DEVELOPMENT AND EVALUATION OF AN [CUSTOMIZABLE MOBILE] APPLICATION FOR BEHAVIORAL RESEARCH IN DATA ANALYTICS

Master Thesis
of

MAX DARMSTADT

June 28, 2023

Matriculation Number
1820000

Submitted at the Chair of Enterprise Systems
University of Mannheim

Reviewer: Prof. Dr. Hartmut Höhle
Supervisor: Jan Schilpp

Contents

List of Figures

List of Tables

List of Abbreviations

IEEE	Institute of Electrical and Electronics Engineers
ISO	International Organization for Standardization

Abstract

Over the past few years, data analytics has become increasingly important for companies across all industries. With the massive amount of data that is now available, companies can use data analytics to gain valuable insights into consumer behavior, market trends, and internal operations, among other things. As a result, data analytics has become a critical tool for companies looking to gain a competitive edge in today's rapidly evolving business environment. However, while data analytics has become an essential tool for businesses, there has been relatively little research done in the area of behavioral research. Specifically, there is a lack of research on the decision-making process involved in data analytics, and how individuals and organizations use data analytics to inform their decisions. One of the major challenges in conducting research in this area is the high cost of developing custom applications for each study. The development of such applications can be time-consuming, expensive, and often requires specialized expertise. To address this challenge, this master thesis develops a generic application that streamlines the process of conducting studies in the field of data analytics. This application enables researchers to design, conduct, and analyze studies more efficiently and cost-effectively, allowing them to explore the field in greater depth. This will be accomplished by using the design science research approach. Firstly, the problem of a lack of behavioral research in data analytics is identified. Then, the objectives for a solution are defined through a literature review and the use of requirement engineering to gather requirements for the application. Next, the application is design, implemented prototypically and its functionality demonstrated. Finally, solution is evaluated through the usages of the requirements.

1 Introduction

1.1 Background and Motivation

1.2 Research Problem and Objectives

1.3 Contribution and Scope of the Study

2 Theoretical foundations

This section represents the theoretical fundamentals of this elaboration by defining the terms “data analytic”, “information value chain”, and “boundaries and conflicts” as they are used in the context of this literature review.

2.1 Definition of terms

2.2 Data Analytics

The term “data analytics” originated in the early 2000s and describes an interdisciplinary field that combines areas such as statistics, machine learning, pattern recognition, system theory, operations research and artificial intelligence