 May 2020.

Fig. 3 Search history view

ID	Name	Query	Documents	Date last run	
90	deep learning	[Search Domain: research] ∩ [Query Term: deep learning] ∩ [Year of Publication: 2014,2019] ∩ [Content Type: Video] ∩ [Lecture] ∩ [Poster] ∩ [Invited Talk] ∩ [Tutorial] ∩ [Event] ∩ [Keynote] ∩ [Event Section] ∩ [Panel] ∩ [Project] ∩ [Best Paper] ∩ [Debate] ∩ [Opening] ∩ [Summary]	725	2019-03-07	Run
89	Support vector machines	[Search Domain: research] ∩ [Query Term: Support vector machines]	1709706	2019-03-07	Run
88	deep learning	[Search Domain: research] ∩ [Query Term: deep learning] ∩ [Year of Publication: 2014,2019] ∩ [Content Type: Video] ∩ [Lecture] ∩ [Poster] ∩ [Invited Talk] ∩ [Tutorial] ∩ [Event] ∩ [Keynote] ∩ [Event Section] ∩ [Panel] ∩ [Project] ∩ [Best Paper] ∩ [Debate] ∩ [Opening] ∩ [Summary]	725	2019-03-07	Run
87	machine learning	[Search Domain: research] ∩ [Query Term: machine learning]	618569	2019-03-07	Run

such as ATS on the right. The search history of the current user can also be viewed, to support future searches.

To enable platform users to view and replicate their previous searches, the search history view is connected with WevQuery (Apaolaza and Vigo 2017). WevQuery serves as an interface to the data generated by UCIVIT (Apaolaza et al. 2013), the tracking tool of which logs user-interaction data. From WevQuery, we get the information about the previous user searches, time when the user performed the search query, and the number of documents retrieved. This information is then utilized to build the search history view, an example of which is shown in Fig. 3.

To present the results of a user query effectively, several visualizations have been implemented. Four characteristic ones are:

- **Concept Graph.** For the discovery and exploration of relationships between documents and their properties.
- **uRank.** A tool for the interest-driven exploration of search results.
- **Top Properties.** A bar chart displaying aggregated information about the properties of the retrieved documents.
- **Tag Cloud.** A visualization for the analysis of keyword frequency in the retrieved documents.

Concept Graph: an interactive network visualization the Concept Graph (Fig. 4) visualizes direct and indirect connections between retrieved search results. For example, a single, disambiguated author of two different publications is visualized as a node in the graph connecting the corresponding publications. Further extracted and disambiguated entities are visualized in a way that users can grasp, quickly, such as research networks. The initial graph visualization starts with a few collapsed nodes. These nodes can be expanded to visualize initially hidden nodes and to incrementally add more information to the graph. Thus, users are not overwhelmed with too much information when they start their search.

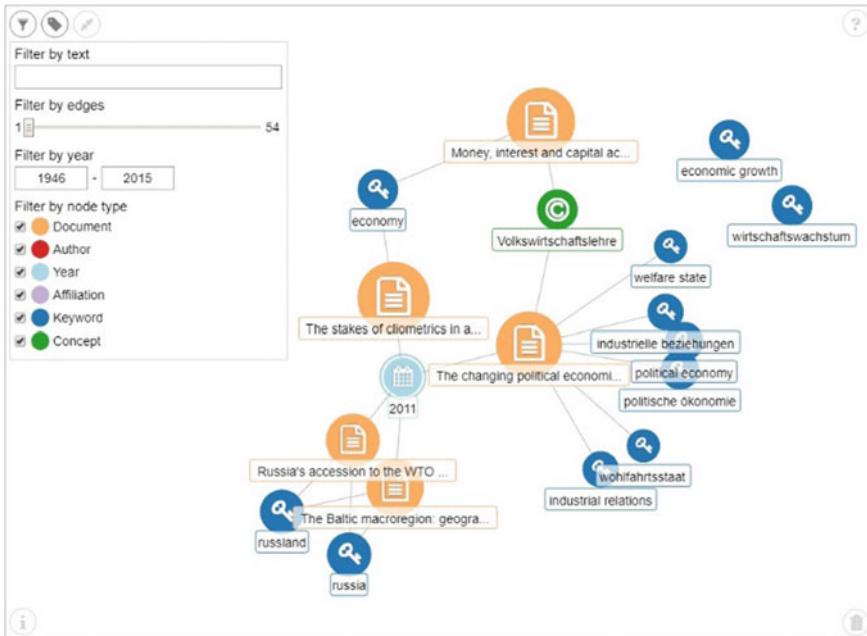


Fig. 4 Concept Graph with opened filter menu

uRank: interest-based result set exploration. Based on the search query the top 100 retrieved results are displayed as a ranked list. The keywords extracted from the results are presented in the Tag Cloud in the right sidebar of uRank (Fig. 5, point A). By selecting keywords of interest, the results in the list (Fig. 5, point C) are re-ranked in such a way that the results containing the selected keyword move to the top. The ranking view (Fig. 5, point D) provides visual feedback on the relevance of the result. It is possible to select multiple keywords and even fine-tune their importance by using the slider under the selected words (Fig. 5, point B). Clicking on a result opens a dialogue box, which presents additional information about the retrieved document. The user can export the current view of uRank, with the current search configuration, by clicking on the export button, which initiates the download of a zip file containing an image and a report text file.

Top Properties: the Top Properties visualization uses 100 of the most relevant results from the current search query. It shows a bar chart visualization presenting one of the following properties of the available results: Authors, Keywords, Concepts, Sources, and Year of Publication. The results are ordered according to the most frequent values of the selected property, as can be seen in Fig. 6. When the publication year is selected, the sorting order changes so that the years are displayed in chronological order to make it easier to identify year-on-year changes. Clicking on one of the bars shows the results associated with this property in a small dialogue box. The results in this dialogue are sorted in the order provided originally by the

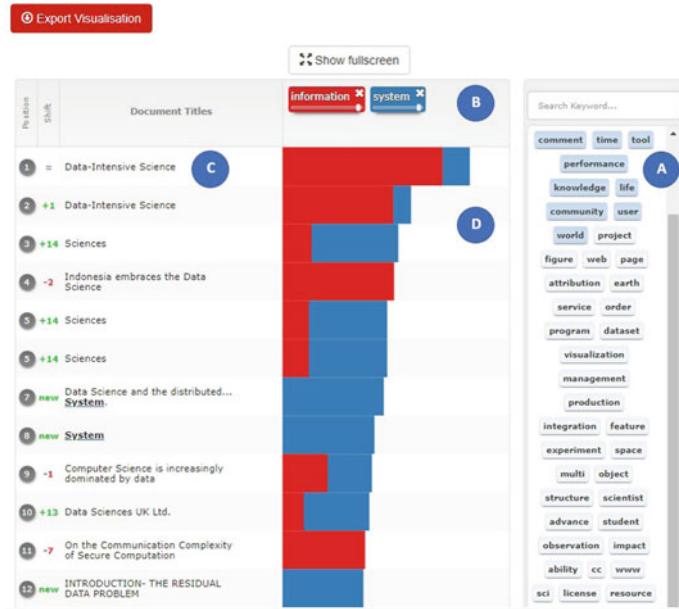


Fig. 5 uRank and its components—(A) tag cloud, (B) tag box, (C) result list, (D) ranking view

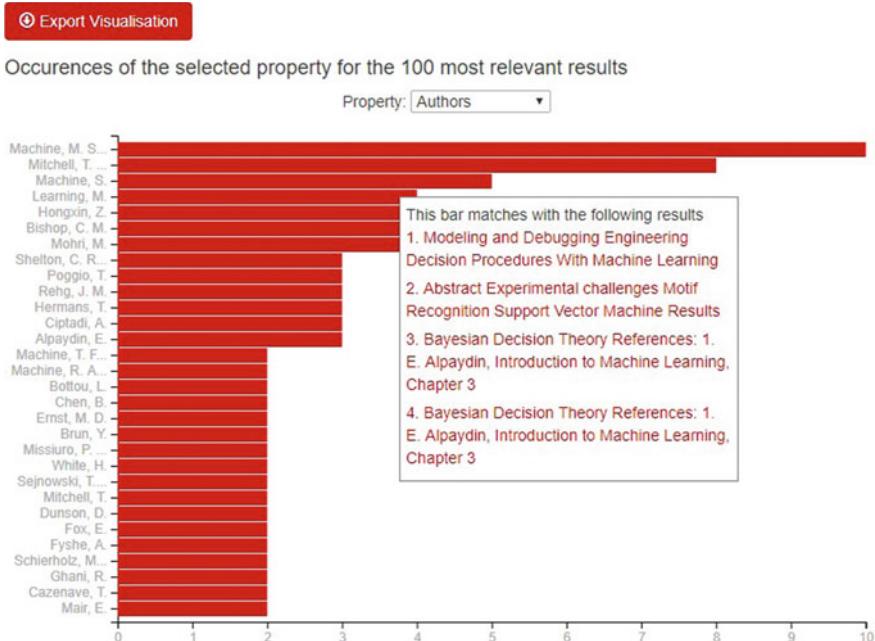


Fig. 6 The Top Properties visualization with the dialogue box showing the result list for a bar of interest

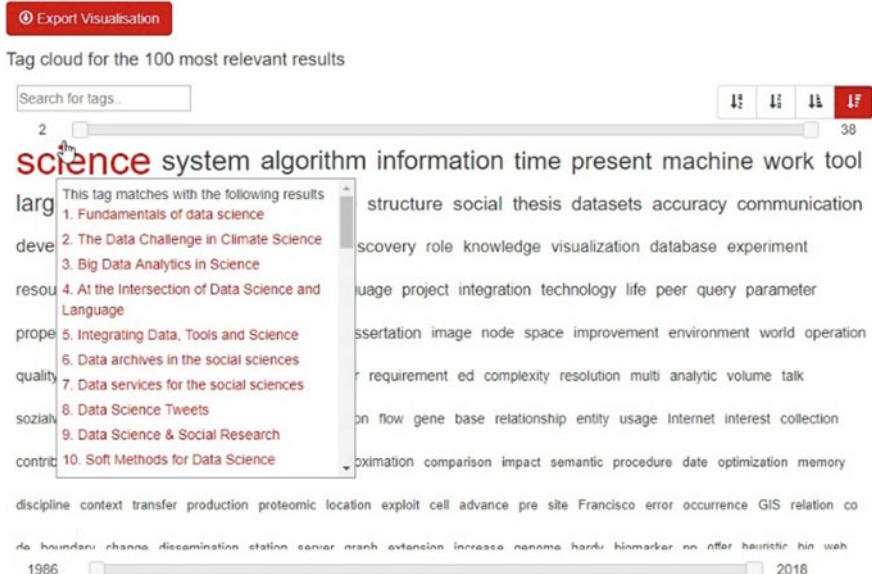


Fig. 7 Tag Cloud visualization with a dialogue box showing the result list for a keyword

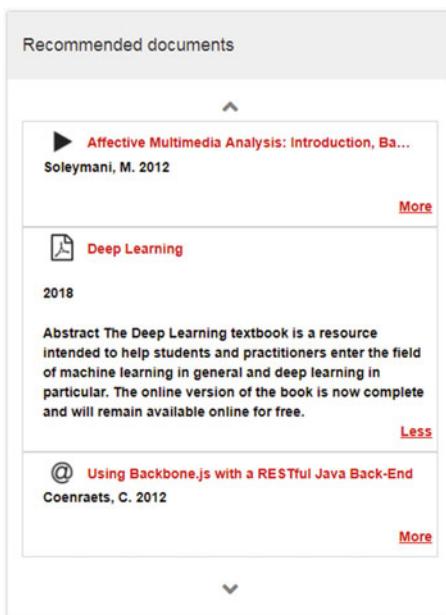
search engine. The Top Properties visualization also supports an export functionality, which exports the current view of the visualization with its search configuration.

Tag Cloud: the Tag Cloud visualization (Fig. 7) retrieves the 100 most relevant results from the search query and displays them by showing the most frequent keywords that occur in the corresponding titles and abstracts. The displayed keywords are initially sorted by their frequency and can be filtered by occurrence, year, or text. Clicking on one of the keywords shows the results associated with this property. The results are sorted in the order provided originally by the search engine.

4.2 Recommender System

The RS widget, depicted in Fig. 8, is part of the search page. It gives users additional suggestions for resources of which they may not be aware. The RS interacts with the search engine, user-interaction tracking, and dashboard (WevQuery), hence bridging the back and front ends of the MOVING platform. To build user profiles, it obtains the search history from the user data previously logged through UCIVIT and then retrieves the documents to suggest from the index, depending on the user's profile. The MOVING RS is based on HCF-IDF (Nishioka and Scherp 2016), a novel semantic profiling approach that can exploit a thesaurus or ontology to provide better recommendations. Further information on the MOVING RS is available elsewhere (Vagliano and Nazir 2019).

Fig. 8 The Recommender System widget suggesting three new items to the user: a video, an article, and a web page (Vagliano and Nazir 2019)



4.3 Communities

Open collaboration and communication are the foundations of open innovation and open science. MOVING communities offer users a powerful tool to organize group collaboration and communities of practice on the MOVING platform (see Fig. 9). MOVING communities are part of the working environment of the platform and offer a range of social technologies with knowledge and information management, including wikis, forums, blog functions, and group news. MOVING communities are based on the project management tools and technologies of the eScience platform on which the MOVING platform is based. The existing eScience modules, which enabled cooperation in closed teams of researchers, were adapted to the goals of the MOVING platform to provide an open innovation environment and foster open collaboration, communication, and knowledge exchange between its users.

Registered users who want to create a new community are offered different options. First, users can create public communities that are visible to everyone in the MOVING platform and can be accessed and edited by anyone interested in the topic. Second, users who want to organize specific project teams or research groups can create private communities that users have to join before they can access and edit content. Private communities are not visible to other users but can be shared with collaborators via email.

The screenshot shows the MOVING platform's interface. At the top, there is a navigation bar with links for Search, Communities, Learning, Contacts, MOOC, My page, and Sign out. Below the navigation bar, a red header bar contains the word "Communities". Underneath this, there is a button labeled "New community" and two sorting options: "by last change (descending)" and "by name (ascending)". A dropdown menu titled "Current view mode: Community" is open, showing four options: All, Privat, Team, and Community, with "Community" being the selected option. Below the dropdown, there are several sections: "Curriculum implementation" (with a sub-note about topics like curriculum implementation, adoption, and adaptation), "Digital Auditing" (with a note about connecting professionals working on EY's Digital Audit), "Eye-tracking in Visual Analytics Recommendations" (with a note about the implementation of Eye-tracking based visual analytics systems for recommendations), and "MOVING MOOC Science 2.0 and open research methods 2019" (with a note about the MOOC showing young academics how to utilize Web 2.0 technologies for research). The background of the main content area has vertical grey bars on the left and right sides.

Fig. 9 MOVING communities

The MOVING CK Editor⁴ enables the creation of formatted text and the integration of multimedia content in HTML pages that are created by users in the MOVING communities. Videos, pictures, GIFs or documents, and social media content from Twitter⁵ and YouTube⁶ can all be easily integrated. Features like the accordion and the option to include expandable items make it easy to structure content in the page. It is a WYSIWYG editor (What You See Is What You Get) so even users that are not familiar with HTML can use it easily to create and edit web-based content within MOVING communities.

The wiki module is useful for creating and collaboratively managing large knowledge repositories with a community. The forum module provides space for

⁴www.ckeditor.com, last accessed on 7 May 2020.

⁵www.twitter.com, last accessed on 7 May 2020.

⁶www.youtube.com, last accessed on 7 May 2020.

The screenshot shows the MOVING MOOC website. At the top, there's a navigation bar with 'MOVING' on the left and 'Register' and 'Sign in' buttons on the right. Below the header is a banner featuring a keyboard background with the 'MOVING MOOC SCIENCE 2.0 AND OPEN RESEARCH METHODS' logo. The main content area has a teal header 'Course Information (#MoMoScience20)'. It lists course details: Start date January 21, 2019; Effort 2-4 hrs/week; Language English; Duration 4 weeks; Level Beginner; Credentials Certificate of participation; and an 'Enrol now' button. Below this is a section titled 'What is this course about?' with a detailed description of Science 2.0 and its principles. A 'Keywords' section follows, listing terms like Science 2.0, Open Science, Open Access, Open Data, OER, Web 2.0 technologies, Creative Commons, open peer review, Altmetrics, and social media. The next section, 'Who is this course for?', describes the target audience as young scholars, post-docs, and students at an advanced level. To the right of the main content is a video player showing a 'MOVING MOOC Science 2.0 Teaser' video. The video thumbnail features the MOVING logo and several circular icons representing different concepts: Science 2.0 & open research, Exchange, Tools & technologies, and open & collaborative. The video player includes standard controls like play, pause, volume, and a YouTube logo.

Fig. 10 MOVING MOOC community

open communication and information exchange—a precondition for open innovation processes. The forum module contains a user rating functionality that allows the community to publicly rate the content of individual forum entries. Users can vote posts and replies up and down, based on the quality of the contribution. The highest-rated input is highlighted to help users find the best response in a thread, and the summarized score for all received votes is shown on each user profile. The ranking functionality helps communities self-organize and peer assess user-generated content. Community administrators can also choose to assign badges to reward users or motivate them to get actively engaged. Badges can be assigned automatically or manually.

The ease of user-generated content creation and integration combined with the social features of MOVING communities open up a wide range of possible applications. Users can organize group work in small project teams, or create open communities around scientific or technical topics to discuss research or ask questions to an expert community. MOVING communities can be organized as an open innovation tool but also as a learning management system, as the following example shows.

One practical application of MOVING communities is the four-week MOVING MOOC (massive open online course) Science 2.0 and open research methods that was organized on the MOVING platform (see Fig. 10).⁷ The MOOC is organized on the platform as a private team community, so that participants have to register to gain

⁷ moving.mz.tu-dresden.de/mooc, last accessed 7 May 2020.

Which Badges can I get in the MOOC and what do I have to do?

In the MOVING-MOOC you can get the following Badges:

MOVLING



The MOVLING-Badge is one badge for all MOVING MOOC participants. All you have to do is to enroll to the MOVING-MOOC. By getting the MOVLING-Badge you show others, that you are interested in Open Science and MOOCs. If you like to promote the MOVING-MOOC backpack the MOVLING-badge and show others this great opportunity to learn and discuss about open Science.

OPEN SCIENCE AFICIONADO



The Open Science Aficionado Badge shows that you:

- understand the concept of Open Science and know the fundamentals of and open research methods
- know how you can use open research methods for scientific collaboration
- understand why and how social technologies improve open scholarly communication and help you to reach a wider audience for your research
- know the open science research workflow and how it relies on the use of digital and social technologies

Fig. 11 MOVING MOOC badges

access to the learning materials and the forums. For each week of the MOOC, we created a sub-community containing learning materials in different media formats as well as weekly assignments. The forums were used to organize group communication and allow users to share their assignment results. A wiki was created and contained additional information about the course, learning goals, and technical details about using the editor or the MOOC badges that users can earn on the course (Fig. 11). Badges are displayed on the user's profile, My page, along with their personal and contact details (profile picture, science field, skills, hometown, institution, email, ORCID⁸).

⁸www.orcid.org, last accessed on 7 May 2020.

Fig. 12 MOVING learning environment

4.4 Learning Environment

MOVING offers a unique combination of working and training features in one platform. The heart of the training programme is the MOVING learning environment. Here, all the learning content is organized and directly accessible to the users. The landing page (Fig. 12) gives an overview of the learning materials including the platform demo videos and video tutorials, the Learning Tracks for Information Literacy 2.0, and the MOVING MOOC that was discussed in the previous subsection, Science 2.0 and open research methods. The platform demos are videos hosted on videolecures.net and are embedded in the learning environment so that users can learn about the different platform features and technologies developed within the MOVING project. Users can improve their data and information literacy as well as digital competences through Learning Tracks for Information Literacy 2.0 (Fig. 13).

4.5 Adaptive Training Support

The ATS (Fessl et al. 2018) comprises two widgets for learning how to search and curriculum reflection.

The Learning-how-to-search (Fig. 14) widget visualizes information about the use of features provided by the MOVING platform. The widget presents to users how they used the features of the platform in a bar chart to motivate them to explore new features and reflect about their usage behaviour. More information about the widget and its evaluation can be found in (Fessl et al. 2019).

The curriculum reflection widget (Fessl et al. 2019) consists of two parts: the curriculum learning and reflection and the overall progress. The first part consists of two main areas. The upper area either contains a learning prompt (suggesting that

The screenshot shows the MOVING platform's start page for Information Literacy 2.0. At the top, there is a navigation bar with links for Search, Communities, Learning, Contacts, MOOC, My page, and Sign out. On the left, a sidebar lists categories: Information & Data Literacy, Communication & Collaboration, Interacting through digital technologies, Sharing through digital technologies, Collaborating through digital technologies, and Content Creation. The main content area is titled "Welcome to the MOVING learning tracks for Information Literacy 2.0." It describes how MOVING helps users improve their data and information literacy and learn about new web-based tools and technologies for communication and collaboration. It also highlights that Information Literacy 2.0 is invaluable for professionals in various fields. Below this, a section titled "Learning Tracks" explains that the curriculum consists of three major tracks: Track 1 Data and Information Literacy, Track 2 Communication & Collaboration, and Track 3 Content Creation. Further down, a "How it works" section details the adaptive training support feature, which offers training hints and learning suggestions in small microlearning units along the three learning tracks. It also mentions that users can navigate directly through the Learning Tracks using the Adaptive Training Support widget. A red "Let's Go!" button is located at the bottom of the main content area.

Fig. 13 Start page of Learning Tracks for Information Literacy 2.0

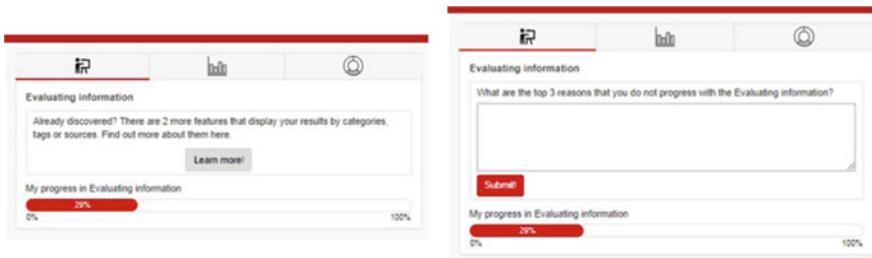
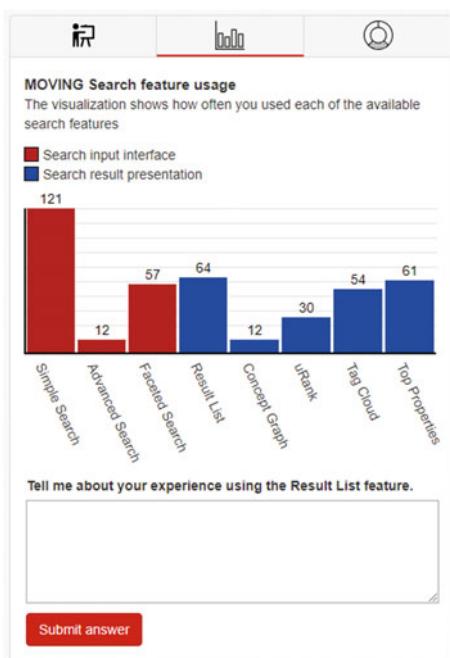
the user learn more about the next topic in the current sub-module) and a button which opens the respective learning unit in a new tab (Fig. 15 left), or it presents a reflective question that motivates the user to think about the current topic of their learning (Fig. 15 right). The user's progress in the current sub-module is displayed at the bottom of the widget.

The overall progress part of the widget shows the user's learning progress through the curriculum using a sunburst visualization. Figure 16 shows that the curriculum is divided into three modules. Each module is represented as a section in the inner circle of the visualization and divided into three sub-modules in the outer circle. Every time a user completes a new learning unit, the percentage in the respective section in the sunburst diagram is updated. Progress in each sub-module is encoded by colour. If the user has not completed any learning units in a sub-module (0%), the respective section will be red. Making progress in a sub-module will turn the section yellow (50%) and completing it will turn the section green (100%).

This is also explained by the legend below the visualization. Moreover, the sections in the sunburst diagram are ordered to mirror the structure of the curriculum. Starting from the top, the sub-modules are completed clockwise, gradually turning the visualization green.

Fig. 14

Learning-how-to-search widget: The tracked features are separated into features of the search input interface and search result presentation

**Fig. 15** Curriculum reflection widget: curriculum learning (left) and reflection (right)

5 Conclusion

In this chapter, we presented the MOVING platform, focusing on the MOVING web application with its search interface and novel results visualizations, community features and learning environment, and components such Adaptive Training Support. These functionalities help users to not only search within and visualize a large multimedia collection using various advanced tools and functionalities, but also to explore the platform more easily, e.g. by showing statistics about their platform use or providing learning guidance. Productive use of the prototype platform in real educational environments, such as the MOVING MOOC, showed how its integrated

Fig. 16 Overall progress widget: The first module was completed and the second module is in progress



training and working environment contributes to making information professionals data-savvy and improving users' information literacy skills.

Acknowledgements This work was supported by the EU's Horizon 2020 programme under grant agreement H2020-693092 MOVING. The mentioned eScience Saxony research network has been supported by the Saxon State Ministry for Science and Art. The Know-Center is funded within the Austrian COMET Programme, Competence Centers for Excellent Technologies, under the auspices of the Austrian Federal Ministry of Transport, Innovation and Technology, the Austrian Federal Ministry of Economy, Family and Youth and by the State of Styria. COMET is managed by the Austrian Research Promotion Agency FFG.

References

- Apaolaza, A., Harper, S., Jay, C.: Understanding users in the wild. In: Proceedings 10th International Cross-Disciplinary Conference on Web Accessibility (W4A'13), Rio de Janeiro, Brazil (2013)
- Apaolaza, A., Vigo, M.: WevQuery: testing hypotheses about web interaction patterns. In: Proceedings ACM Human Computer Interaction 1(EICS), 4:1–4:17. (2017)
- Backes, T.: Effective unsupervised author disambiguation with relative frequencies. In JCDL '18 The 18th ACM/IEEE Joint Conference on Digital Libraries Fort Worth, TX, USA—June 3–7, 2018, edited by Jiangping Chen, Marcos André Gonçalves, and Jeff M. Allen, pp. 203–212. New York: ACM. <http://dx.doi.org/10.1145/3197026.3197036> (2018)
- Backes, T.: The impact of name-matching and blocking on author disambiguation. In: CIKM '18: The 27th ACM International Conference on Information and Knowledge Management Torino,

- Italy—October 22–26, 2018, 803–812. New York: ACM. <http://dx.doi.org/10.1145/3269206.3271699> (2018)
- European Commission: Digital Economy. Online via: <https://ec.europa.eu/digital-single-market/en/economy>, last accessed 13 September 2016 (2016)
- Fessl, A., Apaolaza, A., Gledson, A., Pammer-Schindler, V., Vigo, M.: Mirror, mirror on my search...: Data-Driven Reflection and Experimentation with Search Behaviour. In European Conference on Technology Enhanced Learning, pp. 83–97. Springer, Cham (2019)
- Fessl, A., S'imiric, I., Barthold, S., Pammer-Schindler, V.: Concept and development of an information literacy curriculum widget. In Conference on Learning Information Literacy around the Globe (2019)
- Fessl, A., Wertner, A., Pammer-Schindler, V.: Digging for gold: motivating users to explore alternative search interfaces. In: Lifelong Technology-Enhanced Learning. pp. 636–639. Springer International Publishing, Cham (2018)
- Galanopoulos, D., Mezaris, V.: Temporal lecture video fragmentation using word embeddings. In: Proceedings MMM2019. Springer LNCS vol. 11296, pp. 254–265 (2019)
- Köhler, T., Pscheidla, D., Scherp, A., Koschtial, C., Felden, C., Neumann, J.: Moving research methodology toward eScience. Paper Presentation Track A: Online Research Methodology, General Online Research 2016; Dresden 02.-04.03. (2016)
- Köhler, T., Scherp, A., Herbst, S., Wiese, M., Mezaris, V.: Data driven online research. Potential specifications in relation to user needs; International Science 2.0 Conference; Köln 02.-03.05. (2016)
- Nishioka, C., Scherp, A.: Profiling vs. time vs. content: what does matter for top-k publication recommendation based on twitter profiles? In JCDL '16, pp. 171–180. ACM (2016)
- Pscheidla, D., Köhler, T., Mohamed, B.: What's your favorite online research tool? Use of and attitude towards Web 2.0 applications among scientists in different academic disciplines. In: Marsden, C. Tassiulas, L.: Proceedings of the 1st International Conference on Internet Science; Brussels, Sigma Orionis (2013)
- Pscheidla, D., Minet, C., Herbst, S., Albrecht, S., Köhler, T.: Use of social media and online-based tools in academia. Results of the Science 2.0-Survey 2014; Dresden, TUD Press. SLUB: <http://nbn-resolving.de/urn:nbn:de:bsz:14-qucosa-191110> (2015)
- Scherp, A., Pscheidla, D., Köhler, T., Wiese, M., Nishioka, C., Mezaris, V., Collyda, C.: MOVING: training towards a society of data-savvy information professionals to enable open leadership innovation. In: 13th European Semantic Web Conference (ESWC 2016; Anissaras 29.05.-02.06. (2016)
- Tzelepis, C., Mezaris, V., Patras, I.: Linear maximum margin classifier for learning from uncertain data. IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 40, no. 12, pp. 2948–2962, December (2018)
- Vagliano, I., Nazir, S.: Recommending Multimedia Educational Resources on the MOVING Platform. In: Proceedings of the 8th Workshop on Bibliometric-enhanced Information Retrieval, CEUR-WS.org. pp. 148–158 (2019)
- Vagliano, I., Günther, F., Heinz, M., Apaolaza, A., Bienia, I., Breitfuss, G., Blume, T., Collyda, C., Fessl, A., Gottfried, S., Hasitschka, P., Kellermann, J., Köhler, T., Maas, A., Mezaris, V., Saleh, A., Skulimowski, A.M.J., Thalmann, S., Vigo, M., Scherp, A.: Open innovation in the big data era with the MOVING platform. IEEE Multimedia **25**(3), 8–21 (2018)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



CLARIN-D: An IT-Based Research Infrastructure for the Humanities and Social Sciences



Gerhard Heyer and Volker Böhlke

Abstract The paper discusses the idea of bridging the gap between computer sciences and the humanities by referring to an e-humanities infrastructure that provides tools and services for well-defined and frequently encountered tasks. The main goal of this infrastructure is to enable researchers in the humanities and social sciences to better exploit their potential by reusing available digital resources, and thus to increase the efficiency of e-humanities projects. CLARIN-D is an example of such a research infrastructure. The paper provides a brief overview of the basic principles and services of the CLARIN-D infrastructure, such as metadata harvesting, federated content search, and chaining Web services.

Keywords Digitalization · Humanities · CLARIN-D

1 Introduction

To date, computer science and the humanities have taken different approaches to working methodologies, rather than focusing on the potential synergies. However, recent advances in digitizing historical texts, and the search and text-mining technologies for processing these data, indicate an area of overlap that bears great potential. For the humanities, the use of computer-based methods may lead to more efficient research (where possible) and raise new questions that could not have been dealt with otherwise. For computer science, turning to the humanities as an area of application may pose new problems that require rethinking the approaches hitherto favored by computer science. As a result, new solutions may develop that help to advance computer science in other areas of media-oriented application. At present, most of these solutions are restricted to individual projects and do not allow the digital humanities community to benefit from other advances in computer science,

G. Heyer (✉) · V. Böhlke
Leipzig University, Leipzig, Germany
e-mail: heyer@informatik.uni-leipzig.de

V. Böhlke
e-mail: boehlkev@informatik.uni-leipzig.de

like service engineering. Hence, in this paper we attempt to spell out in detail the idea of an infrastructure for e-humanities. Focusing on the notion of reusability of data and algorithms such as morphological annotation and part-of-speech (POS) tagging, we sketch how a loosely coupled infrastructure based on Web services and a service-oriented architecture (SOA) can help the humanities to better exploit their potential by reusing available digital resources, and thus increase the efficiency of e-humanities projects. As an example, we present a rough overview of Common Language Resources and Technology Infrastructure D (CLARIN-D), a Web-based research infrastructure for the humanities and social sciences.

2 The Impact of Digitization in the Humanities—From Digital Humanities to E-Humanities

To the extent that applications of computer science have always led to a replacement of analog by digital media and processes, digital media and processing models are having an increasing impact on traditional work flows based on analog media in the humanities and social sciences. The interdisciplinary combination of methods from computer science and traditional humanities with large amounts of digital data and advanced tools for processing these is commonly known as e-humanities (cf. McCarty 2005). Although there is no standard definition of terms yet, e-humanities in a broader sense are concerned with the intersection of computing and the humanities in the eScience paradigm, and thus pertain to any digitized data that are subject to investigation in the humanities and the social sciences, such as text, images, and objects (e.g., in archeology).

For the humanities, the use of computer-based methods may lead to more efficient research (where possible) and raise new questions that could not have been dealt with otherwise. For computer science, turning to the humanities as an area of application may pose new problems that lead to rethinking approaches hitherto favored by computer science. As a result, new solutions may develop that help to advance computer science in other areas of media-oriented application. By focusing on text as the main data type in the humanities, we can highlight the benefit that can be gained from the combination of digital document collections and new analysis tools from computer science, mainly derived from information retrieval and text mining. In this way, all kinds of sciences that work with historical or present-day texts and documents are enabled to ask completely new questions and deal with text in a new manner. These methods impact in the following ways:

- qualitative improvement of the digital sources (standardization of spelling and spelling correction, unambiguous identification of authors and sources, marking of quotes and references, temporal classification of texts, etc.);
- the quantity and structure of sources that can be processed (processing of very large amounts of text, structuring by time, place, authors, contents and topics, comments from colleagues and other editions, etc.);

- the kind and quality of the analysis (broad data-driven studies, strict bottom-up approach using text-mining tools, integration of community networking approaches, etc.).

At present, most of these solutions are restricted to individual projects and do not allow the scientific community in the e-humanities to benefit from advances in other areas of computer science. We therefore wish to distinguish between two important aspects of e-humanities:

1. creation, dissemination, and use of digital repositories;
2. computer-based analysis of digital repositories using advanced computational and algorithmic methods.

While the first has originally been triggered by the humanities and is commonly known as digital humanities, the second implies a dominance of computational aspects and might thus be called computational humanities.

A practical consequence of this distinction in organizational terms would be to set up research groups in both scientific communities, computer science, and the humanities. The degree of mutual understanding of research issues, technical feasibility, and scientific relevance of research results will be much higher in the area of overlap between computational and digital humanities than with any intersection between computer science and the humanities.

To empower the humanities to enter into a substantial and mutually beneficial dialog with computer science, however, a research infrastructure is needed that enables researchers in the e-humanities to reuse distributed digitized data and tools for their analysis as much as possible. To use such computational methods, an individual researcher can proceed by employing two strategies, depending on his or her own degree of computer literacy. One strategy is the individual software approach. Given a selection of digital text data, the research question is transferred into a set of issues and methods that can be dealt with by a number of individual programs. This approach allows for highly dynamic and individual development of research issues. It requires, however, a high degree of software engineering know-how. The other approach is to use standard software. For well-defined and frequently encountered tasks, an e-humanities infrastructure will offer solutions that provide the users with data and analysis tools that are well understood, have already delivered convincing results, and can be learned without too much effort (cf. Boehlke et al. 2013).

Both approaches are interdependent. Probably good solutions in one domain of text-oriented humanities can be transferred to other domains by just using different kinds of text. A good infrastructure must be capable of making such solutions accessible as best practices.

3 CLARIN-D—An Infrastructure for Text-Oriented Humanities

Research infrastructures are concerned with the systematic and structured acquisition, generation, processing, administration, presentation, reuse, and publication of content. Content services make available the resources and programs needed for that. Public digital text and data resources are linked together and made accessible by common standards. New software architectures integrate digital resources and processing tools to develop new and better access to digital contents. CLARIN-D¹ is part of CLARIN Europe, which recently² became an independent legal entity according to the ERIC³ statutes. CLARIN-D is primarily designed as a distributed, center-based project (cf. Wittenburg et al. 2010). This means that centers are at the heart of an infrastructure that aims at providing consistent data services. Different types of resource centers form the backbone of the infrastructure, provide access to data and metadata, and/or run infrastructure services. Access to data, metadata, and infrastructure services is usually (but not solely) based on Web services and Web applications. The protocols and formats of infrastructure services (like persistent identifiers or metadata systems and standards that are of interest to the CLARIN initiative on the European level) have been agreed upon in the preparatory phase of the project. Additional infrastructure or discipline-specific services are built upon those basic infrastructure services. The usage of general services like registering and resolving persistent identifiers is not limited to CLARIN itself. Other infrastructure initiatives can and do use such services.

Important metadata on CLARIN centers—for example, technical access points, standards and contact information—is stored in a centralized centers registry that acts as a starting point for service users and enables the automation of various procedures, such as monitoring and visualizing the state of all infrastructure services.

4 Metadata, Citation, and Search

In CLARIN, metadata is usually represented in a component metadata infrastructure (CMDI).⁴ The underlying technology of CMDI is XML-Schema (components, profiles), XML (instances), and REST (component registry). CMDI addresses the problem of various specialized metadata standards used for specific purposes by different research communities. Instead of introducing yet another standard, CMDI

¹<http://de.clarin.eu>.

²<http://ec.europa.eu/research/index.cfm?pg=newsalert&lg=en&year=2012&na=na-290212-1>.

³http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=eric.

⁴<https://www.clarin.eu/content/component-metadata>.

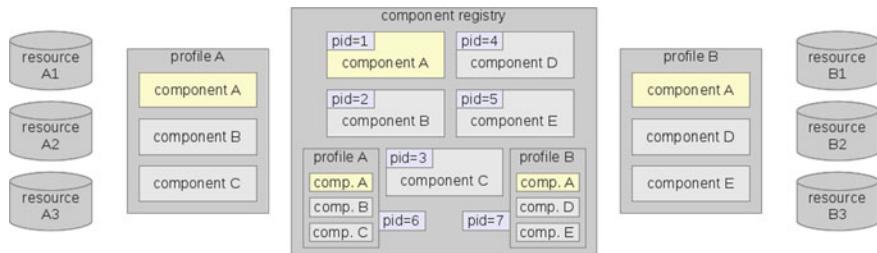


Fig. 1 Components, profiles, and component registry

aims at describing and reusing, and (when used in combination with ISOcat⁵) interpreting and supporting the integration of existing metadata standards. CMDI components act as basic building blocks that define groups of field definitions. These components can be combined into profiles that define the syntax and semantics of a certain class of resources and act as blueprints for metadata instances describing items of this class. These components are managed in a component registry, which allows users to archive and share existing components, thus enabling their reuse (see Fig. 1). Through this approach, CMDI supports the free definition and usage of metadata standards dedicated to specific use cases. As long as metadata is stored in XML, CMDI is able to “embrace” other standards. By combining the data itself with semantic information stored in the ISOcat data-category registry, CMDI forms a solid basis for using sophisticated exploration and search algorithms.

Metadata is the backbone of the infrastructure and publicly available in CLARIN from the resource centers (cf. Boehlke et al. 2012) via the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH).⁶ The openness of metadata is important to CLARIN since it guarantees high visibility of the provided resources in the research community.

OAI-PMH is a well-established standard and is supported by numerous repository systems like DSpace⁷ and Fedora.⁸ The OAI-PMH protocol is based on REST and XML and provides the ability to do two things. It offers full access to the metadata provided by the resource centers and allows for selective harvesting of metadata (see Fig. 2) for search portals like the Virtual Language Observatory (VLO). The VLO enables users to perform a faceted search on the metadata that was harvested from the repositories of all CLARIN centers. By using the information stored in the ISOcat data-category registry (cf. Kemps-Snijders et al. 2008) and the CMDI profiles (see Fig. 3) associated to the CMDI metadata instances, the VLO map information is stored in these instances onto a predefined set of facets (see Fig. 4). The VLO also supports the extraction and usage of additional, CLARIN/CMDI-specific, metadata

⁵ <http://www.isocat.org/>.

⁶ <http://www.openarchives.org/pmh/>.

⁷ <http://www.dspace.org/>.

⁸ <http://fedora-commons.org/>.

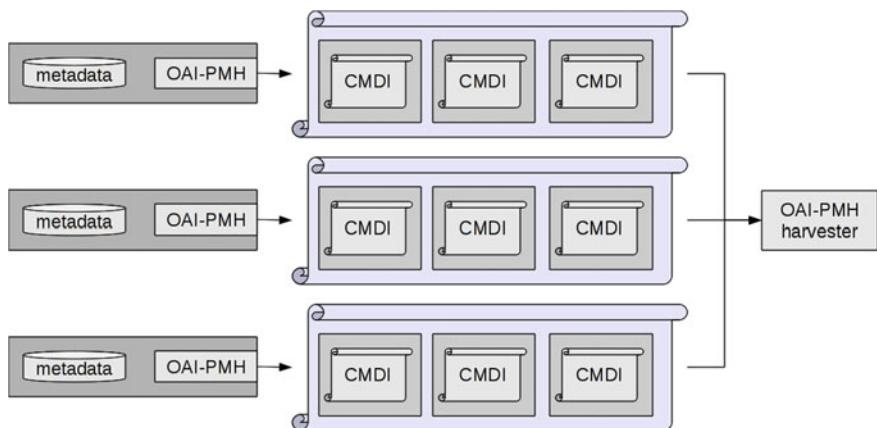
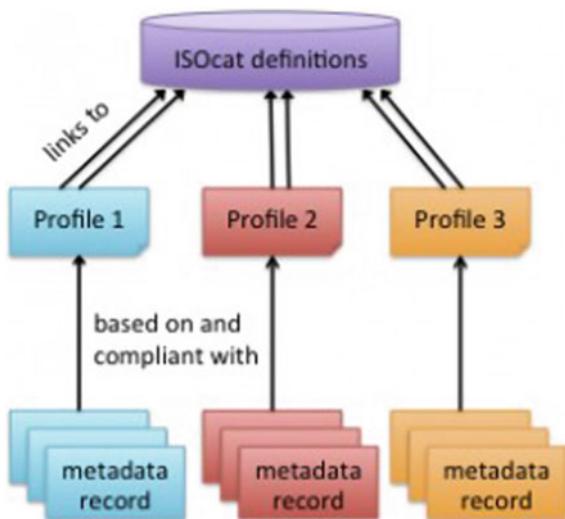


Fig. 2 OAI-PMH harvesting

Fig. 3 Metadata records, profiles, and ISOcat. *Source* https://www.clarin.eu/sites/default/files/styles/openscience_3col/public/cmdi-overview.png



such as ResourceProxy (e.g., link to download, dedicated search portal) and federated content search (FCS) interfaces.

CLARIN also provides support for content-based search. The CLARIN-D FCS⁹ is based on Search/Retrieval via URL (SRU) and Contextual Query Language (CQL) and allows users to perform a CLARIN-wide search over all repositories that offer a FCS interface by using a simple Web application. This Web application and external applications send a request to an aggregator service. This service first queries a repository registry and searches for compatible interfaces. The initial query is then

⁹<https://www.clarin.eu/content/federated-content-search>.



Fig. 4 VLO

sent to all of these interfaces and the individual results are aggregated and sent back to the user or application (see Figs. 5 and 6). Since CLARIN is designed as an open infrastructure, third-party content providers may easily plug their own repository and FCS interface into this process by registering it to the CLARIN repository registry.

Web services in CLARIN are also described via CMDI (which may very well contain a link to a WSDL file). If more specific metadata is provided (i.e., the information enforced by a certain CMDI profile is given), these Web services can be used in a workflow system called WebLicht (cf. Hinrichs et al. 2010). WebLicht allows users to build and execute chains of Web services by analyzing the metadata available for each service and ensuring that the format of the data is compatible; that is, that the output of a predecessor service satisfies the specification of a successor service.

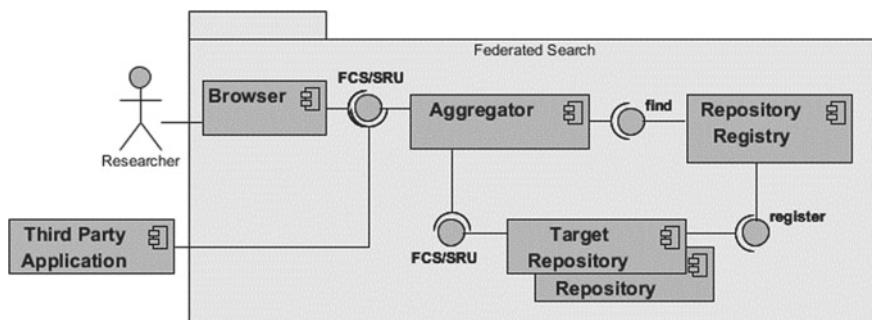


Fig. 5 Federated content search. Source http://www.clarin.eu/sites/default/files/FCS_components.png

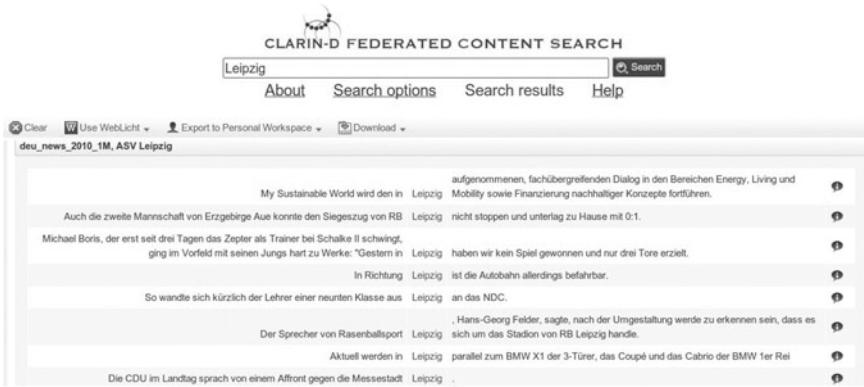


Fig. 6 CLARIN-D FCS Web application

Table 1 Example input specification for a POS tagger Web service

Format	MyFormat
Input	text = UTF-8 language = German tokens = present

When thinking about interchanging neuro-linguistic programming (NLP) data like text, there are several established standards defining how texts can be encoded and how annotations like POS tags may be added. These standardization efforts are supported by WebLicht, hence the following interface definition of a Web service compatible with WebLicht:

- the format used is TCF (or TEI¹⁰ P5, etc.);
- the document contains German text and is annotated with POS tags;
- the POS tags are encoded according to the STTS¹¹ tagset.

A complete interface definition of a WebLicht Web service consists of two identically structured specifications for input and output. Each of these specifications defines the format of a document that is used to represent the data. Additionally, a set of pairs of parameter types is mandatory to invoke the service for the input specification, or is computed and added by the service for the output specification. Each of these parameter types is bound to a standard definition, which binds it to a standardized encoding of the information.

Tables 1 and 2 give example input and output specifications of a POS tagger Web service. This service consumes documents that contain German text that was split

¹⁰An organization which maintains a format for digital text representation. See <http://www.tei-c.org/index.xml>.

¹¹Stuttgart Tübingen Tagset. See <http://www.sfb441.uni-tuebingen.de/a5/codii/info-stts-en.xhtml>.

Table 2 Example output specification for a POS tagger Web service

	Format	MyFormat
	Output	POS tags = STTS

into tokens encoded in an imaginary format. It produces a document of the same format by adding POS tags based on the STTS tagset.

The chaining algorithm of WebLicht (cf. Boehlke 2010) is based on the idea that NLP services usually consume a document of a well-defined standard and will also return such a document. The successful invocation of a service for an input document hence depends on which information is available in that document. A POS tagger Web service may only work if sufficient information on sentence and token boundaries is available, while a named entity recognizer (NER) requires appropriate POS tags. Therefore, the standard used for the input document needs to allow for a representation of this kind of information, and, of course, this information needs to be present in the input document itself. This fact is also represented in the interface definition. Thus, for service chaining to work, it must be ensured that this information is available by using a type checker on each step of a chain.

This check can be done when building the chain, since all the necessary information is already available. Based on a formal Web service description according to the proposed structure, a chaining algorithm, which is basically a type checker, can be implemented. A service can be executed if the previous services in the chain meet the following constraints:

- the format specified in the output is equal to the format specified in the input specification of the service;
- every parameter-type/standard pair defined in the input specification needs to be one of the pairs in the output specifications of services which have been executed (or scheduled for execution previously in the chain, if we stay on build time).

These two constraints are of course a simplification. But in many simple cases, an algorithm like this will be sufficient. A short and simplified example of the chaining logic is given in Figs. 7 and 8, which show part of a chain consisting of Web services A (a tokenizer) and B (a POS tagger). In Fig. 7, Service A can be executed since all constraints defined in its input specification are met. The format of the input document is compatible and its content fulfills the requirements because it contains German text encoded in UTF-8. The tokenizer segments the text into sentences and tokens. After its execution, this information is added to the resulting output document. Service B is checked against this updated knowledge about the content of the output document of Service A (see current metadata in Fig. 8). Service B is compatible since all of its input requirements, format and parameters, are available in the output document of Service A.

Fig. 7 Tokenizer service specification

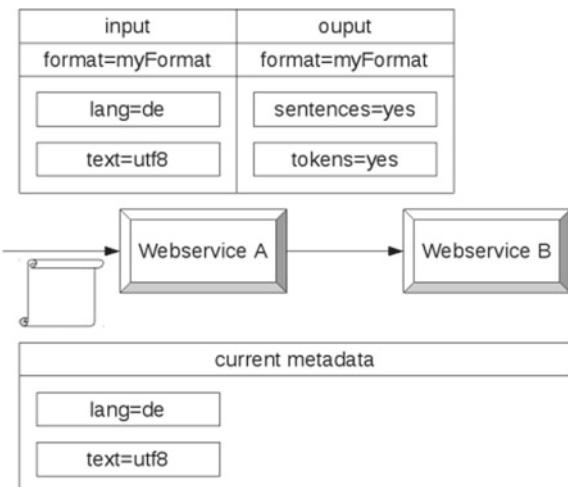
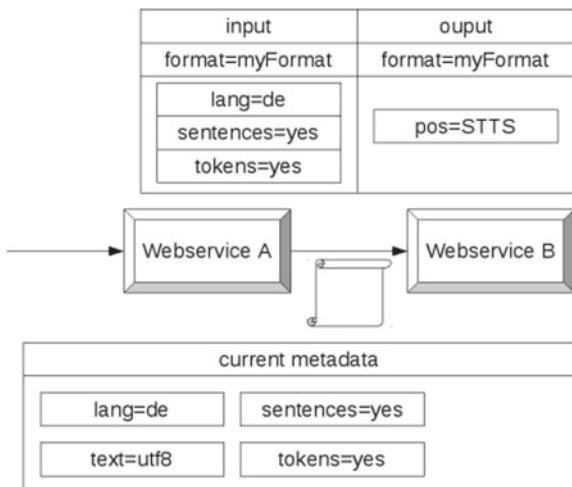


Fig. 8 POS tagger service specification



5 Summary and Conclusion

Research infrastructures for the humanities can help to share digital resources and content services. In particular, they can help researchers in the digital humanities to save time and effort when developing software to deal with specific research issues, while the development of such infrastructures and their key software components is a software engineering task that increasingly poses interesting and challenging research problems for computer scientists. In this paper, we have presented the European Strategy Forum on Research Infrastructures (ESFRI) project CLARIN and some of its key elements as a research infrastructure for the humanities. In detail,

we have presented component metadata infrastructure as a means for unifying metadata descriptions of linguistic resources in the humanities. Based on these metadata, we have also shown how Web services can be built that share data and algorithms in the research infrastructure. Both aspects are closely related: The content-driven use of digitized data and software tools in a specific application scenario in the humanities, and the software and service engineering issues relating to an efficient research infrastructure in the humanities. These two aspects, content and service, clearly need to complement each other in order to establish a culture of best practice in the e-humanities.

References

- Boehlke, V.: A Generic Chaining Algorithm for NLP Webservices. LREC (2010)
- Boehlke, V., Compart, T., Eckart, T.: Building up a CLARIN resource center—step 1: Providing metadata. In: Workshop on Describing Language Resources with Metadata (2012)
- Boehlke, V., Heyer, G., Wittenburg, P.: IT-based research infrastructures for the humanities and social sciences—developments, examples, standards, and technology. Inf. Technol. **55**(1): 26–33 (2013)
- Hinrichs, M., Zastrow, T., Hinrichs, E.: WebLicht: web-based LRT services in a distributed eScience infrastructure. In: Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10). Valetta, Malta (2010)
- Kemps-Snijders, M., Windhouwer, M., Wittenburg, P., Wright, S.E.: ISOcat: Corralling Data Categories in the Wild. In: European Language Resources Association (ELRA) (ed.), Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC 2008), Marrakech, Morocco, May 28–30 (2008)
- McCarty, W.: Humanities Computing. Palgrave, Basingstoke, UK (2005)
- Wittenburg, P., Bel, N., Borin, L., Budin, G., Calzolari, N., Hajicova, E., Koskenniemi, K., Lemnitzer, L., Maegaard, B., Piasecki, M., Pierrel, J.-M., Piperidis, S., Skadina, I., Tufis, D., Van Veenendaal, R., Várdi, T., Wynne, M.: Resource and service centers as the backbone for a sustainable service infrastructure. In: Calzolari, N., Maegaard, B., Marian, J., Odijk, J., Choukri, K., Piperidis, S., et al. (eds.), Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10), pp 60–63. European Language Resources Association (ELRA) (2010)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Toward Process Variability Management in Online Examination Process in German Universities: A State of the Art



Maryam Heidari and Oliver Arnold

Abstract In contemporary organizations, multiple variants of the same business process are often considerable. Such business process variability has caused considerable challenges, both while modeling processes and in their execution. In order to develop a new approach to managing process variants, or extend an existing one, in this research, we review the state of the art in a particular area: online examination processes. We show to what extent variability should be considered in exam processes, whether this is due to special legal restrictions and regulations, different exam frameworks, or even different technical infrastructure. This could be the foundation for developing an approach to managing process variability in the field of e-assessment. Initial findings indicate that examination processes have many similarities, but also considerable differentiation. Therefore, there an appropriate model needs to be developed in order to manage variability in e-assessment and the developed approach must then be validated in identifying faculties. This paper constitutes a first step in this direction.

Keywords Process variability · Online examination · E-assessment process model · Accreditation

1 Introduction

In today's dynamic world, there are often multiple variations of identical business processes. Rosemann and colleagues noted, for instance, that SAP offers 27 different industry solutions with corresponding business process reference models (Rosemann and van der Aalst 2007). These models usually include decisions in the workflow, which could be made before executing process instances. It is impossible for

M. Heidari
TU Bergakademie Freiberg, Freiberg, Germany
e-mail: maryam.heidari@bwl.tu-freiberg.de

O. Arnold (✉)
Westsächsische Hochschule Zwickau, Zwickau, Germany
e-mail: oliver.arnold.1@fh-zwickau.de

both variants of such decisions to coexist in a certain domain or process context. However, conventional modeling approaches do not offer the opportunity to differentiate between such decisions and regular decisions during the runtime of a process instance. An important element of controlling the variability in business process models is to separate the usual runtime decisions from decisions at configuration time, called variation points. The results of such steps are complex artifacts. The number of artifacts makes the manageability of related workflows more complex. Based on the reviewed literature, organizations take different approaches to managing process variability (Ayora Esteras 2012). The existing approaches have limitations in terms of supporting an entire set of elements like control flow, rules, and legal regulations during the construction and execution of business processes.

For this paper, we chose an educational field as an example of process variability, in order to observe effects and causes of variability. The goal is to have a comprehensive overview to address the problem of process variability in online examination processes at German universities and show the necessity of managing it through an appropriate business process model. To achieve this goal, we set out the state of the art in research regarding the process variability of online examinations from different perspectives. This could be the basis for developing a process model to manage existing variability in this field. We evaluate existing approaches and concepts in the context of e-assessment in the literature and clarify current accreditation processes in educational fields. This will help to identify to what extent existing examination procedures reflect variability and demonstrate the necessity of developing a unified e-assessment model to cover all variability in the learning-teaching process.

The paper is organized as follows: After illustrating motivation through existing studies in the following section, the research method is explained. After exploring the literature and data collected in the identified domains, the results are evaluated. Finally, the need for further work is explained in the conclusion.

2 Motivation

In this section, we have a close look at process variability and its challenges, identifying the importance of variability management in the organizations. Basically, process models capture an organization's activities in achieving certain business goals. The aim is to better understand the process, its implementation, and its execution in a workflow (Becker et al. 2013). However, there are a lot of possible variants for one process. Such business process variability creates considerable challenges in process modeling and execution:

- the variants may be modeled in a highly redundant way, so there are many identical or similar parts;
- there is no strong relation between the variants, so there is no support for automatically combining existing variants;

- manual modeling in every single process variant would be time-consuming and error-prone.

In recent years, the proper management of business process variability has been the subject of numerous scientific studies. A very comprehensive survey article about business process variability can be found in Valenca et al. (2013), which describes more than 80 primary sources. Based on this study, significant numbers of variability approaches exist, where each one addresses different issues in terms of process variability. Valenca and colleagues observed 57 new approaches to different aspects of variability in processes (Valenca et al. 2013). They divided these references into five categories: business process configuration, to capture an instance of the reference model; business process correctness, to semantically support correction of the process model; business process flexibility, to change process models fast and easily; business process modeling, to visualize variability in process models; business process similarities, to investigate differentiations between business process models. It is argued that only 30% of solutions are practically evaluated through case studies and surveys, especially of industry: The lack of empirical studies in process variability is considerable, with implications for executing process variability (Valenca et al. 2013).

In the public sector, as in business, processes have a lot in common, but significant differences due to the local conditions and legal regulations are considerable. Vogelaar and colleagues analyze and compare the different processes of ten Dutch municipalities which are found to vary in terms of classical standardization processes (Vogelaar et al. 2012).

In education, Arnold and Laue studied controllability of variability in examination process models (Arnold and Laue 2014). They investigated six different courses at three different universities in the German Federal State of Saxony, to achieve better comparability. Based on this research, considerable variability in examination processes exists, even in one university between different fields of study. The authors tried to provide a solution based on existing variability approaches, in order to manage examination processes. They argued that appropriate process variability modeling requires modeling skills and significant experience in the identified domain (Arnold and Laue 2014).

We focus on online examination processes in higher education, presenting the state of the art in three different domains and observing existing process variability in order to gain a comprehensive overview of process variability in this field, highlighting the need for this to be managed.

3 Research Method

The purpose is to evaluate the existing variability in higher educational e-assessment processes, as a basis for further research into variability management in this field.

The state of the art is identified in five phases (Cooper 1998):

1. Problem formulation: The academic goal and research relevance are defined;
2. Literature search and data collection: Literature and data related to the problem formulation are identified;
3. Literature evaluation: The acquired literature is assessed for relevance and categorized;
4. Interpretation: The results are analyzed and interpreted;
5. Presentation: The results are presented in a suitable fashion.

Assessments are an important part of the educational cycle (Ferrão 2010) and have a great impact on the learning process. They provide valuable information about the effectiveness of a study course in increasing the students' knowledge (Primiano et al. 2004). An appropriate assessment process is not only important in terms of teaching and learning, but also for accreditation processes and educational standards (Ferrão 2010). Recent developments in e-learning can be seen as an accelerator to developing e-assessment alternatives. It is therefore becoming more important to develop methods for e-assessment and to gain feedback on learning and teaching (Sangi and Malik 2007). Furthermore, Dermo has shown that e-assessment can offer different forms of assessment with immediate feedback to both students and lecturers, so it can be recognized as a complementary tool in the learning framework (Dermo 2009). Exam regulation documents, which are the basis for accreditation processes in higher educational institutions, have a lot in common. But in some points, they differ from one university to another or even from one course to another within the same institution. Therefore, there is variability in assessment processes, which is an obstacle to developing a unified process model for e-assessment.

- We reviewed the state of the art in three identified domains within e-assessment:
- IT-related approaches;
- designing study courses and e-assessment concepts;
- accreditation process.

In the following, the results of our literature review and data collection in each domain are explained separately.

3.1 Literature Search and Data Collection in Three Domains

3.1.1 Domain: IT Approaches

Data sources: AISel and EBSCO

Research period: 2000–2013

Search terms and keywords: e-assessment, education, online examination, e-test, computer-based exam (in abstract and title)

Number of related articles:

AISel: 51 after reviewing and removing doublets and non-related articles: 14

AISel: 51 after reviewing and removing doublets and non-related articles: 14

Table 1 Pedagogical aspects of e-assessment

Pedagogical issues	References
Evaluation of the impact of e-assessment on learning processes (strengths and weaknesses)	Becker et al. (2008, 2013), Braun (1998), Coyle (2009), Dermo (2009), Dittman and Deokar (2008), Ferrão (2010), Gruhn and Laue (2007), Hodgson and Pang (2012), Impagliazzo et al. (2002), Johnson-Glenberg (2010), and Jordan and Mitchell (2009)
Formative e-assessment (feedback)	Charoen (2009), Davis et al. (2001), El-Ashy (2006), Hall et al. (2010), Hallerbach et al. (2010), Hallerbach et al. (2008), Hodgson and Pang (2012), and Irani et al. (2000)
Collaborative e-assessment	Ayora Esteras (2012), Boyle (2010), and Davis et al. (2001)
Different types of e-assessment and questions	Coyle (2009), Gorgone (2006), Gruhn and Laue (2007), and Jordan and Mitchell (2009)

Table 2 Technical aspects of e-assessment

Technical issues	References
E-assessment tool evaluation (strengths and weaknesses)	Accountants find e-assessment ... (2011), Boyle and Hutchison (2009), Braun (1998), Charoen (2009), Cooper (1998), Dascalu and Bodea (2010), Davis et al. (2001), Davis (2007), Dermo (2009), Gorgone (2006), and Impagliazzo and Gorgone (2002)
E-assessment implementation (challenges: validity, security, task assessment, adoption, etc.)	Attali and Burstein (2006), Braun (1998), Campbell (2008), Daly et al. (2010), Ferrão (2010), Impagliazzo and Gorgone (2002), Irani et al. (2000), Jacob et al. (2006), Johannsen and Leist (2012), Johnson-Glenberg (2010), and Jordan and Mitchell (2009)

EBSCO: 62 after reviewing and removing doublets and non-related articles: 18

Based on the reviewed articles, a main classification can be recognized in the context of e-assessment in higher education:

- pedagogical issues (educational view)
- technical issues

Different kinds of terms and concepts are used, based on pedagogical and technical approaches: Each one addresses one or more aspects of e-assessment. These issues are summarized in the following Tables 1 and 2.¹

Most of these references include multiple issues from the technical and pedagogical perspectives. These issues are connected and cannot be separated.

¹Numbers in brackets refer to the references in Appendix 1.

Jacob and colleagues deploy an e-assessment tool, the Black Board Learning System (BBLs), as a comprehensive e-learning software to facilitate continuous assessment and evaluate its effects on learning processes (Jacob et al. 2006). It reveals that the biggest advantage of e-assessment in this system is immediate feedback, which bolsters the formative assessment.² One weakness of the system is the lack of the automatic evaluation of essay-writing exams. Kehily analyzed the impact of a Web-based e-learning platform that can support effective teaching (a course management system for lecturers) and formative assessment (a computer-assisted learning tool for students) in a case study (Kehily 2011). Venkatraman developed a four-step student-centered approach to an effective e-learning process and, in a case study of information system (IS) courses, evaluated this approach for different assessment methods including individual, group, peer, and self-assessment (Venkatraman 2007). These four steps are:

1. understanding the students' learning style and their learning expectations;
2. identifying suitable assessment models;
3. designing a set of assessments;
4. evaluating the impact of the assessment on the learning process.

Dermo evaluates the possible risks in planning e-assessments such as computer stress, fairness of choosing questions randomly from a bank, accessibility, and the contribution of e-assessment to students' learning, through six dimensions in a case study (Dermo 2009). These six dimensions are: affective factors, reliability and fairness, validity, security, practical issues, and teaching and learning terms, which are a mixture of pedagogical and technical issues. Daly and colleagues argue that existing e-assessment solutions focus on developing technical and infrastructural issues more than educational aspects (Daly et al. 2010). McCann identifies different factors which affect real implementations of e-assessment systems based on two IS theories: Roger's theory³ and Eckel and Kezar's theory⁴ (McCann 2010).

3.1.2 Domain: Designing Study Courses and E-Assessment Concepts

Data sources: German university homepages

Research period: 2000–2013

Search terms and keywords: e-assessment, online examination, project, computer-based exam, e-exam, e-test

²Formative assessment encourages deeper engagement with learning and is a motivation and progressive force in learning. The key element of formative assessment is feedback.

³It identifies five variables to demonstrate how and why new ideas are adopted: relative advantage, compatibility, trialability, observability, and complexity (McCann 2010).

⁴It identifies five core strategies that explain change across institutions: senior administrative support, collaborative leadership, flexible vision, staff development, and visible actions (McCann 2010).

In this domain, we began by finding some case studies of e-assessment or online examination at German universities. In order to have an appropriate sample, we selected universities which conducted online examinations or had a project to recognize a unified approach or process regarding e-assessment in Germany. Different kinds of projects in terms of computer-based examinations have been in progress from the year 2000 onward. Table 3 summarizes all these projects with their functionality and their relation with exam regulation documents.

Of all the universities studied, only the University of Duisburg-Essen proposes a process model for implementing online examination. Proposing such a process model for online exams has the following advantages:

- improves the speed of feedback to students;
- motivates students and lecturers to increase computer skills through the integration of multimedia elements such as audio and video, and using complex digital systems;
- significant time saving for lecturers by automatic correcting process.

By reviewing the exam regulations and conditions, it becomes obvious that there is no identified exam process model in the administrative processes at different universities. A process model not only supports understanding the complexities of processes properly, but also contributes to advancing and improving defined processes (Irani et al. 2000). Therefore, in order to understand the examination workflows carried in the universities, a process model is required to analyze exam processes comprehensively.

3.1.3 Domain: Accreditation Process

Data sources: Accreditation agencies authorized by AISel⁵ and EBSCO

Research period: 2000–2013

Search terms and keywords: e-assessment, education, online examination, computer-based exam, accreditation process, accreditation criteria (in abstract and title)

As a definition, accreditation is a criteria-based procedure to assess and evaluate the admissibility of an educational program in terms of quality (Gorgone 2006; Impagliazzo and Gorgone 2002; Reichgelt and Yaverbaum 2007). The main goal of accreditation is to assess the educational quality of an academic program to ensure that it meets certain quality standards, called accreditation criteria (Reichgelt 2007).

Based on the European Network for Quality (ENQA), each educational program should fulfill the minimum in the following set of requirements to be accredited:

- requirements and objectives;
- teaching–learning process;
- learning resources;

⁵ Association for Information Systems eLibrary.

Table 3 E-assessment projects in German universities

University	Description of e-assessment projects	Relation to exam regulation documents
Free University of Berlin	<ul style="list-style-type: none"> • Since 2005 • Run by the center for digital systems (CeDiS) • Software: LPLUS Test Studio • 407 spaces for e-examinations • Not used for all study courses 	Not mentioned
University of Duisburg-Essen	<ul style="list-style-type: none"> • Since 2007 (in progress) • Software: LPLUS • Faculties: human science, social science, biology and geography, chemistry, engineering, and mathematics • Implementation of e-exams based on a process model 	e-exams are mentioned in some study courses as one of the exam methods beside other traditional forms Details of the e-exam process are not mentioned in these documents
University of Bremen	<ul style="list-style-type: none"> • Since 2004 • Software: LPLUS • 7500 e-exams per semester 	Since 2010, performing the electronic examinations as one of the examination forms is admissible in the examination rules and regulation documents
University of Giessen	<ul style="list-style-type: none"> • Since 2007 • Software: open-source learning platform (ILIAS) • Offers many opportunities such as learning modules for units, online glossary, import of SCORM and HTML tutorials, e-assessment module for online tests, survey module for user surveys, user management in courses and groups • This system is limited due to the lack of appropriate space to increase efficiency of e-examinations 	Not mentioned
University of Mainz	<ul style="list-style-type: none"> • Since 2007 • Software: open-source learning platform (ILIAS) 	<ul style="list-style-type: none"> • Exam regulation documents mention that e-exams should be carried out under the same conditions as traditional examinations • The details of the e-exam process are not mentioned in these documents

(continued)

Table 3 (continued)

University	Description of e-assessment projects	Relation to exam regulation documents
University of Koblenz-Landau	<ul style="list-style-type: none"> • Since 2013 • Software: open-source learning platform (ILIAS) • Two years for pilot and evaluation phase • Online exam submission through electronic exam schedule service (eKLAPS) 	Not mentioned because the project is new and in the pilot phase
University of Regensburg	<ul style="list-style-type: none"> • Since 2010 (in progress) • Road map defined which consists of different tasks and stages 	Legal conditions and e-exam framework during the whole process are stages of the e-assessment road map
University of Leipzig	<ul style="list-style-type: none"> • Since 2005 • Software: open-source platform (elateXam) • Around 1000 students per semester • Next step: develop the project as a pilot within the project iAssess.Sax in the universities of Dresden and Zwickau 	Not mentioned

- monitoring, analysis, and overview.

The assessment or examinations are placed in the teaching–learning process, which is the most complex aspect of this model because it includes a mixture of technical, pedagogical, and social competences and, furthermore, there is a great freedom to manage courses in order to achieve identified objectives.

According to Reichgelt and colleagues, there are two accreditation types:

- institutional accreditation is applied to an academic institution, like an entire university;
- specialized accreditation is applied to a subunit in an institution and includes two levels: program accreditation focuses on the content of the program, while administrative accreditation concentrates on the administrative process in a subunit of an institution (Reichgelt and Yaverbaum 2007).

The authors explain that there are two main approaches to accreditation processes. The first one is the input-based approach which measures various minimal standards through a checklist based on the learning-input processes such as curriculum, teaching resources, library, laboratory, and other facilities. The second is the outcomes-based approach, which considers the program’s outcomes, such as the institution’s educational objectives and student learning. Reichgelt and colleagues argue that a significant shift from the input-based to the outcomes-based approach

has occurred in recent years and academic institutions attempt to conform themselves with outcome criteria (Reichgelt and Yaverbaum 2007).

3.2 Accreditation Processes in Germany

In Germany, the federal states are responsible for accreditation processes and at least 11 authorized accreditation agencies in different fields of education (medical, natural science, engineering, economics, etc.) are in operation at present.⁶ Educational accreditation in Germany is based on two issues:

1. standards of study courses and degrees, which are based on regulations for study and examinations;
2. accreditation in teaching, which is based on quality improvement measures.

The requirements in the examination accreditation process are as follows:

- examinations are coordinated so that students have sufficient time to prepare themselves;
- it must be possible to move directly from the bachelor's degree to the master's degree without loss of time;
- the form of examination is laid down in the description of each module;
- examinations should not cause extensions to the period of study;
- the evaluation criteria are transparent for both lecturers and students;
- the degree program ends with a final thesis;
- it is checked whether students are capable of oral discussion in their specialist area;
- the supervision of the final thesis is subject to precise regulations in the curriculum.

The archived documents in the accreditation process are test results, drop-out rates, any quantity results of examinations, as well as feedbacks from the courses.

This survey of three domains indicates that multiple perspectives exist, which cause variability in performing and evaluating e-assessment processes. It is therefore essential to develop an appropriate model to account for this variability in online examination processes.

4 Literature and Results

This literature review was performed in order to demonstrate existing process variability in e-assessment in different domains of higher education, which occurs for different reasons. The results of each domain and its relations to process variability are analyzed and presented in the following.

⁶December 2013.

Table 4 Summary of e-assessment projects in German universities

Number of studied universities	30
Number of universities which have e-assessment projects (in progress)	18
Number of universities which perform online examination practically in some fields	9
Number of universities which have a paragraph about online examinations in exam regulation documents	3
Number of e-assessment projects in universities in Saxony	2

4.1 Evaluation of IT Approaches

Based on reviewed IT approaches, it can be argued that due to the great advantages of and positive impacts of online exams on learning processes, there is a considerable movement from traditional to electronic assessment in academia today. Furthermore, studies⁷ show that different issues, from pedagogical to technical and even social matters, cause process variability in higher education examinations. To identify a way of managing this inherent variability and to construct unified e-assessment procedures, it seems necessary to have a cross-functional view of e-assessment projects. Some of the issues related to e-assessment are: the impacts on student learning; effects on the teaching method; formative and collaborative e-assessment; immediate feedback and legal issues; evaluating possible risks in planning such as computer stress or fairness impression; developing infrastructural issues, such as security and question banks.

4.2 Evaluation of Study Courses

Thirty German universities from different federal states and different study courses were reviewed.⁸ The results show that e-assessment or online examination projects are currently in progress in 18 universities, but just nine are performing such assessments in practice. It should be noted that although online examinations are in use at some universities, there exists considerable process variability, too, which makes it difficult to extract a unified platform in this area.

Furthermore, legal conditions in regulation documents create practical limitations on performing online examinations. In order to make it an acceptable form of assessment, electronic examination should be referred to in a paragraph or even a sentence in the corresponding exam regulation document. Based on reviewing examination rules documents in all these universities, Table 4 reveals that only three German universities have a paragraph stating that online examination is an admissible examination form.

⁷All these studies are summarized in a table in Appendix 1.

⁸These universities are listed in Appendix 2.

4.3 Evaluation of Accreditation

The results revealed that accredited educational programs are subject to a variety of quality criteria. These criteria depend on the level and the goals of the courses. Based on the Information Model from the European Network for Quality (ENQA), assessment and examination processes are placed within the teaching–learning process, which is one of the complex parts of accreditation. It involves different aspects such as pedagogical, technical, legal, and even social issues.

No framework for examination forms (traditional or computer-based) was identified in this study of the German accreditation process. In exam regulation documents, assessments were only described in the terms listed at the end of Sect. 3.3 above. The form of examinations is not restricted by the accreditation process, but is under the authority of the educational systems and based on the identified objectives of courses. This causes process variability from one educational institution to another.

4.4 Summary of the Results

For clarity, we summarize the results of the three domains in Table 5.

In sum, exam processes show variability for the following reasons:

- variability in legal restrictions and regulations in different educational institutions;
- variability in exam framework, not only in different universities but also in one university between different courses;

Table 5 Summary of the results

Domain	E-assessment limits on PV ^a	Why PV management is necessary
IT approach	Different perspectives: – teaching–learning issues; – legal issues; – technical issues; – social issues	To have a cross-functional and integrated view
Course and assessment design	Legal conditions are not mentioned in exam regulation documents Different processes are used to perform online examination	To develop an appropriate approach to support legal conditions in exam documents
Accreditation	The form of the exam is not restricted by accreditation criteria, which leads to variability	To develop a business process model for clear e-assessment criteria

^aPV = process variability

- variability in study course design;
- variability in technical infrastructure for online examination processes.

Based on the obtained results, it can be argued that an appropriate design for variability management has to be aligned with the specified domain of projects and conditions to cover all variability within the identified domain. There is only one process model for performing online examinations in some study courses at University of Duisburg-Essen, which could be an appropriate model for conducting online examinations at German universities.

We studied existing process variability (effects and causes) in online examination processes from different perspectives. We found that to manage process variability, an appropriate business process model is necessary to cover all examination processes and features of educational institutions.

According to the existing literature, one approach to process variability management is the single model approach, which models all known variants of the process in one common model (Hallerbach et al. 2010; Kumar and Yao 2012). The alternative is to model every variant of a single model, which is called the multi-model approach. The latter models will have a simpler structure (Hallerbach et al. 2010; Kumar and Yao 2012). Some advanced modeling approaches explicitly deal with families of process models, such as configurable event-driven process chains (C-EPCs) (Rösemann and van der Aalst 2007), PROcess Variants by OPtions (PROVOP) (Hallerbach et al. 2010), ConDec (Petic and Aalst 2006), and feature modeling, which is normally used in software engineering. Each of these approaches has their own advantages and drawbacks but there is a lack of adaptability between different approaches (Valenca et al. 2013).

The next step in this research is to conduct a comprehensive overview of available variability approaches, to identify which of them are more appropriate for online examinations in higher education. Further research is needed to evaluate how existing process variability approaches can be compatible with exam processes, and to what extent existing process variability in this field can be controlled and managed.

5 Conclusion and Further Work

The aim of this research is to demonstrate existing variability in examination processes and emphasize the need for variability management. This could be the basis for developing a new approach to business process variability, or extending an existing one. To reach this goal, we concentrated on online examination processes in higher education.

As a preliminary stage, we performed a literature review in three different domains: IT approaches; concepts; course design and the accreditation process. This paper demonstrates the important role of business process management in improving and promoting the design of e-assessment processes.

Combined with this literature review, an analysis of current e-assessment case studies and projects in German universities yielded the following results. Although different kinds of projects under the name of e-assessment or online examination are in progress in German universities, the variability in these processes is recognizable. In other words, similar processes exist, but distinctions and variability are observable as well. Furthermore, the review of the exam regulation documents for various study courses at different universities revealed that these regulations do not yet mention an acceptable framework for performing e-assessments.

Process variability exists in e-assessment at German universities. It is necessary to manage this variability through an appropriate business process model to support online examination procedures.

This paper describes research in progress which clarifies and identifies the necessity of developing an existing approach to manage and control the process variability in university e-assessment. The next step is to study process variability management to identify an appropriate model in the context of e-assessment. This could be followed by the development of a prototype for the identified approach and finally the validation of the developed method.

Appendix 1: Summary of IT Approaches to e-Assessment

No.	Reference (author, year)	Summary
1	F. Braun, 1998	It introduces some web-based tools for online assessment such as Learning space, Web forms, Microtest, and Quizmaker.
2	Charoen, 2007	It indicates the benefits and limitations of implementing an e-learning system in an educational institution in Thailand based on expert interviews.
3	Dittman, 2008	It introduces a structure, the collaborative e-learning template (CET), to improve collaborative activities in the learning process such as collaborative assessments through a course.
4	Papp, 2000	It indicates different critical factors in distance learning.
5	Squires et al., 2004	It is about education policy analysis in Arizona University.
6	Shen et al., 2000	It analyzes the effectiveness of an online collaborative examination process through a questionnaire.
7	D. Hall, 2010	It introduces some tools to assess teachers.

(continued)

(continued)

No.	Reference (author, year)	Summary
8	Jacob et al., 2006	It reveals the importance of assessments in the learning process through different learning approaches. The Black Board Learning System (BBLs) is used as a comprehensive e-learning software platform in the context of continuous assessments.
9	Campbell, 2008	It proposes an application called the Electronic Performance Support System (EPSS) to enhance the performance of assessment tasks.
10	Kehily, 2011	It uses a web-based platform, DIT learning teaching and technology center, as a course management system for lecturers and as a computer-assisted learning tool for students. It is a case study in the Dublin Institute.
11	Attali and Burstein, 2006	It proposes a flexible modeling procedure which can be used to as a basis for expert judgment.
12	Venkatraman, 2007	It is a case study which develops a four-step student-centered approach to increase positive impacts on the learning process. It analyzes the strengths and weaknesses of assessments in IS courses through a survey.
13	Maheswari, 1998	It is an empirical study of technical problems in web-based systems for assisting education.
14	Sybol, 2005	It investigates how use of an online assessment tool, the Computer-Assisted Personalized Approach (CAPA), helps achieve teacher and students learning goals. It is a case study of urban courses in Florida University.
15	Webb, 2010	It evaluates how digital technologies such as tablet PCs, wireless technology, and Web 2.0, can facilitate formative and collaborative assessment effectively.
16	e-learning age magazine, 2011	It is a case study of The Association of Chartered Certified Accountants (ACCA), which wants to deliver all accounting examinations electronically through an e-assessment program.
17	Miller, 2011	It is a case study of the positive role of aesthetics design on learner perceptions and task performance in an e-assessment environment.

(continued)

(continued)

No.	Reference (author, year)	Summary
18	Dascalu and Bodea, 2010	It proposes a qualitative analysis of an e-assessment and its role to increase knowledge creation and knowledge management.
19	Sainburg and Benton, 2010	It is a case study of the importance of formative assessment in the learning process in schools.
20	Dermo, 2009	It evaluates the possible risks in planning e-assessment through six identified dimensions in a case study.
21	Jordan and Mitchell, 2009	It explains some methods to facilitate answer and assessment of free-text questions in e-assessment. It examines its model (IAT) in the UK Open University, using information extraction techniques and NLP.
22	Ferrao, 2010	It shows that an e-assessment system can be a good alternative to encourage students and has positive effects on the learning process.
23	British Journal of educational studies, 2009	It summarizes different works which all encourage educational institutes to use e-assessment technologies to support the learning process, with emphasis on feedback from e-assessments.
24	Hodgson and Y. C. Pang, 2012	It is a study of statistics students to evaluate the advantage of formative assessment.
25	Glenberg, 2010	It examines the impact of formative quizzes on knowledge construction.
26	Daly et al., 2010	It argues that existing assessment solutions focus on developing technical and infrastructure issues more than educational aspects of assessment through different educational areas.
27	McCann, 2010	It is a case study of a US campus to explore how a new e-assessment system is implemented and to identify factors which affect how it is adopted.
28	Nicol, 2007	It suggests a set of principles for the effective design and evaluation of a formative assessment and feedback process through two different case studies.
29	Boyle, 2010	It is about forecasting models like the Bass model, used to evaluate how educational systems can adopt themselves to e-assessment methods.

(continued)

(continued)

No.	Reference (author, year)	Summary
30	Boyle and Hutchison, 2009	It makes the case for the substantial effects of e-assessment on the assessing process and the technical difficulties involved in developing sophisticated assessment tasks.
31	McNeil et al., 2011	It looks at e-assessment from both sides: educational and technical. It develops a model for analyzing the life cycle of institutional assessments.
32	Sangi and Malik, 2007	It reviews existing e-assessment models and practices and identifies the challenges facing developers in South Asia.

Appendix 2: List of German universities reviewed

1. Freie Universität Berlin
2. Universität Bremen
3. Universität Duisburg-Essen
4. Justus-Liebig-Universität Gießen
5. Philipps-Universität Marburg
6. Johannes Gutenberg-Universität Mainz
7. Karlsruher Institut für Technologie
8. Rheinisch-Westfälische Technische Hochschule Aachen
9. Hochschule Koblenz
10. Leibniz Universität Hannover
11. Westfälische Wilhelms-Universität Münster
12. Universität Koblenz-Landau
13. Universität Regensburg
14. Universität Kassel
15. Technische Hochschule Ostwestfalen-Lippe
16. Technische Hochschule Wildau
17. Hochschule Fulda
18. Universität Trier
19. Universität zu Köln
20. Albert-Ludwigs-Universität Freiburg
21. Universität Düsseldorf
22. Hochschule Wismar
23. Universität Ulm
24. Eberhard Karls Universität Tübingen
25. Georg-August-Universität Göttingen
26. Rheinische Friedrich-Wilhelms-Universität Bonn

27. Otto-von-Guericke-Universität Magdeburg
28. Technische Universität Dresden
29. Universität Leipzig
30. Technische Universität Chemnitz

References

- Arnold, O., Laue, R.: Untersuchungen zur Beherrschbarkeit von Variabilität in Prüfungsprozessmodellen deutscher Hochschulen. In INFORMATIK 2014, 22–26. September 2014: GI-Edition—Lecture Notes in Informatics (LNI), P-232, S. 1135–1148. Bonner Köllen Verlag, Stuttgart (2014)
- Accountants find e-assessment adds up., e-Learning Age news, p. 4 (2011)
- Attali, Y., Burstein, J.: Automated essay scoring with e-rater® V. 2. J. Technol. Learn. Assess. **4**(3) (2006)
- Ayora Esteras, C.: Modeling and Managing Variability in Business Process Models (2012)
- Becker, J., Bergener, P., Kleist, S., Pfeiffer, D., Raeckers, M.: Business process model-based evaluation of ICT investments in public administrations. In: AMCIS 2008 Proceedings. Paper 353 (2008)
- Becker, J., Eggert, M., Bruning, D., Saat, J.: Applicability of business process model analysis approaches—a case study in financial services consulting. In: Wirtschaftsinformatik Proceedings 2013. Paper 74 (2013)
- Boyle, A.: Some forecasts of the diffusion of e-assessment using a model. Innov. J. **15**(1), 2–30 (2010)
- Boyle, A., Hutchison, D.: Sophisticated tasks in e-assessment: what are they and what are their benefits? Assess. Eval. High. Educ. **34**(3), 305–319 (2009)
- Braun, G., Crable, E.: Administering exams electronically: issues, techniques, and assessment. In: AMCIS 1998 Proceedings. Paper 353 (1998)
- Campbell, A.: Performance enhancement of the task assessment process through the application of an electronic performance support system (2008)
- Charoen, D.: Challenges and opportunities of elearning: a case study of higher education in Thailand. In: AMCIS 2009 Proceedings. Paper 132 (2009)
- Cooper, H.M.: Synthesizing Research: A Guide for Literature Review. Sage, Thousand Oaks (1998)
- Coyle, E.: Engineering Education in the US and the EU. Chapter 5. In: Engineering in Context (2009)
- Daly, C., Pachler, N., Mor, Y., Mellor, H.: Exploring formative e-assessment: using case stories and design patterns. Assess. Eval. High. Educ. **35**(5), 619–636 (2010)
- Dascalu, M., Bodea, C.: Challenges in building e-assessment services from project management knowledge perspective. Inter. J. Glob. Manag. Stud. Prof. **2**(1), 35–50 (2010)
- Davis, R.: Business Process Modeling with ARIS. Springer, London (2001)
- Davis, R., Brabänder, E.: ARIS Design Platform: Getting Started with BPM. Springer, London (2007)
- Dermo, J.: e-Assessment and the student learning experience: a survey of student perceptions of e-assessment. Br. J. Edu. Technol. **40**(2), 203–214 (2009)
- Dittman, D., Deokar, A.V.: Collaboration process patterns and integrated assessment in e-learning environments. In: MWAIS 2008 Proceedings. Paper 15 (2008)
- El-Ashy, A.K.: A methodology using UML & SADT to develop an effective business model in healthcare. Paper 449 (2006)
- Ferrão, M.: E-assessment within the Bologna paradigm: evidence from Portugal. Assess. Eval. High. Educ. **35**(7), 819–830 (2010)

- Gorgone, J.T.: Information systems accreditation: preparation, process, and standards. In: Communications of the Association for Information Systems, Vol. 17, Article 19 (2006)
- Gruhn, V., Laue, R.: Approaches for business process model complexity metrics. In: Technologies for Business Information Systems, pp. 13–24. Springer, Heidelberg (2007)
- Hall, J.D.: Fourth and Fifth Grade Student Achievement and Possible Relationship to Three Teacher Selection Tools (2010)
- Hallerbach, A., Bauer, T., Reichert, M.: Configuration and management of process variants. In: vom Brocke, J., Rosemann, M. (eds.) Handbook on Business Process Management 1, pp. 237–255. Springer, Berlin Heidelberg (2010)
- Hallerbach, A., Bauer, T., Reichert, M.: Issues in modeling process variants with provop. In: Business Process Management Workshops—BPM 2008 International Workshops, Milano, Italy, 1–4 September. Revised Papers, pp. 56–67. Springer, Berlin Heidelberg (2009)
- Hodgson, P., Pang, M.: Effective formative e-assessment of student learning: a study on a statistics course. *Assess. Eval. High. Educ.* **37**(2), 215–225 (2012)
- Impagliazzo, J., Gorgone, J.: Professional accreditation of information systems programs. In: Communications of the Association for Information Systems, Vol. 9, Article 3 (2002)
- Irani, Z., Hlupic, V., Giaglis, G.: Business process reengineering: a design perspective. *Inter. J. Flex. Manuf.* **12**(4), 247–252 (2000)
- Jacob, S.M., Issac, B., Sebastian, Y.: Impact on student learning from traditional continuous assessment and an e-assessment proposal. In: PACIS 2006 Proceedings. Paper 63 (2006)
- Johannsen, F., Leist, S.: Erratum to: Wand and Weber's decomposition model in the context of business process modeling. *Bus. Inf. Sys. Eng.* **4**(6): 375–375 (2012)
- Johnson-Glenberg, M.C.: Embedded formative e-assessment: who benefits, who falters. *Educ. Media Inter.* **47**(2), 153–171 (2010)
- Jordan, S., Mitchell, T.: e-Assessment for learning? The potential of short-answer free-text questions with tailored feedback. *Br. J. Edu. Technol.* **40**(2), 371–385 (2009)
- Kehily, D.: Review of Dublin Institute of Technology's E-Learning Platform for Effective Teaching Formative Assessment and Feedback. In: EDUCLEARN 2011, Barcelona, Spain, 4–6 July (2011)
- Kumar, A., Yao, W.: Design and management of flexible process variants using templates and rules. *Comput. Ind.* **63**(2), 112–130 (2012)
- McCann, A.L.: Factors affecting the adoption of an e-assessment system. *Assess. Eval High Educ* **35**(7), 799–818 (2010)
- Pesic, M., Aalst, W. van der: A declarative approach for flexible business processes management. In: BPM'06 Proceedings of the 2006 International conference on Business Process Management Workshops, pp. 169–180 (2006)
- Primiano, D.N., Pirjo-Liisa, O., Schade, A.: Accreditation models in higher education: experiences and perspectives. In: ENQA Workshop Report 3. Helsinki (2004)
- Reichgelt, H., Yaverbaum, G.: Accountability and accreditation: putting information systems accreditation into perspective. In: Communications of the Association for Information Systems, Vol. 20, Article 27 (2007)
- Rosemann, M., van der Aalst, W.M.P.: Configurable reference modelling language. *Inf. Sys. Databases Creation Manage Utilization* **32**(1), 1–23 (2007)
- Sangi, N. A., Malik, I.: E-assessment models and methods for student evaluation. In: 21st Asian Association of Open Universities Conference Putrajaya: Malaysia: Kuala Lumpur (Malaysia) University (2007)
- Valanca, G., et al.: A systematic mapping study on business process variability. *Inter. J. Comput Sci. Inf. Technol.* **5**(1), 1–21 (2013)
- Venkatraman, S.: Student-centred assessments and their impact on learning—an IS case study. In: PACIS 2007 Proceedings. Paper 50 (2007)
- Vogelaar, J., Verbeek, H.M.W., Luka, B., van der Aalst, W.M.P.: Comparing business processes to determine the feasibility of configurable models: a case study. In: Business Process Management Workshops, BPM 2011 International Workshops, Clermont Ferrand, 2011. 2012; S. 50–61. Springer, Berlin Heidelberg (2012)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Designing External Knowledge Communication in a Research Network

The Case of Sustainable Land Management



Thomas Köhler, Thomas Weith, Sabrina Herbst, and Nadin Gaasch

Abstract Designing knowledge communication with external partners is a core activity of research networks. In science, such communication has been addressed only recently and is still considered as non-academic activity. Successful communication with practitioners, that is knowledge transfer, is a crucial factor for effective research. In the age of online communication, this requires special attention and skills, for example related to social media communication. Based on our own empirical results derived from interviews, the authors identify what factors affect the communication process and how the design of communication content may be influenced.

To do so, successful examples of communication with external stakeholders are presented. For the theoretical basis, science communication, knowledge communication, knowledge management, and knowledge transfer were selected and consolidated. Although the findings stem from a research network specializing in sustainable land management, they can be transferred to other academic collaborations. Our results indicate that external communication is effective when knowledge has been transferred between academics and practitioners.

Keywords Research network · Knowledge management · Open science · Qualitative research · Land management

T. Köhler

Faculty of Education/Institute for Vocational Education, TU Dresden, Dresden, Germany
e-mail: thomas.koehler@tu-dresden.de

T. Weith

Leibniz Centre for Agricultural Landscape Research, Müncheberg, Germany
e-mail: thomas.weith@zalf.de

S. Herbst (✉)

Media Centre, TU Dresden, Dresden, Germany
e-mail: sabrina.herbst@tu-dresden.de

N. Gaasch

TU Berlin, Office of the First Vice President for Research, Appointment Strategy, Knowledge & Technology Transfer, Berlin, Germany
e-mail: nadin.gaasch@tu-berlin.de

1 Background: Theory and Project

The results presented in the article are developed in the context of funding measures of the German Ministry of Education and Research (Sustainable Land Management FKZ 033L004, Agricultural Systems of the Future - ZenKO FKZ 031B736, Urban–Rural Stadt-Land-plus ReGerecht FKZ 033L205) as well as a research and qualification project of TU Dresden. Selected parts of the article base on former work published as Zscheischler et al. (2012) as well as Härtel et al. (2015).

1.1 Sustainable Communication in the Sciences

Information and communication processes and the related content form an important basis for defining the principles of sustainable spatial development, at least since the 1992 United Nations Conference on Environment and Development in Rio de Janeiro. While forms of information provision and strategic use of communication are fundamental, renewed mediated information and communicative approaches have become widespread since the 1990s (Lievrouw et al. 2000). These can be applied fruitfully in different academic disciplines such as education or engineering, or in spatial planning and development processes (Weith et al. 2020). Respective information and communication technologies are now seen as part of different governance forms and are recommended for tackling various problems. For example, in 2003 the German Council for Sustainable Development initiated a “dialogue area” to strengthen understanding of processes of changing land use. Also the German federal government, whose goal was to reduce land use by settlement and infrastructure by 2020 to 30 hectares per day, began using such new communication instruments and triggered activity of further groups (even though the original timeline has meanwhile been extended to 2030¹). While in a first step tools such as education material, brochures, cartoons, and computer games were produced to sensitize the relevant actors (Bock et al. 2009), in the almost two and a half decades toward the Sustainable Development Goals (SDGs) the focus has changed from information to knowledge management (Weith and Köhler 2019). Specifically knowledge management is addressed in three of the SDGs (4, 16 and 17) and at the same time linked to education and lifelong learning. Digitization, although relevant for many goals, is explicitly addressed in sub-goal 9c (Industry, Innovation, and Infrastructure): “Significantly improve access to information and communication technology and ensure universal and affordable access to the Internet in the least developed countries by 2020” (United Nations 2015).

Non-governmental organizations (NGOs) such as the Nature Conservation Federation of Germany (Naturschutzbund Deutschland NABU), the Federation for the Environment and Nature Conservation Germany (Bund für Umwelt und Naturschutz

¹<https://www.umweltbundesamt.de/themen/boden-landwirtschaft/flaechen sparen-boeden-landshaf ten-erhalten#flachenverbrauch-in-deutschland-und-strategien-zum-flaechen sparen>.

Deutschland BUND), or the international World Wide Fund For Nature WWF now develop targeted campaigns to raise awareness and promote more sustainable use of natural resources. Initiatives such as the International Year of Biodiversity 2010, the International Year of Forests 2011, or the International Year Plant Health 2020 share mainly attempts to do this on an (inter)national level. Attention-grabbing activities must be taken rather frequently in efforts to change land use, because long-term changes do not have the direct “media marketing value” of disasters like the Fukushima tsunami in 2012. Against all communication efforts, discussion of topics such as soil conservation, land management, or the establishment of regional material cycles remains largely restricted to professional circles.

Despite comprehensive knowledge about communication theories and models, and especially about the concept of a sustainability communication, it can be stated that communication processes are not always effectively implemented. Today we may observe a pronounced awareness of sustainability in general and environmental issues in particular; in Germany, 64% of the population consider environmental and climate protection as an important challenge (BMU and UBA 2019) and the German Parliament may state in its 2019 Environmental Report that a “demanding environmental policy with effective environmental laws and competent environmental administrations is widely accepted by the population” (Deutscher Bundestag 2019, p. 4). Still there is a discrepancy between this awareness and individual behavior in Germany. For example, correlations of affect and cognition with environmental behavior are not particularly strong, but still substantial ($r_{aff} = 0.51$ and $r_{cog} = 0.48$). This means that people who agree with the affective and cognitive statements generally act more environmentally conscious (BMU 2019, p. 68).

In the view of the authors, this is due to the fact that the variety of existing means of communication are not used strategically and thus not exploited to their full potential (Kriese and Schulte 2009; Leipziger 2007). This is especially true in the “bulky” field of sustainability. The Sustainable Land Management (SLM) funding program of the German Federal Ministry of Education and Research (BMBF), described below, is used to critically investigate current practices in research and planning and to identify options for future activity. From 2008 to 2017, the BMBF initiated the SLM program to create a knowledge and decision-making basis for sustainable use of land resources. Already in designing the program the funder considered communication efforts as a central requirement for a successful implementation of this objective. This is only possible if all actors are willing to actually apply the new knowledge gained as a result of the program (Hinzen 2009). Targeted communication efforts played a central role in the management of inter- and transdisciplinary research networks of that funding scheme: It was a condition of information exchange, successful collaboration, and collaborative learning.

Successful communication not only creates awareness of new challenges, but also acceptance for new options, and may initiate behavioral change. It thus contributes significantly to the successful transfer and implementation of scientific findings into practice, in this case regarding the SLM program. But how can communication processes be designed in a targeted and successful way? How can existing knowledge of strategic communication sciences be linked to communicative requirements? How

can means of communication be used strategically in such a complex field? What specific challenges have to be considered? And where are the limits of professional communication?

This paper presents some initial answers to these questions which were developed by one of the scientific projects accompanying the BMBF-funded SLM network. To achieve better insight, the authors first present the core topics of the network and then explain the role of communication in this context. Subsequently, specific challenges and influencing factors are discussed in order to finally outline a strategic approach.

1.2 Theoretical and Conceptual Considerations for the Design of Communication Processes

Human communication is a constant, every day, yet highly complex process that predominantly occurs unconsciously. It is social behavior at a time determined by many factors which accompany the message intendedly sent by a person. These factors include emotions, situational circumstances, and the knowledge and cognition abilities of the communication partners involved and its variety makes the communication process complex. In designing effective communication, it is therefore essential to be aware of the most important factors for the sender and the receiver; the latter include attention, the everyday ecology, and the personal and situational capacity (Kuckartz and Schack 2002). Moreover, communication is expected producing a social exchange of constructions, orientations, ideas, etc., about the world, exclusively created in social discourse and checked for their suitability (Frindte and Geschke 2019, p. 107). By social interaction those individual communications form entities of organizational character (Köhler 2014), which lead to an inter-institutional, i.e., external communication and exchange of knowledge, for example in networks.

Designing knowledge communication with external partners is a core activity of all research, especially of research networks. In science, such communication has been addressed only recently and is still considered as non-academic activity. In the age of online communication, this requires further special attention and skills (Köhler et al. 2019). With Web 2.0, communication technology shifts to a new social form in which content is produced jointly, incorporating all those interested in a certain topic even if they do not have scientific backgrounds. In a society defined by mass media, rivalry for the attention of various target groups is intense, so attracting attention to an issue may need to be the first step. To be perceived is a basic condition for successful communication. But this attention has consequences: Those who create attention must also create content. The term “everyday ecology” takes the real life of the recipient into account. Information may only have an impact if it has a meaning for the recipient, that is, if it can be linked to their real life. Strategic communications utilize this relationship to their advantage by considering the consequences, benefits, and options for the recipient, and presenting them consciously. Basically, a subject should not overwhelm a recipient or a target group. If they do not have the capacity to process

a topic intellectually or emotionally, they may reject or avoid the information. All of these factors need to be analyzed and adapted to specific audiences. A communication strategy must therefore be adapted to the needs of the intended recipient. Senders have more options, which are determined by variables like authenticity, professionalism, and the available financial resources (*ibid.*).

The authenticity of a source is relevant to its visibility (Köhler 2016). Communication functions less on the level of the actual content than in terms of the type and way it is communicated. Credibility, competence, and empathy are the central determinants here. Increasingly, the communicator needs to be professional in order to compete for the “scarce resource” that is the attention of each target group. This professionalism includes organizational and technical know-how, knowledge about methodology, that is, how to address specific target groups, and practical experience. This can be achieved with further training and the help of external communications consultants or agencies.

Experience has shown that too often the only aspect of communication to be considered was the means, and that this was hardly ever strategically communicated (Kriese and Schulte 2009; Leipziger 2007). Recent findings focus on the need for human and financial resources as key to planning, designing, and implementing communication of innovations successfully (Pscheida et al. 2013). This means scheduling appropriate resources and setting goals for communication activities right from the beginning, at the initial planning stage of a research project. Yet, before the actual communicative tasks and related objectives are formulated, the means are often already fixed, usually without any consideration of whether they meet the purpose. Researchers and communicators need to consider the following questions: Is the chosen means useful in view of the objective? Which channel should be designed to address the target audience? What must be communicated and what must not? What steps need to be taken and in what order to achieve the goal?

1.3 Knowledge Management in the Sustainable Land Management Program as a Challenge for External Communication

The Sustainable Land Management program (SLM) had to meet a number of specific challenges toward developing an integrative communication approach. First, the organizational structure of the program was very complex (cf. Fig. 1). In more than two dozen collaborative projects and its 120 subprojects, scientists and practitioners from over 170 organizations were involved. The scientific disciplines involved in SLM brought very different perspectives, methods, and understandings to the overall SLM program. Unsurprisingly, science and practice often have different preferences, and thus communicative goals could be very heterogeneous. It is therefore obligatory to develop a comprehensive communication strategy that is accepted by the parties.

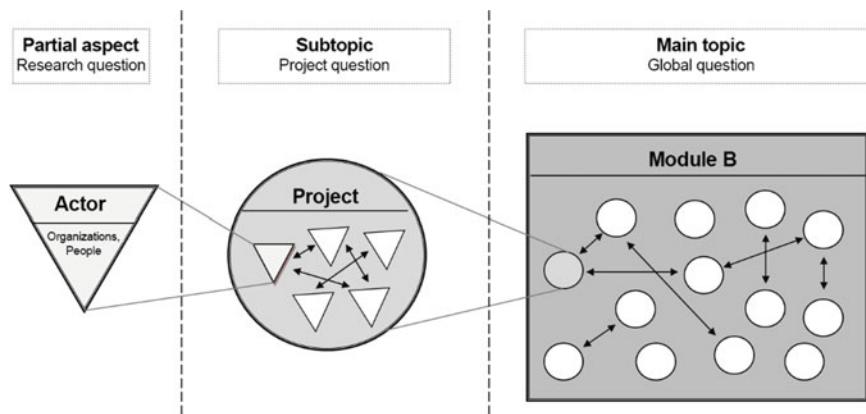


Fig. 1 Schematic representation of the network representing the funding program “Sustainable Land Management” (cf. Härtel et al. 2015; translated by authors)

Another challenge was to establish communication structures at the beginning of the program. In our case, a new organizational context with new communication channels needed to be defined and then perpetuated in the new SLM research network. This was time consuming and required resources, as individual experience from completed projects could not have been reused one to one. But this is a general challenge in science as nowadays, research is often project-based and short-term, that is, the organizational structures are frequently terminated and re-established again. Although bilateral or multilateral research and practice networks remain, the topic-overarching management structure, which includes integrated communication, usually dissolves. This is one reason why it is difficult to implement and perpetuate the results, knowledge, and experiences obtained. In addition, when research starts, the results are not yet available and cannot be presented quickly. But media products to be communicated must be developed first, i.e., cannot be finalized only when advertising needs to begin. At this stage, researchers are still developing models, principles, strategies, and combinations of instruments, which are complex and may be highly abstract. Accordingly, the “new” knowledge is owned only by a small group of experts only but has not been transferred to the target audience yet.

Further, SLM is an overarching term, so its actions are not clearly defined. Communications had to clarify what is meant by all three ambiguous and much-debated parts of the term, “sustainability,” “land,” and “management.” This means that all participants of the program had to deal with a high complexity and enormous variety of subjects. In fact SLM combined many issues which embody enormous communicative challenges. The collaborative projects, for example, were dedicated to sustainable water management, regional material and energy cycles, renewable resources, ecosystem services, sustainable urban development, and urban–rural linkages. The actors, interests, and target groups were also numerous. Thus, from a scientific point of view SLM is a highly complex field, which is typical for many research networks, especially for those that link research with its application in practice.

Yet, mass media requires a high level of simplification, which is contrary to the claims of many scientists. They often have difficulty handling active (non-technical) media and do not want their highly complex topics to be reduced to simple, striking stories. This resonates with a common concern about losing their reputation in the scientific community. Scientists feel that they eventually lose control and sovereignty of interpreting their results through publication in the mass media. Black-and-white arguments such as “renewable energy is positive” or “nuclear energy is bad” contradict not only the scientific, but also the communicative self-understanding of science. Due to that fear, scientists begin stepping into public in order to engage for their research, i.e., start acting as lobbyists. Respectively, they become aware to generate findings which may be associated with social consequences—what calls for a renewed consideration of research ethics (Dobrick et al. 2017) and can only be achieved with assessment based on normative values. Roose (2006) therefore speaks of an increasing politicization of science while Weingart (2001) points out the danger of its political exploitation.

Altogether, an intelligent communication strategy is required that centers on targeted but achievable action, even though very limited financial resources are available. Indeed communication for a typical research network like SLM cannot follow the rules of classic advertising because of the special funding conditions for such non-commercial topics. Yet, knowledge of the discussed key determinants of communication is essential. Therefore, in the following, the methodology for evaluating the most significant factors empirically is briefly introduced.

2 Approach and Methodology

2.1 Data Collection

Social science research objects and their stakeholders, as in the present case of the SLM funding program, are characterized through a complex and procedural context (Witzel 1985, p. 227). Following the research question, it was necessary to identify exemplary information transfer and implementation strategies within SLM funding program represented by the collaborative projects. The problem-centered interview was selected as the survey method to investigate the communication structures of the individual projects (Kaiser et al. 2012). We were interested in both internal and external transfer and implementation strategies. Twelve typical stakeholders concerned were interviewed, in order to collect their experiences and establish a systematic knowledge base. Following a qualitative approach, there was an equivalent consideration of both researchers and practitioners, covering all types of projects. In the course of the investigation, a research guide was developed as a basis for discussion, including aspects related to both content and communication. All researchers of the collaborative projects contributed to the guide, which covered all subject areas of interest regarding transfer and implementation. For the

present analysis, the authors focused on the concept of transfer, and especially on communication-related aspects.

2.2 *Evaluation Method*

To analyze the interviews, the authors applied the qualitative method of content analysis developed by Mayring (2010), which can be used for communication text data. To cope with the length of the text and to serve the purpose of the problem-centered interviews, the authors decided to complete a structured content analysis. Guided by theory-based main categories, we systematically worked through the transcribed material and passages assigned to the categories.

By focusing on the transfer and implementation of knowledge in the project network and the resulting questions, main forms of practice can be concluded along the theory in a deductive way, forming main and sub-categories (see Härtel and Hoffmann 2013). In order to address the criterion of openness of the research process, the authors created an inductive category using the summary content analysis. Overall, the focus was on structured content analysis, in particular on structuring the content: “to filter out certain topics, content, aspects of the material and to summarize it” (Mayring 2010, p. 98).

The evaluation was conducted using the software MAXQDA. Specifically developed for structured content analysis, the software allows for the definition and use of codes and sub-codes which reflect the categories in an orderly manner (Kuckartz 2010, p. 114). Methodologically controlled compression of the material was used to work out cross-case statements on regularities in the terms of the research question (*ibid.*, p. 110ff.).

3 Results

3.1 *Practitioners and Civil Society as Target Groups of External Knowledge Communication*

In the course of the interview analysis, we identified the target groups of external knowledge communication: These were practitioners in the economy, society, politics, and administration. Addressed economists often represented the agricultural and forestry sector; the former included the food manufacturing, and the latter the wood processing industry. Other practitioners were at the interface between the public and private sectors such as health care (doctors, health insurance), mobility and transport (transport networks), the energy sector, and the private education and research sector. Administrative practitioners covered a variety of responsibilities (municipal and state level) and subject matter including conservation, transport infrastructure,

and health. Civil society actors, in a broader sense, included voluntary clubs and NGOs, such as nature conservation associations, support groups, and interest groups such as farmers' organizations.

3.2 Effects and Interactions of Factors Influencing External Knowledge Communication

Various aspects of inter-institutional, i.e., external knowledge communication (which we find in networks as well, going beyond bilateral exchanges) can be derived from the interviews, in terms of both content and process. First, the means of communication are selected with different intentions. Content includes project content and results usually developed for two reasons: to transfer knowledge from science to practice, and to provide feedback from actors during the research process. At the process level, it was observed that region and theme often influenced whether communicated content was picked up. Further, different practitioners often have different expectations on knowledge transfer. In the interview analysis, authors identified the following dimensions: efficiency in developing solutions and economy of the provided solutions (project results, scientific knowledge); practicability of the developed solutions to concrete problems or at least not making such problems worse (positive and negative movement between actors and researched problems); and the meaning of personal attitudes (in form of expectations regarding the research topic). These dimensions influence perception, acceptance of, and willingness for further communication in networks. The success of external knowledge communication is also affected by legal and statutory conditions, such as funding and copyright, the available human and financial resources, and scientists' capacity for such communication.

3.3 Selecting a Suitable Means of Communication

The available financial and human resources often limited the choices regarding means of communication in the collaborative projects. For example, limited resources hindered knowledge transfer between science and practice: "And then it was evaluated how expensive it is (...) and then it was determined that well that would surely exceed the budget" (Interview 1.1). Legal and statutory conditions had a similar effect: Privacy policies impeded access to the target group and restricted the means of communication. External communication needed to be adapted to the concerns of target groups, seasonal or other variations. "[W]e have always started public relations work in May, June, and not before, because [...] this is a seasonal theme, and you can't kindle a fire which keeps [burning] year round" (Interview 1.3). The general attitude of stakeholders, key players, and the audience to the project problem and results (e.g., environmentally sustainable agriculture and renewable energies) could

impede or even prevent access to the target group, regardless of the means of communication used: “Of what avail is it, if the owner [of an agricultural land] tells you at the end (...): ‘No, I don’t like it because I have something against renewables anyway.’ (...) There are some, very flat opinions” (Interview 2.9).

Obstacles related to the character of the actors have been found to be surmountable using appropriate means of communication. “And it takes a long time for these introverted groups. You can’t hope at this moment. I have just spoken with one of them: ‘Yes, yes, mhm, yes it’s good’. And he didn’t even say goodbye. And then you sit on the phone and think, ‘What just happened?’” (Interview 2.5). It was in particular difficult to access practitioners who demonstrated a lack of trust. Reasons may include negative experiences in the past: “[T]his is certainly the downside we have in East [Germany], that people had pretty big security needs from the political system of the GDR [...] And today it is different and therefore people are sometimes overwhelmed and in certain places have been, I’d say, fooled, and that’s why they tend to be careful” (Interview 2.9).

To reach as many people as possible and to give actors an insight into the status of the project, open-access publications were recommended: “You have to say that very clearly, this is a public [research] project, (...) so that we also see our obligation to make all our results publicly available. (...) [W]e want [the results] to actually be disseminated and accepted and we will then make the best possible information available to the public” (Interview 1.6). Direct face-to-face contact with the target audience could help too to identify representatives who could spread the message: “We went to the event and just talked with people. And at the agricultural fair you get very direct contact with the people. And, in fact, an assumption that I had proved to be correct. Namely, that a well-defined type of farmer [...] is the first who we can connect with” (Interview 2.5).

3.4 Selecting and Preparing the Communications Content

The content to be communicated has to meet the expectations of the target audience. In the course of the interviews, it came out that certain scientific project content, despite its practical relevance, was too complex and abstract for industry partners to see its relevance. Indeed, the wish was expressed that “the topic is somehow prepared either for the target group or scientifically [...] But, do not tell the whole world. Such a claim can only go wrong” (Interview 2.8). For scientists, this means “that science must speak increasingly in the language of the local partner when initiating contacts. So, not the language of science. [...] they need to translate for ‘the average Joe’” (Interview 2.9).

Content prepared without a target group in mind, such as an exclusive scientific publication of project results, could “not reach all who work in practice” (Interview 1.6). For many actors, cost-effectiveness and efficiency are key criteria for the measure which is communicated. This may be a precondition for any dialogue with

actors outside the scientific community. If project content or results are not considered economically or efficiently, it will be very difficult to communicate them to stakeholders. Certain topics in the field of SLM are perceived as “unattractive” or cannot be communicated easily to the wider public per se, such as the issue of short rotation coppice or biodiversity. This is often due to general attitudes of actors in the field of environmental sustainability.

If the impact of the project on a target group is expected to be negative, it is necessary to reflect on when and what content can be communicated: “The word ‘re-watering’, that’s what you say after a half-hour conversation. When people know that they will not be inundated. [...] You cannot come in with that” (Interview 2.5). Nevertheless, direct involvement can have positive effects, especially in the case of problems that would otherwise enjoy little attention or which create little incentive to generate scientific knowledge. This is true of non-tradable areas in the health sector: “My concern is conducive for health projects [...] if they are not able to be commercialized. I’m not talking about pharmaceutical development, where big profits attract attention but where it actually comes to service” (Interview 1.4). This partly precluded the need to prepare the project content, resulting in rather low commitment from scientists. “We collect the messages, see what is framed by this and then make a nice communication profile. What was the result? After two reminders came nothing at all. Only after a third, relatively nasty email [...], then came the usual suspects [...] the ones that had always made them anyways. And then we have with very full, very, very, very petty, very painstaking work somehow collected the messages. But now looking back, they were not really new messages. There were the usual messages” (Interview 2.8). When project outcomes have been prepared properly, focused on the target group by integrating a science journalist, the message reached people who can pass it on, such as journalists (Interview 1.5).

The content of external knowledge communication must be targeted to its audience to ensure successful knowledge transfer. In order for recommendations to be adopted in practice, it is important that “you very strongly address [...] the participants and pick them up thematically where they are anchored, that is, when I talk to a farmer who might not necessarily be interested in the depth of the bird world, but who cares more about the agricultural effects” (Interview 1.6).

Legal frameworks, such as copyright and intellectual property, can hinder knowledge transfer between science and practice, as certain technologies cannot be readily used by the practitioner: “There is a little problem: This is patented. One cannot simply be reconstructed, there are costs. But [...] it works, if constructed properly” (Interview 2.4).

Another hurdle for knowledge transfer was the profitability and efficiency expectations of practitioners. The cost of project results is even described as “the most inhibitory factor” (Interview 2.7) for successful knowledge transfer. This applies not only to economic practitioners but also local governmental ones, such as mayors. The latter could be encouraged to support and potentially pass the message on if they could identify potential for regional development: “because a small community in rural areas has to simply see what options are there to generate added value” (Interview 2.9).

3.5 Addressing the Attitude of Stakeholders

Market conditions and industry policy often influence the attitude of stakeholders to problems, such as climate change adaptation in the food sector. It required special treatment of the content—differences and similarities must have been disclosed and potentials illustrated (see Interview 1.1). Some actors doubt whether the problem exists: “Is the climate really changing? So, are they telling us the right thing? And if they are not telling us the right thing, can I still do what I’ve been doing so far? Yes, they are easy and economic decisions … decisions of habit: I have done it this way before, so why should I do it differently now? And you cannot answer that, you can just say, ‘From OUR perspective if you do it, this and this will happen, and that and that will not happen, and you will have this and that risk’” (Interview 1.5). Another important factor in acceptance of research results and policy implications is vividness: “once it has something presentable. It is precisely what you can show this clientele—the farmers—something that is useful” (Interview 2.5). This may be “something photogenic” (Interview 2.5), but may also include specific transfer measures such as “field days.” Events must specifically be relevant to the practitioners who attend, “but this being-on-site feeling and talking about it is what makes something. On that level I can facilitate the transfer easily when I create environments that are unusual” (Interview 2.5). The extent to which communication activities can be institutionalized successfully often depended on the available financial and human resources (Interviews 1.5 and 2.2).

The degree of concerns of target audience about the problem addressed in project has partly affected the communication with actors in a positive and negative way. Existing networks could simplify the access to the actors: “facing the transport sector was [...] in many respects beneficial that they knew each other and that communication then took place without problems” and “it was always a very good basis to get in contact with these target groups” (Interview 1.2). The means of communication thus became less relevant: “if now by email or [...] the better the connection is, the less important the communication means becomes and the more likely the success of communication” (Interview 1.2). Actors in a network make networking work: “if we have three, four who really want it in the county or district, that’s enough. We don’t need much more” (Interview 2.5). If it was in the interest of certain groups of actors to process a problem, this could encourage their commitment to successful knowledge transfer: “the local players participated because they also partly had a personal interest” (Interview 2.3). At the same time, actors did not give their support when a scientific problem has been considered irrelevant in practice: “without personal involvement, you may encounter limited interest. Because many other things are more important in the view of the people” (Interview 1.4). “Due to the positive development for farmers on the world market and also here in Germany for agricultural products, they are not dependent on this new product [...]. Everything is going well for him in the field. Why should he tackle these uncertainties?” (Interview 2.4).

In general, different communication tools and channels, and how the content is revised, influence the potential for interacting with external knowledge communication. The analysis shows, the more specific and relevant to practice the content was, the more likely it was to get feedback from the addressed actors, because “people ask only if they know that they can ask. [...] You can only communicate that. Or I have a basket full of messages and can precisely position the target and target group” (Interview 2.8). For successful knowledge transfer to a wider public, acceptance and civil society participation was important: “you just involve people that are really objectively confronted with the background information. And people can simply decide what they, as it were, want. And then people can decide” (Interview 2.1). The potential for interaction was increased by extending the interfaces between the project researchers and the target audience, for example by including opinion leaders and key players on external project advisory boards.

Online media could increase the potential for interaction and facilitate target group feedback: “On our website we also get questions, requests for information. And so the messages that are received they have recorded. And [...] we hear, for example, that rental charges are a problem at the moment. And we take a look: so, does this really have an economic impact on this calculation, cost calculation, or is this actually a side issue, perhaps with a psychological effect, but has no economic meaning? And we grab the topics and try to then integrate them within our considerations and in our presentations” (Interview 2.4). Mere marketing efforts in external communication were not conducive, but rather direct exchange and regular contact with practitioners and key players: “We just don’t do marketing. Instead we explain to them, we tell the tale. From our experience, from the first research results come conversations with practitioners. And then they share something” (Interview 2.5).

4 Conclusions

In order to cope with the variety of factors in communication processes and to keep a clear head, theorists from marketing and communication studies recommend a systematic approach with a specified sequence of operations, especially with a clear concept of how social media communication is embedded (cf. Leipziger 2007; Hansen and Schmidt 2010; Kreutzer 2018). Respective steps usually include an analysis of the current situation, the definition of the communicative tasks, the development of communication goals, the identification of target groups, the development of messages, and the designation of a strategy to implement the selected communication approach. To some extent such an approach matches the generic idea about implementing communication technologies into an organizational configuration in an ideal manner as developed by Munkvold (2003), who divides into four sub-areas of implementing collaboration technologies:

- Organizational context;
- Implementation project;

Technological context;
Implementation phase.

Authors did previously explain how such can be transferred easily to the context of information exchange, learning and education (Köhler et al. 2010).

4.1 *Background and Communicative Tasks*

Sustainability communication is about to become a topic in scholarly publications at the intersection to citizen science practices (Weith and Köhler 2019), specifically, the influence of digitization on the genesis of knowledge in the context of a sustainable, fair development as discussed. In the context of land management, there is an overall strong focus on specific branches like tourism (Tiago et al. 2020). In any case that process begins with the collection of information describing the initial communication problem. As it is important to identify the significant data, this includes the consideration of relevant target groups and goals from the very beginning. Initially, only those facts relevant to communication problems are included. This process will identify communicative tasks that derive from the description of the initial situation and the necessity to modify or enhance communication actions. Its interpretation explores the problems to be solved, but does not yet offer solutions. In the specific case addressed here, the task is to identify how land users can most effectively acquire knowledge and develop a decision-making basis for SLM.

4.2 *Definition of Communication Objectives*

After analyzing the current situation and defining the communicative tasks, communication objectives should be formulated in a communication strategy. Objectives describe the desired end state of a process. Those are measurable and thus represent a kind of commitment. However, goals can change over the course of a project and then need to be adjusted.

Goal setters must ask whether the objectives can be reached at all. As well certain practices like the sustainability digital communication relationships are especially effective (Tiago et al. 2020). Of course unrealistic goals may not serve as appropriate, reliable basis for a successful communication concept. For example, including expensive measures within a modest budget will jeopardize the objectives.

4.3 Definition of Target Groups

The more precisely a target group is defined, the better it can be addressed. This definition determines the communication tools and approach to be used. When identifying target groups, it can help to be guided by demographic, lifestyle-related, or functional factors (Hansen and Schmidt 2010). In some cases, target audiences may be divided into subgroups with different patterns of media reception (Fischer 2012). This information is used to decide how to access the target group. It should be noted that people play different roles, at work or at home, with family or friends.

A survey of collaborative projects in a workshop of the SLM network revealed that due to the wide range of actors relevant to the topic of land use, a variety of target groups exist. Some target groups could be clearly identified, such as stakeholders from management, agriculture, associations, academia, and research. Other target groups were described only very generally and imprecisely, such as “local people.” This is problematic, in terms of selecting both the communication tools and content to be communicated.

4.4 Formulating Messages

Messages are content to be communicated to representatives of a target audience. The larger the target group selected, the simpler the messages should be. Such simple messages should consist only of models for everyday use. Complicated theoretical concepts have no place in the mass media, which presents a major challenge for scientists. It should be clear which effect model shall be implemented and why—if there is a marketing communication or an educational communication addressed. In a complex topic such as SLM with very different target groups and representatives on different levels of influence, it is also advisable to limit the selection of subjects and to focus on key contents. Such contents should contain only the most relevant information and consequences for the selected audience and the respective recipient. If the aim is behavioral change, the messages should present options for action.

4.5 Definition of Communication Strategies

The next step is to determine how to achieve the designated goals with the resources available. This means looking for the cheapest “lever” with which the target can be achieved most efficiently and effectively. The strategy combines all the resources for a specific parent maxim (Leipziger 2007). More recently, it is suggested that when analyzing sustainability communication a typology of three different communication modes would be appropriate: communication of, about, and for sustainability (Fischer

et al. 2016). Obviously the SLM network applied all three dimensions simultaneously as these have been components of the developmental approach.

Well-known and frequently used approaches include the piggyback, testimonial, and provocation strategies (see Hansen and Schmidt 2010; Leipziger 2007). The aim of the piggyback strategy (also known as issue management) is to attach the desired message to a consistent public relations action or to current media issues. The testimonial strategy generates attention through celebrity ambassadors, and the provocation strategy seeks to attract attention by breaking taboos or challenging competitors. So far, our network has favored a piggyback strategy.

4.6 Activity Planning and Scheduling (“Concerted Activity”)

Only when the strategy has been defined does the implementation begin. Now is the time to clarify what measures will be taken when, where, and how often? An appropriate mix of measures is in line with the strategy and aims to attain the objectives set. Not every interesting idea is to become an appropriate measure. The results are an action and a schedule that represent all the measures in chronological order, the so-called communication plan. This becomes a management tool, provides an overview of all parties, and allows accurate budgeting.

However well thought out a plan may be, its implementation depends on many events which are not clearly predictable. There is the momentum of cross media communication as well as specific preferences of single stakeholders, especially in science, who are eventually not skilled for supporting social media communications (Pscheida et al. 2015; Albrecht et al. 2020). Resulting, delays can occur or journalists may suddenly no longer be interested in the subject because breaking stories take priority. Subsequently, deviations between plan and reality arise and one needs to respond promptly and derive new action consequences.

4.7 Limitations of the Study

The empirical data used for the study is limited to just one research network, in which stakeholders often share a single focus, embedded into a single domain. Even though the configuration of the network is overarching sectors and includes research as well as public administration as well as stakeholders from industry, its outreach is limited. Mainly representativity is hindered by the missing link to the individual citizen as well as the missing direct link to the media sector.

Additionally, the role of project advisory boards has not been addressed in detail, that is, political influence may overlap with other effects described. In this case, the direct involvement of the target group also increased the potential for interaction in the communication process. It would also have been possible to assess external knowledge communication needs by questioning the target group directly. Finally,

authors did not explicitly address the presumably high potential for interaction offered by the concept of citizen science.

4.8 *Lessons Learned*

The study demonstrated that sustainable land management is a case which does have specific communicative affordances due to its complex, multi-actor character that brings together different perspectives. Still communication in and for research networks is not consequently addressed, literature both on practice and research is rather limited. Not only with the increasing meaningfulness of digital formats there is an increasing need for a well thought, analytically proven approach in designing communication of and within research networks strategically. Obviously scientists are not easily capable of developing such and are hindered first by their individual characteristics (limited competences skills, etc.) but as well by the ecological conditions (economic and structural deficits). In that sense the paper has collected theoretical and empirical evidence of how research may deal with the expectations toward the influence of (mainly digital) communications on the genesis of knowledge in the context of sustainable, fair development.

References

- Albrecht, S., Minet, C., Herbst, S., Pscheidla, D., Köhler, T.: The use of digital tools in scholarly activities. Empirical findings on the state of digitization of science in Germany, with special focus on Saxony. In: Koschcial, C., Köhler, T., Felden, C. (eds.) e-Science—The Enhanced Science. Progress in IS Series. Springer, Berlin (2020)
- BMBF (Bundesministerium für Bildung und Forschung): Bekanntmachung der Richtlinien über die Fördermaßnahme Nachhaltiges Landmanagement, 24. Oktober 2008. <https://www.bmbf.de/foerderungen/13138.php> (2008). Accessed 5 Apr 2020
- BMU (Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit) & UBA (Umweltbundesamt) (ed.): Umweltbewusstsein in Deutschland 2018. Ergebnisse einer repräsentativen Bevölkerungsumfrage. https://www.umweltbundesamt.de/sites/default/files/medien/1410/publikationen/ubs2018_-_m_3.3_basisdatenbroschuere_barrierefrei-02_cps_bf.pdf (2019). Accessed 23 Jan 2020
- Bock, S., Hinzen, A., Libbe, J. (eds.): Nachhaltiges Flächenmanagement—in der Praxis erfolgreich kommunizieren. Ansätze und Beispiele aus dem Förderschwerpunkt REFINA, Berlin (2009)
- Deutscher Bundestag: Umwelt und Natur als Fundament des sozialen Zusammenhaltes. Umweltbericht 2019; Unterrichtung durch die Bundesregierung, Drucksache 19/13400. <https://dip21.bundestag.de/dip21/btd/19/134/1913400.pdf> (2019). Accessed 5 Apr 2020
- Dobrick, F.M., Fischer, J., Hagen, L.M.: Research Ethics in the Digital Age: Ethics for the Social Sciences and Humanities in Times of Mediatization and Digitization. Springer, Berlin (2017). https://link.springer.com/chapter/10.1007/978-3-658-12909-5_6
- Fischer, D., Luedcke, G., Godemann, J., Michelsen, G., Newig, J., Rieckmann, M., Schulz, D.: Sustainability Communication. (2016). https://doi.org/10.1007/978-94-017-7242-6_12

- Fischer, H.: *Know your types—Konstruktion eines Bezugs zur Analyse der Adoption von E-Learning-Innovation in der Hochschullehre*; Doctoral Dissertation, TU Dresden/Uni Bergen (2012)
- Frindte, W., Geschke, D.: *Lehrbuch Kommunikationspsychologie*. Beltz, Weinheim (2019)
- Härtel, L., Hoffmann, M.: Transfer- und Implementierungsstrategien von Wissen im Projektverbund. Empirische Untersuchung am Beispiel der BMBF Fördermaßnahme Nachhaltiges Landmanagement. Forschungsbericht in Form einer Masterarbeit, TU Dresden: not published (2013)
- Härtel, L., Hoffmann, M., Köhler T., Weith, T.: Wissenskommunikation und Transfer für die Landschaftsentwicklung. Eine Analyse im Forschungsnetzwerk „Nachhaltiges Landmanagement“. *Zeitschrift für Gruppendynamik und Organisationsberatung* **46**, 289–312 (2015). <https://doi.org/10.1007/s11612-015-0296-0>
- Hansen, R., Schmidt, S.: Konzeptionspraxis. Eine Einführung für PR- und Kommunikationsfachleute: mit einleuchtenden Betrachtungen über den Gartenzwerg. 5. Aufl. Frankfurter Allgemeine Buch, Frankfurt am Main (2010)
- Hinzen, A.: Kommunikation und Bewusstseinsbildung. In: Bock, S., Hinzen, A., Libbe, J. (eds.) *Nachhaltiges Flächenmanagement – in der Praxis erfolgreich kommunizieren. Ansätze und Beispiele aus dem Förderschwerpunkt REFINA*. Berlin (2009)
- Kaiser, D.B., Köhler, T., Weith, T.: Informations- und Wissensmanagement im Nachhaltigen Landmanagement (IWM im NLM). In: Köhler, T., Kahnwald, N. (eds.) *Virtual Enterprises, Research Communities & Social Media Networks. Proceedings of the GeNeMe 2012*, pp. 121–133, Dresden (2012)
- Köhler, T.: Patterns of inter-institutional and inter-organizational collaboration. Strengthening the relationship between VET and the labor market for developing a professional work force. In: Muslikhin, M., Surono, I.M. (eds.) *Empowering Vocational Education and Training to Elevate National Economic Growth. Proceedings of the 3rd International Conference on Vocational Education and Training ICVET 2014*, UNY Press, Yogyakarta (2014)
- Köhler, T.: Visual anonymity in online communication. Consequences for creativity. In: Skulimowski, A.M.J. Kacprzyk, J. (eds.) *Knowledge, Information and Creativity Support Systems: Recent Trends, Advances and Solutions*. Springer, New York (2016)
- Köhler, T., Neumann, J., Saupe, V.: Organisation des Online-Lernens. In: Issing, L.J., Klimsa, P. (eds.): *Online-Lernen. Handbuch für Wirtschaft und Praxis*. 2. Korrigierte Auflage. Oldenbourg Wissenschaftsverlag, München (2010)
- Köhler, T., Schoop, E., Kahnwald, N.: Communities in New Media. Researching the Digital Transformation in Science, Business, Education & Public Administration. Proceedings of 22nd Conference GeNeMe 2019. TUD Press, Dresden (2019)
- Kreutzer, R.T.: *Social-Media Marketing kompakt*. Springer Gabler, Wiesbaden (2018)
- Kriese, U., Schulte, P.: Flächenakteure zum Umsteuern bewegen! Analyse und Bewertung vorliegender Kommunikationsansätze – Ausgangspunkt für neue kreative Marketingstrategien. In: Bock, S., Hinzen, A., Libbe, J. (eds.) *Nachhaltiges Flächenmanagement – in der Praxis erfolgreich kommunizieren. Ansätze und Beispiele aus dem Förderschwerpunkt REFINA*, pp. 47–56. Berlin (2009)
- Kuckartz, U.: *Einführung in die computergestützte Analyse qualitativer Daten*. VS Verlag für Sozialwissenschaften, Wiesbaden (2010)
- Kuckartz, U., Schack, K.: Umweltkommunikation gestalten. Eine Studie zu Akteuren, Rahmenbedingungen und Einflussfaktoren des Informationsgeschehens. Leske + Budric, Opladen (2002)
- Leipziger, J.W.: Konzepte entwickeln. Handfeste Anleitungen für bessere Kommunikation; mit vielen praktischen Beispielen. 2. Aufl. Frankfurter Allgemeine Buch im F.A.Z.-Inst., Frankfurt am Main (2007)
- Lievrouw, L.A., Bucy, E., Frindte, W., Gershon, R., Haythornthwaite, C., Köhler, T., Metz, J., Sundar, S.S.: Current research in new media: an overview of communication and technology. In: Gudykunst, W. (ed.) *Communication Yearbook 24*. Lawrence Erlbaum Publishers, Mahwah (2000)

- Mayring, P.: Qualitative Inhaltsanalyse. Grundlagen und Techniken. Beltz, Weinheim (2010)
- Munkvold, B.E.: Implementation and Use of Collaboration Technologies in a Multinational Engineering Group: The Case of Kvaerner. In: Munkvold, B.E. (ed.) Implementing Collaboration Technologies in Industry: Case Examples and Lessons Learned, pp. 109–128. Springer Verlag, Berlin (2003)
- Pscheida, D., Köhler, T.; Mohamed, B.: What's your favorite online research tool? Use of and attitude towards Web 2.0 applications among scientists in different academic disciplines. In: Marsden, C., Tassiulas, L.: Proceedings of the 1st International Conference on Internet Science. Sigma Orionis, Brussels (2013)
- Pscheida, D., Minet, C., Herbst, S., Albrecht, S., Köhler, T.: Use of social media and online-based tools in academia. Results of the Science 2.0-Survey 2014. <https://nbn-resolving.de/urn:nbn:de:bsz:14-qucosa-191110>. TUD Press, Dresden (2015)
- Roose, J.: Lobby durch Wissenschaft. Umweltverbände und ökologische Forschungsinstitute im Vergleich. Online J. Environ. Policy Stud. (OJEPS) **1** (2006).
- Tiago, F., Gil, A., Stemberger, S., Borges-Tiago, M.: Digital sustainability communication in tourism. J. Innov. & Knowl. (2020). <https://doi.org/10.1016/j.jik.2019.12.002>
- United Nations: Sustainable development goals. <https://sustainabledevelopment.un.org/sdgs> (2015). Accessed 5 Apr 2020
- Weingart, P.: Die Stunde der Wahrheit? Zum Verhältnis der Wissenschaft zu Politik, Wirtschaft und Medien in der Wissensgesellschaft. 1. Aufl. Velbrück Wiss, Weilerswist (2001)
- Weith, T., Barkmann, T., Gaasch, N., Rogga, S., Strauß, C., Zscheischler, J.: Sustainable Land Management in a European Context: A Co-design Approach. Springer, Berlin (2020)
- Weith, T., Köhler, T.: Der Einfluss der Digitalisierung auf die Wissensgenese im Kontext einer nachhaltig-gerechten Entwicklung; Synergie. Fachmagazin für Digitalisierung in der Lehre, Ausgabe #07, <https://www.synergie.uni-hamburg.de> (2019). Accessed 5 Apr 2020
- Witzel, A.: Das problemzentrierte Interview. In: Jüttemann, G. (ed.): Qualitative Forschung in der Psychologie: Grundfragen, Verfahrensweisen, Anwendungsfelder, Beltz, Weinheim, pp. 227–255, <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-5630>, (1985)
- Zscheischler, J., Weith, T., Gaasch, N., Strauß, C., Steinmar, R.: Nachhaltiges Landmanagement – eine kommunikative Herausforderung. Flächenmanagement und Bodenordnung **5**, 37–45 (2012)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Researching Scientific Structures via Joint Authorships—The Case of Virtual 3D Modelling in the Humanities



Sander Münster 

Abstract One of the topics addressed by e-science research is the measurement of academic knowledge production based on electronic data and its relevance in defining the academic landscape. The author employs e-science methods to research cooperative authorships and scientific structures in a specific area of applied e-sciences: virtual 3D modelling in the humanities. Based on the findings, possibilities for cross-disciplinary and international cooperation are discussed. The number of international publications and average number of authors involved in each publication are lower than those found in other scientific fields. Moreover, research indicates that in the humanities, 3D modelling is relatively new and still emergent. Besides such general indications, several key players as people and institutions which interconnect groups of researchers could be identified on a structural level.

Keywords Cooperative authorships · 3D modelling · Humanities

1 Introduction¹

A major issue related to the measurement of academic knowledge production is the distinction between disciplines and the mapping of scientific structures. The vast and heterogeneous variety of possible indicators result in a lack of standardisation and homogenisation. Joint standards to measure academic performance—as intended by the German Research Council (Wissenschaftsrat: Empfehlungen zu einem Kern-datensatz Forschung Berlin 2013)—are still being established. Our field of research is a specific area of applied e-sciences: virtual 3D modelling in the humanities. The research started with selecting a sample of publications in order to investigate current trends, scenarios and workflows in this field, and to quantify the scholarly field. The

¹This article reflects the state of my research in 2015. a more recent state was presented in: Münster, S.: Digital cultural heritage as scholarly field—topics, researchers and perspectives from a bibliometric point of view. *J. Comput. Cult. Heritage* **12**, 22–49 (2019).

S. Münster ()
Friedrich-Schiller-Universität, Jena, Germany
e-mail: sander.muenster@uni-jena.de

initial challenges were to (a) develop a suitable research instrument and to (b) perform an investigation. Due to the limitations of the included information and the magnitude of the data sample, many potentially interesting research approaches—such as a quantification of current topics, standard references and citation networks—are not applicable. The author examines the scientific community involved in this specific area and their level of cross-disciplinary and international cooperation. Furthermore, we identify the key people and institutions, which interconnect groups of researchers.

1.1 Defining Disciplines

To start, a definition: disciplines are characterised by common methods and theories and have similar “reference systems, disciplinary ways of thinking, quality criteria, publication habits and bodies” (Schophaus et al. 2003) as well as similar institutionalisation. Likewise, Knorr-Cetina thought that each discipline has its own “epistemic culture” in the sense of different “architectures of empirical approaches, specific constructions of the referent, particular ontologies of instruments, and different social machines” (Knorr-Cetina 1999). Although disciplines and their boundaries are results of social construction processes (Weingart 1987), a number of phenotypic fields can be identified (Knorr-Cetina 2002). One basic classification scheme is the distinction between humanities and sciences. In a more elaborate classification, the OECD distinguishes between six scientific fields containing about 40 disciplines (OECD 2002, 2007). Furthermore, especially library classification delivers highly sophisticated distinction categorisation schemes (Semenova and Stricker 2007).

1.2 Defining Cross-Disciplinarity

Cross-disciplinarity refers to a “confrontation of several disciplines with a [joint] topic or issue” (Schophaus et al. 2003). In regard to this, Schelsky speaks of a “partial scientific development unit at the empirical object” (Schelsky 1966). Cross-disciplinary collaboration is characterised by developing a multidisciplinary terminology and a joint methodology (Gibbons 1994; Münster et al. 2014). The degree of institutionalisation of cross-disciplinary fields ranges from temporary collaborations to the creation of new “hybrid” research disciplines (Klein 2000) such as the digital humanities, in which computing is applied to foster humanities research.

2 The Case of Virtual 3D Modelling in the Humanities

2.1 Field of Research

3D models and visualisation have always been an important medium for teaching, illustrating and researching historical facts and items. While historical picture sources usually provide elusive and fragmentary impressions, digital three-dimensional models of historical objects and their depictions offer the chance to convey holistic and easily accessible impressions. Until 2000, virtual 3D modelling technologies and computer-generated images of cultural heritage objects were used merely as a digital substitute of physical models (Novitski 1998). Nowadays, 3D models are widely used to present historic items and structures to the public (Greengrass and Hughes 2008) as well as in research (Favro 2004) and education (El Darwich 2005). In addition, 3D technologies can obviously serve cultural heritage management and conservation tasks, and even their advertising. An important distinction needs to be drawn between still extant, no longer extant, and never realised objects. 3D modelling technologies make it possible not only to digitise historic objects which are still extant, but even to virtually reconstruct objects that are no longer extant physically and only known from descriptions.²

Research design

This investigation of scientific structures related to the usage of 3D modelling techniques for both extant and no longer extant types of historical objects is based on an analysis of published project reports and presentations. An upstream problem was the identification of relevant publications. Unlike, for example, in medicine, there are no comprehensive publication databases extant for cultural studies and humanities. Prior to creating the database, three experts—chairholders in the fields of archaeology, art history and geomatics—were queried to identify relevant journals and conferences. This yielded the following findings:

- On the one hand, in the field of cultural and history studies, no multidisciplinary, periodically held international conferences are known that deal specifically with 3D modelling. However, there are a number of local or non-periodic conferences and workshops that deal with specific questions or topics.
- On the other hand, there are four major conferences on the topic in the fields of archaeology and cultural heritage: The International Workshop for 3D Virtual Reconstruction and Visualization of Complex Architectures (3DARCH); the Computer Applications and Quantitative Methods in Archaeology Conference (CAA); the International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST); and the Visualisation in Archaeology workshop (VIA).

²Originally published in: Münster, S., Köhler, T., Hoppe, S.: 3D modeling technologies as tools for the reconstruction and visualization of historic items in humanities. A literature-based survey. In: Traviglia, A. (ed.) Across Space and Time. Papers from the 41st Conference on Computer Applications and Quantitative Methods in Archaeology, Perth, 25–28 March 2013, pp. 430–441. Amsterdam University Press, Amsterdam (2015).

- The Journal for Digital Heritage represents an overarching publication organ for digital content on all humanities.

2.2 Data Sample

These findings formed the basis for collecting the data sample presented in Table 1. As a scope for conference proceedings, entire volumes were included and relevant journal articles were identified via keyword search. A sample of 452 journal articles and conference proceedings was included during the first stage of the analysis. The articles selected were written in English and, for practical reasons, had to be available electronically. Especially the latter selection criterion meant that no publications of the VIA conference and only single volumes of CAA and VAST could be included. In addition to these conference papers, relevant articles from the Journal for Digital Heritage and other periodicals were included using a keyword-based search.

One major obstacle to building a research database was the fact that most of the included conferences and journals were not listed in citation repositories or in publication databases such as ISI Web of Science, Scopus, or Google Scholar in 2012 when the database was compiled. This made it necessary to retrieve metadata by crawling data from each single contribution. For each article, the following information was obtained:

- names and affiliations of contributing authors
- names and addresses of affiliated institutions
- source data (conference and publication name, year, type of document)
- title of publication.

Moreover, conference contributions were classified based on their content. As pointed out in (Münster et al. 2013), one-third (37%) of these articles deal with neither 3D modelling nor historical objects. Nearly the same number of articles report about single projects. This means that they describe workflows for rebuilding certain historic items as 3D models. Another group of contributions deals with certain aspects

Table 1 Sample ($n = 452$)

Publication	Volume	No.
3DArch Conf.	2005–2009	112
CAA Conf.	2007, 2009	130
VAST Conf.	2003–2007, 2010	105
J. Digital Heritage	From 2000 ^a	52
Various project reports and publications dealing with no longer extant objects	1999–2011	79

^aImportant articles were selected via a keyword-based search

of 3D modelling for historical purposes, such as presentation and modelling strategies, data acquisition methods, or handling and classification of 3D data. Focussing on project reports only, a further investigation takes into consideration whether an original object is still extant. To quantify, more than 2/3 of project reports deal with extant objects or their fragments, while another 1/3 focus on non-extant objects. While digitisation of extant objects is mostly based on acquired data and uses widely automated algorithms, reconstruction of no longer extant or never realised objects usually involves manual model creation using CAD or VR software tools. For each type of object, the model creation processes are very different. For this reason, one aim here is to investigate whether both topics might attract different contributors and build slightly different sub-communities.

2.3 *Scientific Approach: Analysis of Scientific Authorship Relations*

From a disciplinary point of view, an investigation of “laws governing the production, flow and application of information in science” (Vinkler 1996) by a numerical analysis of publications is part of bibliometrics. This discipline contains a wide spectre of measures and methods to investigate scientific structures and output. Based on former categorisation and formalisation attempts (Vinkler 2001; Egghe 2009; Gauffriau et al. 2007), several bibliometric approaches are distinguishable, according to their objects of study (Table 2). Not all research approaches are applicable to the described data sample. The limiting factors are the low number of samples and the types of information collected.

The metadata of **publications** are a major object of study. Related approaches include classification of various attributes, like publication type, journal, disciplinary

Table 2 Brief overview of bibliometric approaches

Object of study	Approach (example)
Publications	Classification (i.e. De Solla Price 1963) Scaling laws (i.e. Bettencourt et al. 2008)
Authors	Key numbers (i.e. De Solla Price 1963) Clustering of authors (i.e. Glänzel 2001) Disciplinary productivity (i.e. Lotka 1926)
Topics	Topic graphs (i.e. Schoepflin and Glänzel 2001) Scientograms (i.e. Vargas-Quesada and Moya-Anegón 2007) Epidemiology of ideas (i.e. Garfield 1980)
Citations	Impact (i.e. Hirsch 2005; Smith 2012) Co-citation analysis (i.e. Bellis 2009)
Communities	Structures (i.e. Newman 2001a; Glänzel and de Lange 2002) Protagonists (i.e. Otte and Rousseau 2002; Newman and Girvan 2004; Kretschmer and Aguillo 2004)

backgrounds or dates. These classes allow for comparison and evaluation of distribution functions, as well as monitoring of trends and prediction of emergent fields of research based on time rows (Bettencourt et al. 2008). While the latter approach in particular relies on plenty of lossless data, it does not seem applicable to our research.

Another important object of study is related to **authors** of publications. One approach is to calculate key numbers in various ways, such as an average count of authors per publication or a rate of publications authored by single individuals. As one example, the cutting-edge analyses of De Solla Price (1963) in the early 1960s employed key numbers to investigate the transformation processes of scientific production. A second approach is to cluster authors, for example by nationality, to study preferences for international joint authorships (Glänzel 2001). Both research approaches are employed in our study to investigate **cooperative authorship**. A third approach uses author data to measure disciplinary characteristics such as disciplinary productivity, used in this article by employing the **Lotka Coefficient** (Egghe 2009; Lotka 1926). Furthermore, Schubert and Glänzel studied preference patterns of cross-national authorships (Schubert and Glänzel 2006) and stated that there was a “major influence [of] historical, cultural and linguistic proximities” (p. 426). Such an approach is not applicable to this investigation due to the small number of samples.

Several investigational approaches focus on **topics** described in researched articles. As one example, a topic graph structure classifies current research topics in a certain scientific area (Glenisson et al. 2005; Schoepflin and Glänzel 2001). Moreover, approaches like epidemiology of ideas (Goffman and Newill 1964) or scientograms focus on predicting emergent trends based on an evolution of the importance of topics.

Over the last few years, **citations** have become a very popular object of research into scientific performance. This includes measuring individual impact factors via indexes, most popularly the h-index invented by Hirsch (2005), or the total impact of certain journals via the Garfield index (Vinkler 2012). Furthermore, co-citation analysis provides clues about the evolution of a scientific area over time and its standard works (Bellis 2009). Neither citations nor topics are covered by the available data, so these objects of study are not included in our investigation.

Scientific communities as a “group of scientists [...] agreed on accepting one paradigm” (Jacobs 2006) are another research object of bibliometrics. One particular approach, the study of co-authorship networks, focusses on detecting structures of scientific cooperation employing graph analysis methods (Vargas-Quesada and Moya-Anegón 2007). Although such research approaches are limited to a structure representation (Hardeman 2013) and include a number of potential sources of error and limitations, computer-based analysis and evaluation of co-authorships fosters several new insights related to scientific cooperation (De Stefano et al. 2011; Lu and Feng 2009). For example, a comprehensive investigation of publications in the fields of medicine, science and computing (Newman 2001a, b, c) reveals that the “small-world phenomenon” (Milgram 1967) (i.e. any two authors are connected in a chain of on average five to six parties) could be identified for these scientific communities. A number of smaller studies also deal with co-authorship within individual disciplines (Aleixandre-Benavent et al. 2012), or for individual countries or regions

(Abramo et al. 2010; Morelli 1997; Gaillard 1992). Besides describing scientific networks, another issue is to identify important players as **protagonists of scientific communities** (Kretschmer and Agullo 2004; Hou et al. 2007). This latter aspect is of interest regarding the community dealing with 3D modelling in the humanities.

3 Findings

3.1 Indication 1: Cooperative Authorship

One of the essential characteristics of modern research is the large number of authors involved in a single publication. In 1962, De Solla Price pointed out that in 1900, more than 80% of publications had a single author (De Solla Price 1963). In 2000, a study of scientific articles listed in the Science Citation Index (Glänzel et al. 2004) revealed an average contribution of 4.2 authors per article, wherein the proportion of articles written by individual authors was only 11%. Within our research sample, an average of 3.4 authors was involved in each publication. From the perspective of cross-disciplinary and international cooperation, the disciplinary affiliation of the author collectives seems especially interesting. As shown in Fig. 1, the majority of the studied publications were written by authors or author collectives belonging to the same area of research and only a limited number of publications were cross-disciplinary. The author's disciplinary affiliation was identified from the correspondence addresses noted in publications. However, such data only provides information about the disciplinary focus of an employing institution and not on the author himself. To overcome this potential flaw, an alternative method which takes personal disciplinary backgrounds into account—self-sorting by authors via questionnaire—is intended for the next stage of the research, but not yet realised for this set of data. In this data, for 21% of authors the respective disciplines at affiliated institutions could not be identified or distinguished precisely.

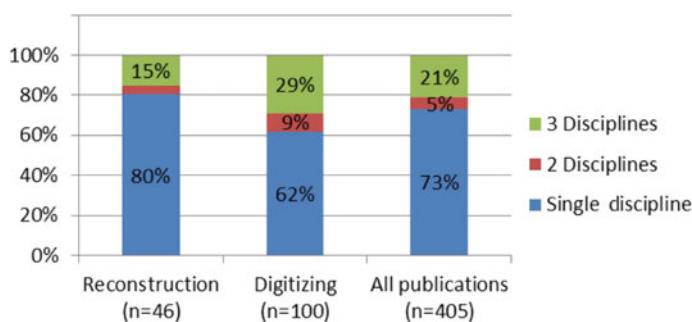


Fig. 1 Number of participating disciplines

With regard to the distinction between types of modelling, the number of cross-disciplinary publications describing the digitisation of extant objects is significantly higher than for reconstruction projects.

Another interesting aspect is the disciplinary background of the authors' employing institutions depending on the type of object modelled. In the table shown in Fig. 2, cross-disciplinary collaborations were included proportionately and each cross-disciplinary publication has been counted with 1, while for publications including two disciplines, each of them has been counted with 0.5. It seems remarkable that a large number of articles describing reconstruction projects of unrealised or non-extant objects were written by authors affiliated with institutions in the field of architecture, while publications for digitisation projects were often written by authors with a background in engineering and geosciences. A plausible explanation is provided by the competence profiles of these departments. For example, automated data acquisition via remote sensing techniques is a focus of the geosciences, while architectural studies incorporate extensive know-how about both architectural history and CAD modelling. Figure 3 shows cross-disciplinary authorships in the researched publications. Each node stands for a single publication and each edge represents the disciplinary assignment of the participating authors. The graph shows that authors from institutions in the digital humanities are especially frequently involved in cross-disciplinary cooperative authorships. Preferred partners are authors from institutions in the field of computer science, while joint publication with authors from the humanities tends to occur rarely.

As shown in Fig. 4, a significant number of publications were written by authors whose employing institutions are in the same nation. Compared to findings related to other scientific domains, which estimate an overall rate of international publications at 35% (Acosta et al. 2010), the number of international publications in the sample is

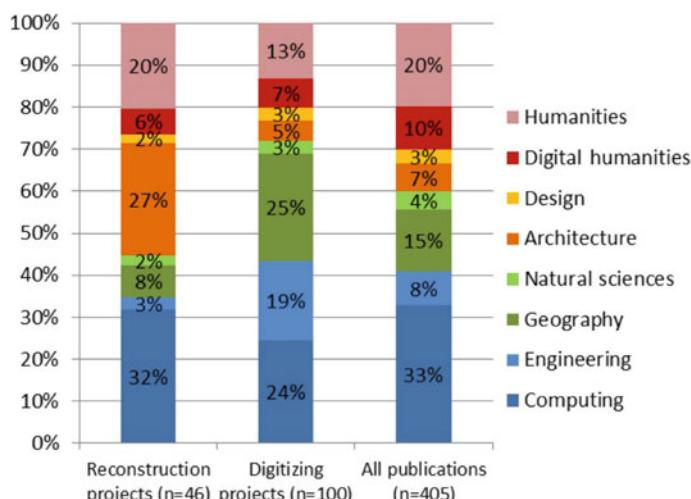


Fig. 2 Disciplinary affiliation of publication authors

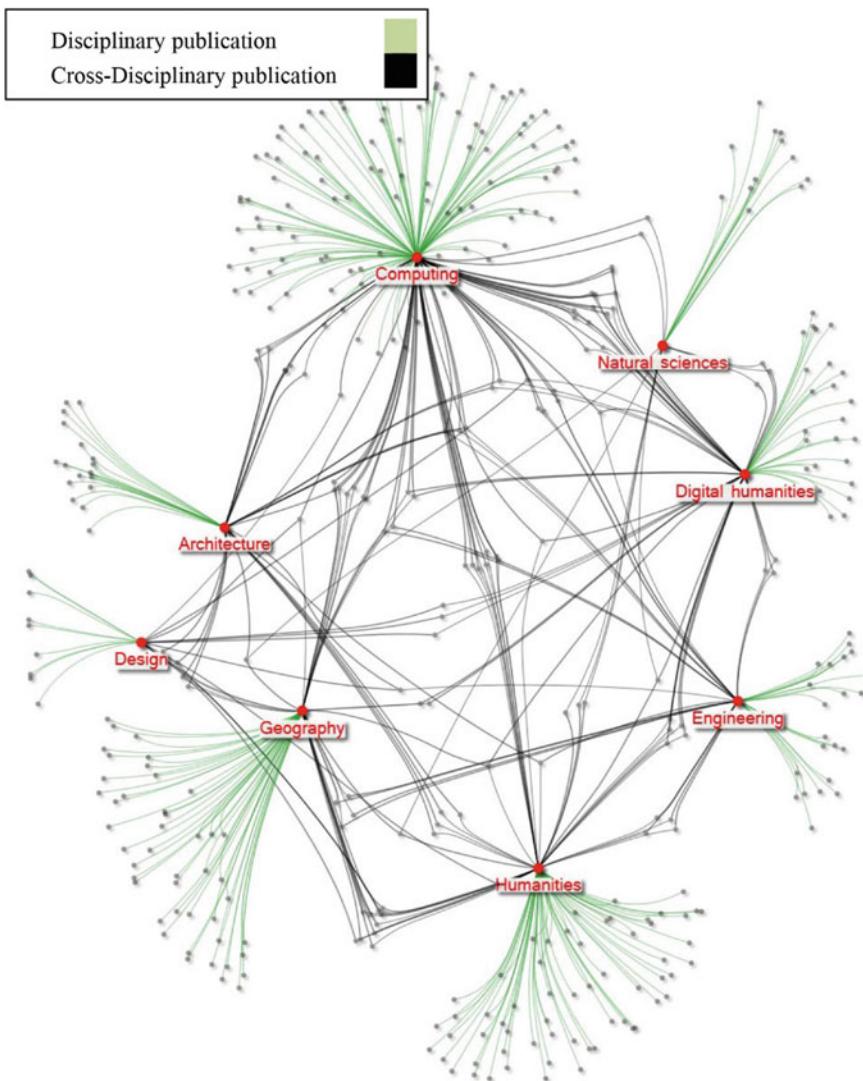


Fig. 3 Cross-disciplinary authorship

significantly lower. Analogous to the findings related to interdisciplinary cooperation, the number of international publications describing digitisation projects is above average, while only 8% of the publications describing reconstruction projects were written by international teams.

The findings of a below average rate of international and cross-disciplinary authorships in combination with a large variety of involved disciplines indicate the fuzzy

demarcation of the field. This assumption is supported by the finding that only 30% of authors are employed in institutions which prioritise humanities or digital humanities.

3.2 Indication 2: Lotka Coefficient

One of the most common indicators is the number of publications per author. Relatively, Lotka (1926) developed a distribution function for the publication frequency of individual authors which covers a wide range of disciplines and their publications. The distribution curve shows that a large number of authors with only one publication are contrasted by a very small number of authors with multiple publications.

Related to the investigated publication data, a classical Lotka distribution already revealed an extensive congruence (Fig. 5). This follows the formula

$$Y = C/X^n$$

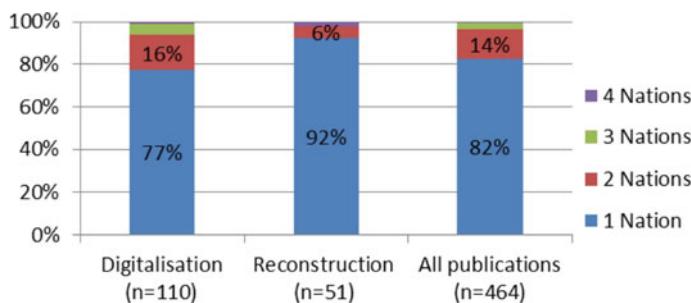


Fig. 4 Cross-national publications

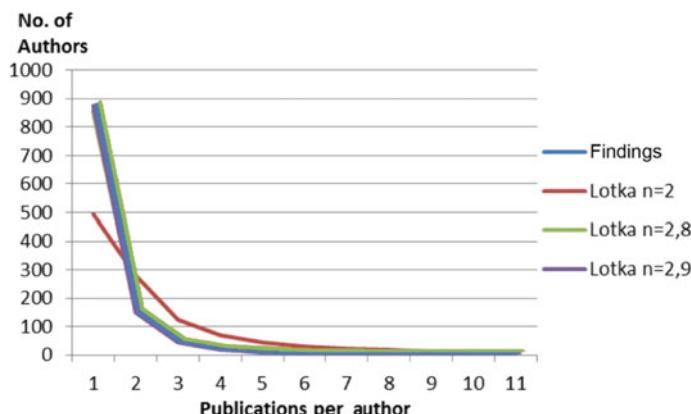


Fig. 5 Frequency distribution curves for publications

where C is the total number of authors included ($n = 1120$) and X indicates the number of publications of each cohort (authors with 1, 2 or more publications). The exponent n is a constant. From his studies, Lotka postulated an average exponent of $n = 2$, which varies significantly depending on the investigated discipline (Egghe and Rousseau 1990; Egghe 2000), while recent studies assumed an average value of 2.3 to 2.5 (Chung and Kolbe 1992; Pulgarín 2012). In the empirical findings, the distribution function of the investigated publications coincides with $n = 2.8 \dots 2.9$. Any further interpretation of these values must be estimated in the context of the relatively small and potentially flawed sample. Compared to the lower mean values of the exponent mostly cited in literature, the above-average exponent found here indicates low publication productivity with a disproportionate number of authors who are only occasionally involved.

3.3 Indication 3: Key Players³

Another hypothesis is that collaborative publications establish knowledge communication between authors. The basic idea is that, in most cases, common authorship would be related to a personal connection and interaction between all included authors. Depending on sociological role theory, such a connection between people—regardless of its strength (Granovetter 1973)—could foster sharing and exchange of ideas and information. Regarding structure, connections between people across disciplinary and national borders play a key role in disseminating information in social communities.⁴ Nevertheless, information transfer in the context of joint publications is just assumed and intensity or even information transfer between authors cannot be reconstructed based on empirical data.

The sample publications were authored by 1500 individuals who were connected by over 3000 links (Fig. 6). Most of the publications were written by authors belonging to institutions of the same discipline and nationality. All the individuals at each institution were incorporated into Fig. 7. Key players were highlighted in the graphs: these were the people and institutions that were in the top ten in the categories of (a) number of connections to other authors (*degree*), (b) the relevance as a connecting factor between author groups (*betweenness centrality*) or (c) the number of publications. (Wasserman and Faust 1994) But there are also several international or cross-disciplinary networks visible whose members have written more than just one joint publication. It was possible to identify some important key players

³Originally published in: Münster, S., Köhler, T., Hoppe, S.: 3D modeling technologies as tools for the reconstruction and visualization of historic items in humanities. a literature-based survey. In: Traviglia, A. (ed.) Across Space and Time. Papers from the 41st Conference on Computer Applications and Quantitative Methods in Archaeology, Perth, 25–28 March 2013, 430–441. Amsterdam University Press, Amsterdam (2015).

⁴There are several studies on scientific communities and inherent social interaction, i.e. Stützer, C.: Knowledge transfer in web-based collaborative learning systems (PhD-Thesis), Dresden 2013.

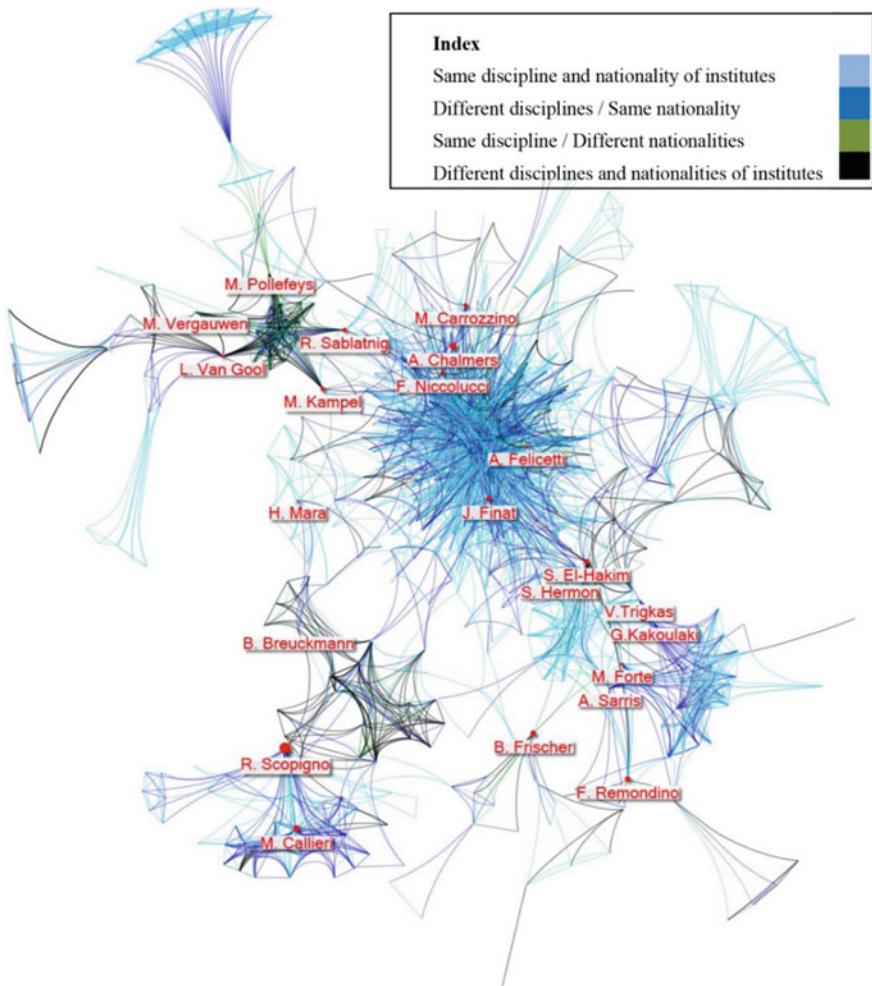


Fig. 6 Author—co-author relations—individuals (*key players* highlighted)

who connect groups of researchers. From an institutional perspective, the cooperation between the University of Leuven and the technical universities of Vienna and Zürich has produced a particularly large number of cross-disciplinary and international publications. A further, if smaller, cluster includes mostly French and Italian institutions, but also encompasses authors from Japan and Germany. Generally, there is a high level of networking and number of publications from people and institutions working on data-based visualisation. Finally, the key players are most connected, both internationally and cross-disciplinarily. To validate this, the results were discussed with experts. Generally, these key players are not only active publishers, but often

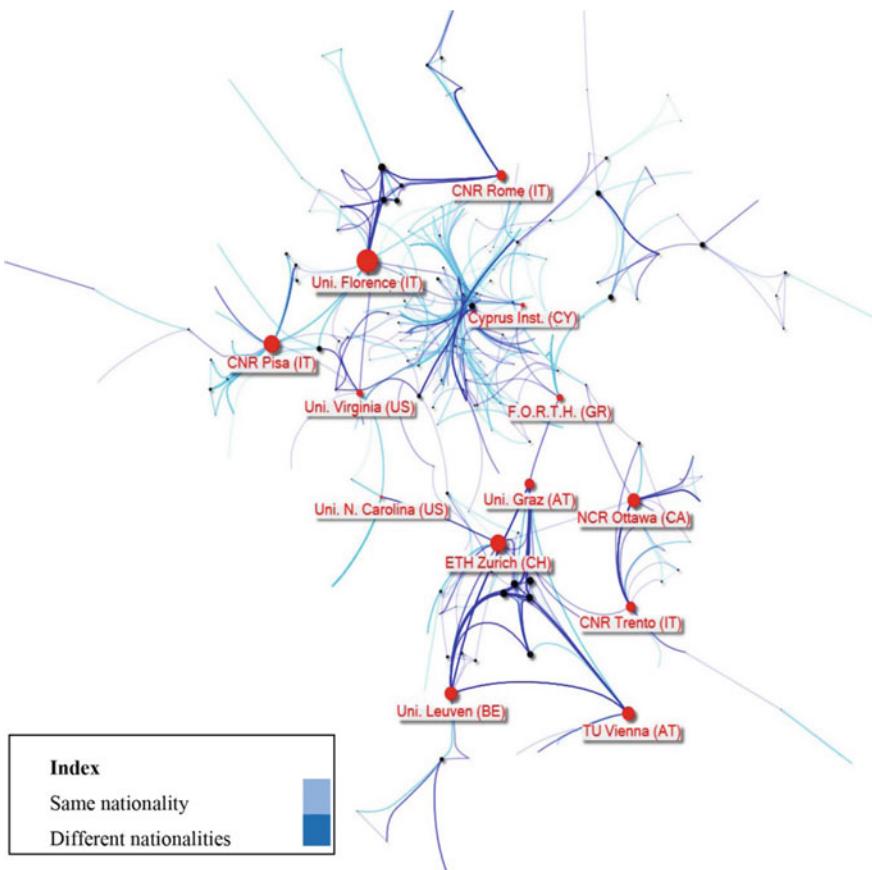


Fig. 7 Author—co-author relations—institutions (*key institutions* highlighted)

also play key roles in the community in other ways, too, whether as members of scientific committees, conference chairs, and initiators or leaders of projects.

We also investigated the connection between theory and practice in the field of 3D modelling in the humanities. We compared, for each institution, the number of participating digitisation and reconstruction projects described in articles and the number of publications. The results show that institutions with a high publication output are usually also involved in an exceptionally large number of projects. A significant difference between ranks of publication activity and project participation was identified in just a few institutions, such as the TU Wien or the Istituto di Scienza e Tecnologie dell'Informazione (Table 3).

This leads to the assumption that a scientific community is primarily a *community of practice* (Lave and Wenger 1991), with a close link between practical project work and theory, while specific think tanks as theory building institutions are currently not visible for this field of research.

Table 3 Ranks of project and publication participation by institution

Institution	Projects (Rank)	Publications (Rank)
Politecnico di Milano	10 (1)	14 (1)
ETH Zürich	8 (2)	12 (4)
National Research Council Canada	7 (3)	11 (5)
University of Florence	7 (4)	12 (3)
Istituto di Scienza e Tecnologie dell'Informazione	6 (5)	13 (2)
Centre for Scientific and Technological Research	6 (6)	7 (7)
University of Virginia	5 (7)	6 (14)
Tokyo Denki University	4 (8)	6 (12)
CNR Institute of Technology Applied to Cult. Heritage	4 (9)	6 (15)
TU Wien	4 (10)	11 (6)

4 Conclusion

With regard to the aim of identifying scientific structures via co-authorships, it was found that the field of 3D modelling in the humanities at an international level is widely dominated by research interests and approaches from archaeology and cultural heritage research. However, the authors involved come from a large variety of disciplinary backgrounds.

Another finding is that the number of publications written by international teams and average number of authors involved in each publication are lower than in other scientific fields. It seems remarkable that publications about digitisation projects which deal with extant objects are significantly more often written by cross-disciplinary and international teams than publications describing a reconstruction of no longer extant or never realised objects. This may be caused by the slightly different disciplinary constitutions and uses of publications in digitisation and reconstruction projects. Taking the relatively small and potentially flawed sample into account, further investigations and additional data are required for a valid evaluation.

3D modelling in the humanities is relatively new and emergent field of research. This is indicated by the above-average coefficient for a Lotka distribution describing the frequency of publications per author. Nevertheless, contributors from various disciplines were involved in the researched publications, which may indicate a currently blurry demarcation of the scientific field. Even if these findings are endorsed by other studies (Albrecht 2013), both indications provide only a hint that the field is becoming established.

What are the implications for research on e-sciences? This article described several strategies for investigating scientific structures in the field of 3D modelling in the humanities, based on electronic data and using software tools for graph analysis and QDA software for qualitative content analysis. Information was retrieved about structures and publication practices in the field. It was found that the investigated publications were mostly about archaeology and cultural heritage, while other research

interests like aspects of cultural or art history were treated mostly via national communities and published in offline media. While the research objects and issues are closely related to the humanities, just a minority of authors are affiliated with (digital) humanities, and authors with a background in computing are very prominent in the publications. Even though these findings require further investigation, they may indicate that an international community in digital humanities is less influenced by practitioners whose competence relates to the research questions and objects than by those who provide digital research methods.

References

- Abramo, G., D'Angelo, C.A., Solazzi, M.: The relationship between scientists' research performance and the degree of internationalization of their research. *Scientometrics* **86**, 629–643 (2010)
- Acosta, M., Coronado, D., Ferrández, E., León, M.D.: Factors affecting inter-regional academic scientific collaboration within Europe. the role of economic distance. *Scientometrics* **87**, 63–74 (2010)
- Albrecht, S.: Adoption of eScience Practices among Scholars in the Humanities and Social Sciences (Presentation). 2nd Network Conference des E-Science Forschungsnetzwerks Sachsen, June 11–13, 2013, Dresden (2013)
- Alexandre-Benavent, R., Aleixandre-Tudo, J.L., Alcaide, G.G., Ferrer-Sapena, A., Aleixandre, J.L., Toit, W.d.: Bibliometric analysis of publications by South African viticulture and oenology research centres. *S. Afr. J. Sci.* **108**, 74–84 (2012)
- Bellis, N.D.: Bibliometrics and Citation Analysis from the Science Citation Index to Cybermetrics, Plymouth (2009)
- Bettencourt, L.M.A., Kaiser, D.I., Kaur, J., Castillo-Chávez, C., Wojick, D.E.: Population modeling of the emergence and development of scientific fields. *Scientometrics* **75**, 495–518 (2008)
- Chung, K.H., Kolbe, P.T.: Empirical regularities in the market for real estate research output. *J Real Estate Res* **7**, 115–124 (1992)
- De Solla Price, D.: Little Science—Big Science. Columbia University Press, New York (1963)
- De Stefano, D., Giordano, G., Vitale, M.: Issues in the analysis of co-authorship networks. *Qual. Quant.* **45**, 1091–1107 (2011)
- Egghe, L.: A heuristic study of the first-citation distribution. *Scientometrics* **48**, 345–359 (2000)
- Egghe, L.: Performance and its relation with productivity in Lotkaian systems. *Scientometrics* **81**, 567–585 (2009)
- Egghe, L., Rousseau, R.: Introduction to Informetrics. Quantitative Methods in Library, Documentation and Information Science. Elsevier, Amsterdam (1990)
- El Darwich, R.: Modelle. In: Pandel, H.-J., Becher, U.A.J. (eds.) *Handbuch Medien im Geschichtsunterricht*, pp. 580–591. Wochenschau-Verlag, Schwalbach (2005)
- Favro, D.: In the eyes of the beholder. Virtual Reality re-creations and academia. In: Haselberger, L., Humphrey, J., Abernathy, D. (eds.) *Imaging Ancient Rome: Documentation, Visualization, Imagination: Proceedings of the 3rd Williams Symposium on Classical Architecture*, Rome, 20. 23 May 2004, pp. 321–334. J. Rom. Archaeolo, Portsmouth (2006)
- Gaillard, J.: Scientific production in developing countries. *Scientometrics* **23**, 57–73 (1992)
- Garfield, E.: The epidemiology of knowledge and the spread of scientific information. *Essays of an information scientist* **4**, 586–591 (1980)
- Gauffriau, M., Larsen, P.O., Maye, I., Roulin-Perriard, A., Ins, M.: Publication, cooperation and productivity measures in scientific research. *Scientometrics* **73**, 175–214 (2007)

- Gibbons, M.: The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies. Sage, London (1994)
- Glänzel, W.: National characteristics in international scientific co-authorship relations. *Scientometrics* **51**, 69–115 (2001)
- Glänzel, W., De Lange, C.: A distributional approach to multinationality measures of international scientific collaboration. *Scientometrics* **54**, 75–89 (2002)
- Glänzel, W., Schubert, A.: Analyzing scientific networks through co-authorship. In: Moed, H.F., Glänzel, W., Schmoch, U. (eds.) *Handbook of Quantitative Science and Technology Research. The Use of Publication and Patent Statistics in Studies of S&T Systems*, pp. 257–276. Kluwer Academic Publishers, Dordrecht (2004)
- Glenisson, P., Glänzel, W., Janssens, F., De Moor, B.: Combining full text and bibliometric information in mapping scientific disciplines. *Inf. Process. Manag.* **41**, 1548–1572 (2005)
- Goffman, W., Newill, V.A.: Generalization of epidemic theory, an application to the transmission of ideas. *Nature* **204**, 225–228 (1964)
- Granovetter, M.S.: Strength of the weak ties. *Am. J. Sociol.* **78**, 1360–1380 (1973)
- Greengrass, M., Hughes, L.M.: *The Virtual Representation of the Past*. Ashgate, Aldershot (2008)
- Hardeman, S.: Organization level research in scientometrics: A plea for an explicit pragmatic approach. *Scientometrics* **94**, 1175–1194 (2013)
- Hirsch, J.E.: An index to quantify an individual's scientific research output. *Proc. Natl. Acad. Sci. U. S. Am.* **102**, 16569–16572 (2005)
- Hou, H., Kretschmer, H., Liu, Z.: The structure of scientific collaboration networks in scientometrics. *Scientometrics* **75**, 189–202 (2007)
- Jacobs, S.: Models of scientific community. Charles Sanders Peirce to Thomas Kuhn. *Interdisc. Sci. Rev.* **31**, 163–173 (2006)
- Klein, J.T.: A conceptual vocabulary of interdisciplinary science. In: Weingart, P., Stehr, N. (eds.) *Practising Interdisciplinarity*, pp. 3–24. University of Toronto Press, Toronto (2000)
- Knorr-Cetina, K.: *Epistemic Cultures. How the Sciences Make Knowledge*. Harvard University Press, Cambridge (1999)
- Knorr-Cetina, K.: *Die Fabrikation von Erkenntnis*. Frankfurt a. M., Suhrkamp (2002)
- Kretschmer, H., Agullo, I.F.: Visibility of collaboration on the web. *Scientometrics* **61**, 405–426 (2004)
- Lave, J., Wenger, E.: *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, Cambridge (1991)
- Lotka, A.J.: The frequency distribution of scientific productivity. *J. Wash. Acad. Sci.* **16**, 317–323 (1926)
- Lu, H., Feng, Y.: A measure of authors centrality in co-authorship networks based on the distribution of collaborative relationships. *Scientometrics* **81**, 499–511 (2009)
- Milgram, S.: The small-world problem. *Psychol. Today* **1**, 60–67 (1967)
- Morelli, W.V.: *Computergeschützte Bildrecherche und Bildanalyse*. In: Kohle, H. (ed.) *Kunstgeschichte digital. Eine Einführung für Praktiker und Studierende*. Warburg, Berlin (1997)
- Münster, S.: Digital cultural heritage as scholarly field—topics, researchers and perspectives from a bibliometric point of view. *J. Comput. Cult. Heritage* **12**, 22–49 (2019)
- Münster, S., Köhler, T., Hoppe, S.: 3D modeling technologies as tools for the reconstruction and visualization of historic items in humanities. A literature-based survey. In: Traviglia, A. (ed.) *Across Space and Time. Papers from the 41st Conference on Computer Applications and Quantitative Methods in Archaeology*, Perth, 25–28 March 2013, pp. 430–441. Amsterdam University Press, Amsterdam (2015)
- Münster, S., Schlenker, L., Köhler, T.: Common grounds and representations in cross-disciplinary processes. In: Carlucci, D., Schiuma, G., Spender, J. (eds.) *Knowledge and Management Models for Sustainable Growth*, pp. 579–589. Institute of Knowledge Asset Management, Basilicata (2014)
- Newman, M.E.J.: The structure of scientific collaboration networks. *Proc. Natl. Acad. Sci. U. S. Am.* **98**, 404–409 (2001a)

- Newman, M.E.J.: Scientific collaboration networks. I. Network construction and fundamental results. *Phys. Rev. E* **64**, 016131 (2001b)
- Newman, M.E.J.: Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Phys. Rev. E* **64**, 016132 (2001c)
- Newman, M.E., Girvan, M.: Finding and evaluating community structure in networks. *Phys. Rev. E* **69**, (2004)
- Novitski, B.J.: *Rendering Real and Imagined Buildings: The Art of Computer Modeling from the Palace of Kublai Khan to Le Corbusier's Villas*. Rockport Pub, Gloucester (1998)
- Organisation for Economic Co-operation and Development: *Frascati Manual. Proposed Standard Practice for Surveys on Research and Experimental Development*. OECD Publications Service, Paris (2002)
- Organisation for Economic Co-operation and Development: *Revised Field of Science and Technology (FOS) Classification in the Frascati Manual*, Paris (2007)
- Otte, E., Rousseau, R.: Social network analysis: A powerful strategy, also for the information sciences. *J. Inf. Sci.* **28**, 441–453 (2002)
- Pulgarín, A.: Dependence of Lotka's law parameters on the scientific area. *Malays. J. Libr. Inf. Sci* **17**, 41–50 (2012)
- Schelsky, H.: Das Zentrum für interdisziplinäre Forschung. In: Mikat, P., Schelsky, H. (eds.) *Grundzüge einer neuen Universität*, pp. 71–87. Bertelsmann, Gütersloh (1966)
- Schoepflin, U., Glänzel, W.: Two decades of scientometrics—an interdisciplinary field represented by its leading journal. *Scientometrics* **50**, 301–312 (2001)
- Schophaus, M., Dienel, H.-L., von Braun, C.-F.: Von Brücken und Einbahnstraßen. Aufgaben für das Kooperationsmanagement interdisziplinärer Forschung (Discussion paper Nr. 08/03), Berlin (2003)
- Schubert, A., Glänzel, W.: Cross-national preference in co-authorship. *Scientometrics* **69**, 409–428 (2006)
- Semenova, E., Stricker, M.: Eine Ontologie der Wissenschaftsdisziplinen. Entwicklung eines Instrumentariums für die Wissenskommunikation. In: Ball, R. (ed.) *Wissenschaftskommunikation der Zukunft*, 4. Konferenz der Zentralbibliothek im Forschungszentrum Jülich, 6. 8. 11. 2007, vol. Band 18, pp. 61–69. *Schriften des Forschungszentrums Jülich*, Jülich (2007)
- Smith, D.R.: Impact factors, scientometrics and the history of citation-based research. *Scientometrics* **92**, 419–427 (2012)
- Stützer, C.: Knowledge transfer in web-based collaborative learning systems (PhD-Thesis), Dresden (2013)
- Vargas-Quesada, B., Moya-Anegón, F.D.: *Visualizing the Structure of Science*. Springer, Berlin [u. a.] (2007)
- Vinkler, P.: Some practical aspects of the standardization of scientometric indicators. *Scientometrics* **35**, 237–245 (1996)
- Vinkler, P.: An attempt of surveying and classifying bibliometric indicators for scientometric purposes. *Scientometrics* **50**, 539–544 (2001)
- Vinkler, P.: The Garfield impact factor, one of the fundamental indicators in scientometrics. *Scientometrics* **92**, 471–483 (2012)
- Wasserman, S., Faust, K.: *Social Network Analysis*. Cambridge University Press, Cambridge (1994)
- Weingart, P.: Interdisziplinarität als List der Institutionen. In: Kocka, J. (ed.) *Interdisziplinarität. Praxis - Herausforderung - Ideologie*, pp. 159–166. Suhrkamp, Frankfurt a. M. (1987)
- Wissenschaftsrat: Empfehlungen zu einem Kerndatensatz Forschung Berlin (2013)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Visions of a Future Research Workplace Arising from Recent Foresight Exercises



Andrzej M. J. Skulimowski

Abstract The results of recent foresight projects reveal the impact of future ICT tools on the practice of scientific research. This paper presents several aspects of the process of building scenarios and trends of selected advanced ICT technologies. We point out the implications of emerging global expert systems (GESs) and AI-based learning platforms (AILPs). GESs will be capable of using and processing global knowledge from all available sources, such as databases, repositories, video streams, interactions with other researchers and knowledge processing units. In many scientific disciplines, the high volume, density and increasing level of interconnection of data have already exhausted the capacities of any individual researcher. Three trends may dominate the development of scientific methodology. Collective research is one possible coping strategy: Group intellectual capacity makes it possible to tackle complex problems. Recent data flow forecasts indicate that even in the few areas, which still resist ICT domination, research based on data gathered in non-ICT supported collections will soon reach its performance limits due to the ever-growing amount of knowledge to be acquired, verified, exchanged and communicated between researchers. Growing automation of research is the second option: Automated expert systems will be capable of selecting and processing knowledge to the level of a professionally edited scientific paper, with only minor human involvement. The third trend is intensive development and deployment of brain–computer interfaces (BCIs) to quickly access and process data. Specifically, GESs and AILPs can be used together with BCIs. The above approaches may eventually merge, forming a few AI-related technological scenarios, as discussed to conclude the paper.

Keywords ICT foresight • E-science • Technological development trends • Global expert systems • Brain–computer interfaces • Artificial intelligence (AI) scenarios

A. M. J. Skulimowski (✉)

Department of Automatic Control and Robotics, AGH University of Science and Technology,
Kraków 30-059, Poland
e-mail: ams@agh.edu.pl

International Centre for Decision Sciences and Forecasting, Progress & Business Foundation,
Kraków 30-048, Poland

1 Introduction

Based on the results of the foresight project, SCETIST (Skulimowski 2013), and a Delphi study on future development trends of knowledge platforms performed within the recent Horizon 2020 project MOVING (Köhler and Skulimowski 2019), this paper aims to provide an insight into the future of e-science. The focus is on three specific aspects of this perspective: the emergence of new research tools related to global expert systems (GESs), researcher communication with computers through brain–computer interfaces (BCIs), and the role of researchers in shaping holistic knowledge development systems that will emerge over the next few decades.

The aims of the aforementioned foresight projects include making recommendations to R&D and ICT policymakers, while pointing out prospective ICT development and research trends relevant to individual researchers and research teams. The time horizon of foresight was 2025, with an impact analysis of selected anticipated technological breakthroughs up to 2030. Some of the project results related to e-science are presented in Skulimowski (2016b); the results on the emergence of GESs are published in Skulimowski (2013), while the relation to artificial autonomous decision systems (AADSs) is discussed in Skulimowski (2014b, 2016b).

A diverse spectrum of methods was applied to elaborate on technological and social scenarios and forecasts. Those used predominantly included bibliometric analyses, extrapolation Delphi surveys (Skulimowski 2019), group building of a hierarchical state-space model of information society evolution (Skulimowski et al. 2013) and anticipatory networks (Skulimowski 2014a).

For the purposes of e-science foresight, the computer-assisted multi-round expert Delphi questionnaire retrieval (cf. e.g. Skulimowski et al. 2013, 2019), combined with expert panel meetings and outcomes of bibliometric and patentometric research proved most useful within the overall project. The analysis of expert responses was combined with an information retrieval strategy from the open Web and from major bibliographic databases. Different procedures were elaborated for fusing quantitative and qualitative knowledge and providing recommendations to the ICT industry and policymakers. A trust and competence factor system was used to compensate for the impact of diverse expert biases and competences. Each survey respondent was assigned a vector with trustworthiness coefficients of this expert in the particular subject areas of the Delphi exercise. A weighted combination of individual responses with coordinates of the trustworthiness vector was applied, wherever appropriate, to take account of the difference in respondents' credibility.

Section 2 outlines certain basic ICT/AI development trends that may influence future research tools. The roles played by AI-based learning platforms (AILPs) and GESs will gain importance when fusing ever-growing information flows, culminating in deeper automatic data refinery before presenting them to researchers. GESs will be capable of processing “big data” to “big knowledge”. New knowledge fusion methods will be developed, such as hybrid and scenario-based anticipatory networks (Skulimowski 2014a), e-science foresight (Skulimowski 2016b), including combinations of forecasts (Elliott and Timmermann 2004) or recommendations

(Skulimowski 2017a). Finally, Sect. 3 presents the results of the Delphi surveys on information systems prospects, which were conducted for SCETIST and MOVING projects (Skulimowski et al. 2013; Köhler and Skulimowski 2019). We show that different technological trends will have a synergetic impact on e-science. Artificial intelligence-based (AI) tools and approaches will play a major role. New tools will make the research conducted by humans more efficient by reaching predefined goals faster and more accurately.

Recommendations that may be useful to R&D policymakers, artificial intelligence researchers, and innovative companies will be presented in Sect. 4. We will also explore the relationship between BCIs and the future methodology of storing and processing scientific information in GESs and AILPs. Moreover, Sect. 4 discusses the opportunities, challenges, and threats posed by the development of AI tools and how BCIs could be used to quickly overcome the problem of accessing big data streams and knowledge repositories.

2 Integration of Future Research Tools in Global Expert Systems

GESs were originally intended as a generalization of large-scale expert surveys and intelligent digital libraries (Leidig and Fox 2014), capable of merging heterogeneous information. They were defined in Skulimowski (2013, p. 582) as “all knowledge sources, sensors, databases, repositories, and processing units, regardless of whether they are human, artificial, animal, or hybrid, provided that they are all mutually connected and endowed with … the usual expert system functionalities.” Nodes of a GES are marked as “users” and each GES has a specific user hierarchy. Moreover, a GES must offer each user an efficient information management system providing “knowledge transfer on immediate demand” (*ibid.*).

The growing coverage of scientific information by search engines, with an increasing share of open access resources, further enhances the capabilities of autonomous information retrieval, which is the base of the GES paradigm. In the e-science context, the rationale justifying the introduction of GESs is to determine rules and principles for the design of knowledge-based systems capable of gathering and processing big scientific data, information and knowledge at different stages of verification and refinery. The access of autonomous webcrawlers and other GES tools to paid or sensitive information sources may be ensured with automatic subscription passwords or automatic micropayments and may be facilitated by distributed ledger technologies such as Linux Foundation’s Hyperledger Fabric blockchain (Thakkar et al. 2018). It is also assumed that the researchers will pursue the trend to upload the results of their work to public open access repositories such as [researchgate.net](#), [zenodo.org](#), or [academia.edu](#).

The development of GES and the simultaneous emergence of AILPs will ensure similar progress in learning approaches (Skulimowski 2019). It has also been

argued (Skulimowski 2013) that GESs may play an important role in solving the human–computer convergence problem, which touches upon the AILPs as well. The following Internet development trends that support the above claims were identified in Skulimowski (2013, 2014b):

- growing integration of heterogeneous information sources (ISs);
- increasing interconnection of knowledge units, online and offline;
- increasing sophistication of information processing within each knowledge unit;
- growing availability of sensor and other scientific measurement data, including information from Internet of things (IoT);
- growing need to apply big data technologies in scientific information processing driven by the overall growth of the amount of information available online;
- the emergence of common standards for scientific information management (Jeffery et al. 2014).

The above trends are amplified by qualitative and quantitative refinement of the information stored and processed online as well as by the growing availability of the learning content. The latter is fed to AILPs and boosts their development.

The usability of online information for scientific purposes depends upon how well it is structured and accessible via search engines. For instance, the percentage of all data stored on the open Web and indexed by the search engine Google rose from 1% in January 2007 to 6% in January 2010 and exceeded 10% in January 2012. This estimated ratio has been preserved until at least 2019. At the same time, the estimated amount of information available online rose to 800 exabytes (10^{18} B) in 2009 and 1.3 zettabytes (10^{21} B) in 2013. According to the Delphi survey in Skulimowski et al. (2013), question [I.8], it is expected to rise to 1.6 zettabytes in 2020 and to reach the value of 3.5 zettabytes in 2025 and about 7 zettabytes in 2030. The recent Internet metrics data¹ yield the value of 2 zettabytes of information contained in indexed Web sites as of 2019, which does not deviate much from the Delphi forecasts from 2012 to 2013 (Skulimowski et al. 2013). The same survey provided replies to the question of whether the information available online is really useful to scientists. The results are presented in Sect. 3.

The number of Web sites exceeded 1700 million in 2016,² then slightly declined and rose again to 1730 million in 2019 (Mill provides the value of 1.27×10^9 as of December 2019). Only 15% of all Web sites are active.³ They are hosted in about 360 million top-level domains.⁴ Forecasts of a further increase until 2025 and beyond diverge considerably depending on whether exclusively machine-operated and used (M2M) sites in the Internet of things are considered or not. Estimations vary between 3 and 50 billion sites in 2025. The number of Web pages indexed by

¹<https://www.statista.com/statistics/267202/global-data-volume-of-consumer-ip-traffic/> [access Jan 10, 2020].

²An estimate after <http://en.wikipedia.org/wiki/Exabyte> [access Jan 10, 2020].

³<https://www.millforbusiness.com/how-many-websites-are-there/> [access Jan 10, 2020].

⁴https://www.verisign.com/en_US/domain-names/dnib/index.xhtml [access Jan 10, 2020].

Google and Bing rose to 6.27×10^{12} in January 2020.⁵ When the tools offered by search engines become sufficiently sophisticated, this system of interconnected Web sites may become a real GES with strong analytic capacities.

Another salient trend shaping the future of e-science is the emergence of a new form of collaborative learning (Köhler and Skulimowski 2019) that is facilitated and made more efficient with AILPs. This trend supports collaborative research, the overall growth of collective intelligence of research teams (Mohamed et al. 2013) and their fusion in GESs. Although in the mid-term future, the intellectual capacity of scientists can be outperformed by autonomous “global brain” type analytic engines (Heylighen 2017), using GESs and AILPs as the composite tools for learning and research will keep them aligned to the recent progress of autonomously performed research. In addition, the “explainable AI” paradigm (Xu et al. 2019), when commonly applied, can use combined GESs and AILPs as tools to make available the results of any kind of autonomous research in a comprehensible form for any GES/AILP user.

Internet-based information supply chains of constantly growing size and complexity necessitate new approaches to designing search-and-survey procedures and to delegating more of this design work to autonomous agents. In a *creative decision process* (Skulimowski 2011), the user defines an initial subset of ISs according to some criteria, assigns them *trust* or *credibility coefficients* (Gligor and Wing 2011) and activates the procedure that transforms selected IS to autonomous agents with capabilities similar to those of the user. The procedure runs recursively from the initial IS, so that second-stage ISs are selected and activated. This allows the agents to pursue the search autonomously and simultaneously, until a prescribed stack level or the desired retrieval goal is achieved. A creativity-stimulating content-based search and recommendation has been investigated within the recent Horizon 2020 project (Skulimowski 2017a). The design of GES knowledge provision procedures must ensure that the reply to each query is given at a specified level of trust. When trust coefficients φ_i , $0 \leq \varphi_i \leq 1$, are assigned to each source of information available to this GES, the resulting trust $\tau(q)$ in the information retrieved in reply to a query q can be higher than any of its individual sources.

Autonomous management of complex queries processed by a GES is a multicriteria combinatorial optimization problem (Skulimowski 1994). The order of queries from different users and the sequence of information sources to be contacted can be assessed from the point of view of precision, recall, and other information retrieval measures, such as timeliness. The GES functioning proposed in Skulimowski (2013) is based on a snowball principle: The node that generated a query activates other units until the desired information is found. The following principles of query processing in a GES have been defined in Skulimowski (2013).

- (a) Each knowledge unit K activated by another one, K_i with a query q_{ij} returns the information specified by q_{ij} to K_i or passes to (b).
- (b) If the query q_{ij} can be only partly responded by K_j , the latter unit modifies it to q_{jk} to ask for the missing information. Thus K_j activates further knowledge

⁵<https://www.worldwidewebsite.com/> [access Jan 10, 2020].

units $K_{k1}, \dots, K_{kn(k)}$ with the query q_{jk} in the order specified as a solution of the search optimization problem as proposed in Skulimowski (1994). The resulting information search strategy minimizes the number of repeated activations of the same knowledge unit.

- (c) The procedure (b) activates recursively further units. Each unit K_j activated by K_i fuses the information received from units activated by itself and returns them to K_i . All activated units are deactivated after the information requested in q_{ij} is gathered.

As previously mentioned, the above procedure is a special case of a multicriteria search strategy optimization problem, where the resulting strategy maximizes the amount of information, which is to be gathered in the least amount of time, at a minimum effort of all activated units, and at minimum cost for the initial unit. Such a search strategy may be endowed with a certain level of free will and may be designed to fulfill the definition of a creative decision process (cf. Skulimowski 2011).

The natural question of whether science is capable of accommodating any kind of future AI technology for research purposes and how it can be achieved appears when projecting the GES future. From a purely economic standpoint, the role of AADSs in e-science will grow, encompassing new areas of intellectual activity and the replacement of human researchers. Performing a complex Web search strategy by an intelligent autonomous web crawler is a real-life example of such empowerment. The development of GESs will challenge users with a growing complexity of queries, a growing amount of gathered information, and with a need to comprehend the search workflow. Rejecting useful information due to the lack of an appropriate explanation of its provenance (Malaverri et al. 2013) may cause the recipients to lose the reply, but they may prefer to proceed so as to avoid infringing cybersecurity rules.

3 Results of the Delphi Survey on e-Science Tools and Factors

This section highlights a sample of the Delphi survey results (Skulimowski et al. 2013). This survey based on the novel “Extrapolation Delphi” principle was performed twice, the first time within the above-cited project and once during its durability period. Specifically, we present the results concerning the future development of advanced expert systems, heading toward advanced GESs, which were the subject of questions contained in survey Section 11 titled “*Future prospects of knowledge base, expert systems, information streams and decision support systems integration*” (Skulimowski et al. 2013). The replies to five questions most relevant to this article’s topics are presented out of 36 questions in the above mentioned survey section.

Table 1 Estimated share φ_I [in %] of researchers considering the online information widely available through browsers and search engines as fully representative in their areas of scientific research. Analysis of the replies to question No. 11.1a in (Skulimowski et al. 2013) weighted with combined trust/competence coefficients of respondents

Specification of results	Estimates for 2015	Forecasts for 2020	Forecasts for 2025	Forecasts for 2030
No. of replies	47	47	47	47
Shapiro–Wilk test	negative	negative	positive	negative
Unimodality test	positive	positive	positive	positive
Mean weighted value	25,519	28,73	40,146	47,337
Weighted standard deviation	18,006	16,737	21,449	22,66
Weighted left semideviation	15,27	16,906	18,867	22,7
Weighted right semideviation	22,93	17,18	25,806	23,373
Weighted median value:	20	25	35	45
1st weighted quintile	5	10	15	20
2nd weighted quintile	10	20	30	40
3rd weighted quintile	20	30	35	50
4th weighted quintile	50	50	50	60
Interquintile range IQVR	45	40	35	40
Interquartile range IQR	30	30	30	25
No. of reply clusters	1	1	1	1

3.1 Delphi Survey Background and Scope

The survey results are presented in tables, which provide the basic statistical characteristics of replies, together with Delphi-specific consensus measures of experts and a cluster analysis (von der Gracht 2012). The latter is then used to construct the development scenarios of investigated information systems. The survey respondents were requested to define certain numerical development indicators for four time horizons: 2015 (as forecast in 2013 and an estimate in 2016), 2020, 2025, and 2030 (forecasts). The following indicators have been calculated for all replies and for all time horizons:

- the average value, standard deviation, left and right semideviations,
- the median, 1st and 3rd quartile and four quintiles,
- the *interquartile range* (IQR), defined as the difference between the third and first quartile,
- the *interquintile range* (IQVR), defined as the difference between the fourth and first quintile,
- Hartigans' dip test of unimodality (Hartigan and Hartigan 1985); if negative, it was followed by a clustering of replies and the number of clusters of replies was determined,
- the Shapiro–Wilk (log) normality test, applied to replies either directly or to their logarithms when the question touched upon growth ratios.

The consensus indicators IQR and IQVR should be normalized, for example by dividing them by the maximum data range $R := r_{\max} - r_{\min}$ after eliminating the outliers. Then, the consensus is defined by one or both inequalities

$$\text{IQR}/R \leq \eta_1, \quad \text{IQVR}/R \leq \eta_2,$$

where η_k , $k = 1, 2$, are certain threshold values and $\eta_1 \leq \eta_2$. We can clearly see that given the same threshold value, the IQVR provides a stronger consensus test. A positive result of the Shapiro–Wilk normality test indicates a potentially unimodal distribution of replies and rejects the hypothesis that there is more than one cluster of replies.

The statistical analysis was first performed under the hypothesis that the replies be weighted according to a self-assessment of certainty by the respondents' survey, in combination with a self-assessed credibility coefficient of individual replies, and an automatically assigned individual expert competence score. This score was computed by the Delphi support system⁶ (Skulimowski 2017), based on previous survey participation, the record of publications, research projects, and other achievements in the question-related area. It has been observed (Skulimowski 2016a) that for most survey questions, there was no significant difference between the statistical indicators for weighted and non-weighted responses. This observation also touches upon the consensus measures and indicates that the expert group's ability to estimate the future evolution of indicator values was homogeneous. Therefore, in this section we concluded that the resultant analysis variant yields a smaller statistical error (in terms of the standard deviation) for a majority of forecasting horizons. The sum of errors was a decisive factor, for an equal number of dominating values at different horizons. Out of five questions selected for this section, only the replies to question 11.8 (Table 4) exhibited smaller errors when analyzed without weighting coefficients.

The survey in the project SCETIST (Skulimowski 2013) consisted of two rounds and was conducted in 2012 and 2013. There was also a post-project update round with the same participants, questions and Delphi support software. The respondents could select the questions to answer, according to their competences. Therefore, from over

⁶The current version of the system is available at www.forgnosis.eu.

100 respondents, the number of those replying to questions in Section 11.1 varied between 43 and 48 in the first and second rounds.

3.2 The Future Use of Information Systems for e-Science—The Results of the Delphi Survey

The first of the above-mentioned survey outcomes presented in this paper is a basic statistical analysis of question 11.1a pointing out the forecasted shares of scientists that consider online information to be accurately representative of their research. It is shown in Table 1.

The above question did not distinguish between the research areas, so the replies only provide a rough estimate by merging humanities, engineering, etc. However, it shows the average value of online researchers' share almost doubling between 2015 and 2030, while the mean square ex-ante forecast error rose only by about 20%, and the relative error decreased considerably. All but one (2025) reply sets for the estimation (2015) or forecasting (2020, 2025, 2030) horizons were considerably irregular and did not pass the weighted Shapiro–Wilk normality test. However, all value distributions were unimodal and concentrated in one cluster.

Let us note that all quantiles (quartiles, quintiles, median) and consequently, the consensus measures, are integers because the respondents select their replies from the standard integer pick list [0:100]. The same list was used for all questions in Section 11 of the survey where the replies were to be provided in %.

Table 2 shows the breakdown of the verified and raw quantitative information available on the Web for the same estimation/forecasting.

The respondents estimated the amount of trustworthy information (i.e., *knowledge*) to comprise about one-fifth of all quantitative information available. This cannot be seen as an optimistic estimate. The forecast for 2030—about 40% of refined information—presumes the emergence of a new data refinery mechanism. This share is almost double in comparison with the estimate for the present state of the Internet. Nevertheless, the share of unverified Web information will still be close to the larger part of the golden proportion, which is an indication of the power of disinformation and fake data. The question in the first two rounds just touched upon the knowledge, irrespective of whether it was quantifiable or not. Based upon the respondents' postulates, the question for the follow-up round was formulated more precisely, but without a statistically essential impact on outcomes. A characteristic feature of the above replies is smaller than the usual difference between the IQR and IQVR consensus measures, which indicates a relatively large number of equal replies between the 1st quartile and 1st quintile as well as between the 3rd quartile and 4th quintile.

The next question (11.3) assumed the emergence of a next generation of Wolfram's Alpha⁷—an expert system capable of providing informed replies to virtually any

⁷<http://www.wolframalpha.com>.

Table 2 Amount of processed and verified quantitative knowledge available online (in % of all quantitative information available). Replies to question no. 11.2 weighted with combined trust/competence coefficients

Specification of results	Estimates for 2015	Forecasts for 2020	Forecasts for 2025	Forecasts for 2030
No. of replies	47	47	47	47
Shapiro–Wilk test	negative	negative	negative	positive
Unimodality test	positive	positive	positive	positive
Mean weighted value	17,186	22,732	27,664	38,079
Weighted standard deviation	11,66	12,865	16,657	24,339
Weighted left semideviation	9,363	12,572	18,072	21,447
Weighted right semideviation	15,135	13,633	16,008	28,604
Weighted median value:	15	20	30	35
1st weighted quintile	5	10	10	15
2nd weighted quintile	10	15	20	25
3rd weighted quintile	15	25	30	40
4th weighted quintile	25	30	40	50
Interquartile range IQVR	20	20	30	35
Interquartile range IQR	15	20	30	32
No. of reply clusters	1	1	1	1

query. This question touched upon a quantitative characteristic of a future GES capability to reach the existing information, namely, its maximum recall value relative to the query provided by the system user. Replies equal to “0” were representative of the disbelief of this particular survey respondent that such software will be created (Table 3).

Unlike in the case of the two previous questions, the replies to question 11.3 above indicate a sharp rise in the GES search range, from an initial estimate of about 2–27% in 2030, with a high yet relatively decreasing uncertainty, expressed by standard deviation and semideviations.

A symmetrical problem to that shown above was considered in question 11.8 (Skulimowski 2013); namely, we investigated the Internet users’ attitudes to

Table 3 The share of information available on the Web that can be processed by advanced expert software (GES) capable of analyzing heterogeneous data (quantitative economic information, multimedia, publications, video streaming) and providing GES users with informed replies to any given question (in % of available information used for this purpose)

Specification of results	Estimates for 2015	Forecasts for 2020	Forecasts for 2025	Forecasts for 2030
No. of replies	48	47	47	47
Shapiro–Wilk test	positive	positive	positive	negative
Unimodality test	positive	positive	positive	positive
Mean weighted value	2,156	5,525	14,536	26,926
Weighted standard deviation	4,928	6,877	14,319	20,797
Weighted left semideviation	2,091	4,477	10,295	15,467
Weighted right semideviation	11,713	10,861	18,675	30,245
Weighted median value:	0	2	10	20
1st weighted quintile	0	0	2	10
2nd weighted quintile	0	0	5	15
3rd weighted quintile	0	3	15	25
4th weighted quintile	1	10	20	50
Interquintile range IVQR	1	10	18	40
Interquartile range IQR	0	10	10	40
No. of reply clusters	1	1	1	1

searching for solutions to their problems on the Web. The analysis of replies is given in Table 4.

A predominance of solving problems through access to online information is not a surprise. Actually, the above characteristics may be burdened by a relatively high share of elderly people who have Internet access via their mobile phones, but use it sparingly. The most recent research performed within the project (Skulimowski 2019) yields considerably higher estimates for 2025 and 2030, reaching more than 90% of all queries.

The last set of results presented in this section touches upon the emergence of qualitatively new capabilities and phenomena in GESs, manifesting itself through

Table 4 Answers to problems, questions, and queries of all kinds (translations, spelling, definitions, geographical information, graphical object finding, legislation, etc.) that will be sought online: in % of all queries from user with Internet access (mobile or landline); unweighted

Specification of results	Estimates for 2015	Forecasts for 2020	Forecasts for 2025	Forecasts for 2030
No. of replies	46	45	45	45
Shapiro–Wilk test	negative	negative	negative	negative
Unimodality test	positive	positive	positive	positive
Mean value	38,553	49,189	60,676	71,919
Standard deviation	19,134	18,428	18,64	19,836
Median value:	40	55	65	75
1st quintile	15	25	30	40
2nd quintile	25	40	55	70
3rd quintile	40	55	65	75
4th quintile	50	60	70	85
Interquintile range	35	35	40	45
Interquartile range	32,5	28,75	27,5	20
No. of reply clusters	1	1	1	1

solving previously intractable problems or answering unresolved questions. Namely, the integration of knowledge on the Internet will allow for a new level of quality in resolving problems presented by GES users, specifically those intractable problems, and providing replies to queries, which are unavailable through contemporary information processing methods (Table 5).

Both the uncertainty expressed by the standard deviation and semi-deviations, as well as the consensus indicators IQR and IQVR for question 11.9, are relatively lower than in case of the two previous forecasts. Fitting the above replies with the logistic curve (Skulimowski 2017b), we can calculate the expected time when the majority of problems and queries can be better solved by GESs, namely the year 2037. This year can thus be regarded as a kind of a *singularity* (Skulimowski 2014b); however, in a limited sense. To conclude this section, let us note that reaching a consensus need not be the ultimate goal of a Delphi survey. Usually, if the unimodality test is negative, a lack of consensus indicates the existence of several clusters of replies. If this is not the case and the IQR or IQVR values are rather high, while growing more slowly than the trend investigated by the survey, it means that there is a common expectation of a certain trend or event among the survey respondents, with a high uncertainty regarding its time of occurrence, however.

Table 5 The share in % of problems and queries that will be more adequately solved by GES, compared to the solutions and replies provided by human experts (question 11.9)

Specification of results	Estimates for 2015	Forecasts for 2020	Forecasts for 2025	Forecasts for 2030
No. of replies	46	45	45	45
Shapiro–Wilk test	negative	negative	positive	positive
Unimodality test	positive	positive	positive	positive
Mean weighted value	14,477	23,12	31,653	45,813
Weighted standard deviation	9,439	9,786	11,459	18,268
Weighted left semideviation	7	7,436	8,881	16,184
Weighted right semideviation	12,231	12,634	15,542	21,121
Weighted median value:	10	20	30	40
1st weighted quintile	5	15	20	25
2nd weighted quintile	10	15	30	35
3rd weighted quintile	10	20	30	40
4th weighted quintile	20	30	40	60
Interquintile range IQVR	15	15	20	35
Interquartile range IQR	15	10	20	25
No. of reply clusters	1	1	1	1

4 Discussion and Conclusions

The results of the Delphi survey presented in Sect. 3 provide clues, arising from expert judgments, regarding the amount of information available online and its use for e-science purposes until 2030. It is expected that by 2030, the corresponding information retrieval tools will reach sufficient enough levels to provide virtually all necessary scholarly information to researchers. Furthermore, within a similar time frame, GESs are expected to outperform human experts in solving complex knowledge processing tasks.

Another AI trend that may have a relevant impact on e-science is the development of brain–computer interfaces (BCIs) and their deployment in enhancing research, their joint use with GESs and AILPs, as well as in intelligent decision support

systems. The results of a Delphi survey on BCIs are presented in Skulimowski (2014b, 2016b). Here, we briefly discuss a summary of these findings. By definition, in a BCI, outward information is retrieved by recognizing the brain's electromagnetic neural activity, while for the inward transfer direction, a BCI triggers the neural circuits directly (Brunner et al. 2011; Jiang et al. 2019). The best transmission rates and qualities were obtained with invasive BCIs, based on intracranial implants, but the greatest hope in enhancing human capabilities is placed on non-invasive BCIs, such as wearable devices that are used to retrieve EEG or fMRI signals. They are expected to facilitate efficient bidirectional communication with GES (Zhang et al. 2013) as well as direct communication between human brains, called hyperinteraction (Grau et al. 2014; Jiang et al. 2019). The ability of a BCI to directly connect researchers' brains with powerful expert systems will speed up progress in global data integration provided by GESs. It will also increase the efficiency of scientific collaboration (Leidig and Fox 2014; Shi et al. 2017) and the use of AILPs. The positive effect of BCIs on researchers who obtain efficient and instant access to big research data may partly compensate for the negative impact of data explosion. However, the question of whether e-science can fully exploit the capabilities of emerging advanced AI tools and technologies such as AILPs, GESs, and BCIs to increase the quality and efficiency of scientific research remains to be seen.

The analysis of the full set of SCETIST Delphi survey replies resulted in deriving three human–AI interaction scenarios (cf. Skulimowski 2014b, 2016a, b). Here, we adjust them slightly to provide conditional responses to the above question. The full and beneficial use of AI defines the optimistic scenario of human–AADS interaction, while the negative response is associated with the pessimistic scenario, often referred to as the *AI threat problem*. The foresight results presented in Skulimowski (2014b) suggest that the main condition triggered between the positive and negative scenarios is the capability of future BCIs to provide a direct interface to GESs and facilitate the creative process of GES users.

In the optimistic scenario, the growing empowerment of AADSs will be compensated for by the ability of human supervisors and authorized users to control them directly with BCIs. This scenario is backed by results of the Delphi survey presented in Sect. 3, which suggest that GESs and AILPs supported by high-performance BCIs and enhanced reality will ensure control over advanced AI technologies. Further results of the Delphi survey on the development of artificial creativity and creativity support systems performed in SCETIST (Skulimowski 2016a) highlight the importance of coupling human users with GESs and AILPs via BCIs to stimulate their creative abilities.

The pessimistic scenario presumes that a growing share of human creative activity, specifically in research, will be replaced by AADSs due to the ever-growing complexity of research and decision problems to be solved along with increasingly large data volumes. In this scenario, AADSs will specify goals, criteria and constraints, target quality and the scope of applicability of solutions. Human researchers will only perform auxiliary and assistive roles.

In the third, neutral scenario, technological development is generally slowed down in the face of various setbacks. In this case, the AADS/human competition problem

will be deferred to a more distant future, beyond horizon 2030 of the foresight studies presented here.

In conclusion, the results of recent foresight studies highlight the relevance of development trends in selected advanced AI technologies for future e-science, e-learning, and e-research. According to the outcomes of the research projects (Skulimowski et al. 2013; Köhler and Skulimowski 2019), the areas of intensive ICT/AI development efforts that can be of utmost relevance for e-science are GESs driven by autonomous web crawlers and dedicated decision support systems, creativity support systems capable of stimulating or at least preserving human creative abilities, and bidirectional non-invasive BCIs providing direct links to GESs and other researchers to efficiently tackle large amounts of scientific data.

Acknowledgements The author is grateful for the support of the research project “*Scenarios and Development Trends of Selected Information Society Technologies until 2025*” (acronym “SCETIST”), financed from the ERDF funds within the Innovative Economy Operational Programme, contract No. WND-POIG.01.01.01-00-021/09. Some results presented in this paper have been also verified and extended within the Horizon 2020 research project “*Training towards a society of data-savvy information professionals to enable open leadership Innovation*” (acronym “MOVING”) financed by the EC, contract No. 693092. Finally, the author is grateful for the invitation to the Organizing Committee of the e-Science Conference in Leipzig.

References

- Brunner, P., Bianchi, L., Guger, C., Cincotti, F., Schalk, G.: Current trends in hardware and software for brain–computer interfaces (BCIs). *J. Neural Eng.* **8**(025001), 7 (2011)
- Elliott, G., Timmermann, A.: Optimal forecast combinations under general loss functions and forecast error distributions. *J. Econom.* **122**, 47–79 (2004)
- Gligor, V., Wing, J.M.: Towards a theory of trust in networks of humans and computers. *Lecture Notes in Computer Science*, 7114, Security Protocols **XIX**, 223–242 (2011)
- Grau, C., Ginhoux, R., Riera, A., Nguyen, T.L., Chauvat, H., Berg, M., Amengual, J.L., Pascual-Leone, A., Ruffini, G.: Conscious brain-to-brain communication in humans using non-invasive technologies. *PLoS ONE* **9**(8), e105225 (2014)
- Heylighen, F.: The offer network protocol: mathematical foundations and a roadmap for the development of a global brain. *Eur. Phys. J. Spec. Top.* **226**, 283–312 (2017). <https://doi.org/10.1140/epjst/e2016-60241-5>
- Hartigan, J., Hartigan, P.: The dip test of unimodality. *Ann. Stat.* **13**(1), 70–84 (1985)
- Jeffery, K., Houssos, N., Jörg, B., Asserson, A.: Research information management: the CERIF approach. *Int. J. Metadata, Semant. Ontol.* **9**(1), 5–14 (2014). <https://doi.org/10.1504/ijmoso.2014.059142>
- Jiang, L., Stocco, A., Losey, D.M., Abernethy, J.A., Prat, ChS, Rao, R.P.N.: BrainNet: a multi-person brain-to-brain interface for direct collaboration between brains. *Sci. Rep.* **9**, 6115 (2019). <https://doi.org/10.1038/s41598-019-41895-7>
- Köhler, T., Skulimowski, A.M.J.: Triggering research creativity and increasing efficiency through open innovation platforms. In: International Congress on Advanced Applied Informatics (IIAI AAI 2019)—10th International Conference on E-Service and Knowledge Management (ESKM 2019), Toyama, Japan, July 7–12, 2019, pp. 49–55 (2019). <https://doi.org/10.1109/IIAI-AAI.2019.00021>

- Leidig, J.P., Fox, E.A.: Intelligent digital libraries and tailored services. *J. Intell. Inf. Syst.* **43**, 463–480 (2014). <https://doi.org/10.1007/s10844-014-0342-3>
- Malaverri, J.E.G., Mota, M.S., Medeiros, C.B.: Estimating the quality of data using provenance: a case study in eScience. 19th Americas Conference on Information Systems, AMCIS 2013—Hyperconnected World: Anything, Anywhere, Anytime, v.2, pp. 1442–1451 (2013). <http://repositorio.unicamp.br/handle/REPOSIP/88829>
- Mohamed, B., Pscheidt, D., Köhler, T.: The Fish model: a conceptual framework for enhancing e-research collaboration. In: 5th International Conference on Education and New Learning Technologies, July 1–3, 2013, Barcelona, Spain—EDULEARN13 Proceedings, IATED, pp. 3267–3276 (2013)
- Shi Z., Fang, F.: Optimizing peer teaching to enhance team performance. In: Sukthankar G., Rodriguez-Aguilar, J. (eds.) Autonomous Agents and Multiagent Systems, pp. 138–150. AAMAS 2017. Lecture Notes in Computer Science, 10642. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-71682-4_9
- Skulimowski, A.M.J.: Optimal strategies for quantitative data retrieval in distributed database systems. Proceedings of the Second International Conference on Intelligent Systems Engineering, Hamburg, pp. 389–394, September 5–9, 1994; IEE Conference Publication No. 395. IEE, London (1994). <https://doi.org/10.1049/cp:19940655>
- Skulimowski, A.M.J.: Freedom of choice and creativity in multicriteria decision making. In: Theeramunkong, T., Kunifiji, S., Nattee, C., Sornlertlamvanich, V. (eds.) Knowledge, Information, and Creativity Support Systems: KICSS 2010 Revised Selected Papers, LNAI 6746, pp. 190–203. Springer, Berlin; Heidelberg (2011). https://doi.org/10.1007/978-3-642-24788-0_18
- Skulimowski, A.M.J.: Universal intelligence, creativity, and trust in emerging global expert systems. In: Rutkowski L. et al. (eds.) 12th International Conference on Artificial Intelligence and Soft Computing, Zakopane, 2013, Proceedings, Part II. LNAI 7895, pp. 582–592. Springer-Verlag, Berlin–Heidelberg (2013). https://doi.org/10.1007/978-3-642-38610-7_53
- Skulimowski, A.M.J.: Anticipatory network models of multicriteria decision-making processes. *Int. J. Syst. Sci.* **45**(1), 39–59 (2014a). <https://doi.org/10.1080/00207721.2012.670308>
- Skulimowski, A.M.J.: Future prospects of human interaction with artificial autonomous systems. In: Bouchachia, A. (ed.) Adaptive and Intelligent Systems, Third International Conference, ICAIS 2014, Bournemouth, UK, Proceedings. LNAI 8779, pp. 131–141. Springer-Verlag, Berlin–Heidelberg (2014b). https://doi.org/10.1007/978-3-319-11298-5_14
- Skulimowski, A.M.J.: The role of creativity in the development of future intelligent decision technologies. In: Skulimowski, A.M.J., Kacprzyk, J. (eds.) Knowledge, Information and Creativity Support Systems: Recent Trends, Advances and Solutions. Selected Papers from KICSS '2013. Advances in Intelligent Systems and Computing 364. Springer (2016a). https://doi.org/10.1007/978-3-319-19090-7_22
- Skulimowski, A.M.J.: Impact of future intelligent information technologies on the methodology of scientific research. In: 16th IEEE International Conference on Computer and Information Technology, Nadi, Fiji, Dec 7–10, 2016, Proceedings, IEEE CPS, pp. 238–247 (2016b). <https://doi.org/10.1109/CIT.2016.118>
- Skulimowski, A.M.J.: Cognitive content recommendation in digital knowledge repositories—a survey of recent trends. In: Rutkowski, L., Korytkowski, M., Scherer, R., Tadeusiewicz, R., Zadeh, L., Zurada, J. (eds.) Artificial Intelligence and Soft Computing. ICAISC 2017. Lecture Notes in Computer Science 10246, pp. 574–578. Springer, Cham (2017a). https://doi.org/10.1007/978-3-319-59060-8_52
- Skulimowski, A.M.J.: Expert Delphi survey as a cloud-based decision support service. IEEE 10th SOCA 2017, Nov 22–25, 2017, Kanazawa, Japan. IEEE, Piscataway, pp. 190–197 (2017b). <http://ieeexplore.ieee.org/document/8241542>
- Skulimowski, A.M.J.: Forward-looking activities supporting technological planning of AI-based learning platforms. In: Herzog, M., Kubincová, Z., Han, P., Temperini, M. (eds.) Advances in Web-Based Learning—ICWL 2019, Magdeburg, Sept 23–25, 2019. Lecture Notes in Computer

- Science 11841, pp 274–284. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-35758-0_26
- Skulimowski, A.M.J., et al. (ed.): Scenarios and development trends of selected information society technologies until 2025. Final Report, Progress & Business Publishers, Kraków, Progress & Business Foundation. Available at <http://www.ict.foresight.pl> (2013)
- Thakkar, P., Nathan, S., Viswanathan, B.: Performance benchmarking and optimizing hyperledger fabric blockchain platform. In: 2018 IEEE 26th International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS), pp. 264–276. Milwaukee, WI (2018). <https://doi.org/10.1109/MASCOTS.2018.00034>
- von der Gracht, H.A.: Consensus measurement in Delphi studies: review and implications for future quality assurance. *Technol. Forecast. Soc. Chang.* **79**(8), 1525–1536 (2012)
- Zhang, C., Wang, H., Wang, H., Wu, M.-H.: EEG-based expert system using complexity measures and probability density function control in alpha sub-band. *Integr. Comput.-Aided Eng.* **20**, 391–405 (2013)
- Xu, F., Uszkoreit, H., Du, Y., Fan, W., Zhao, D., Zhu, J.: Explainable AI: a brief survey on history, research areas, approaches and challenges. In: Tang, J., Kan, M.Y., Zhao, D., Li, S., Zan, H. (eds.) Natural Language Processing and Chinese Computing. NLPCC 2019. Lecture Notes in Computer Science 11839, pp. 563–574. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32236-6_51

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

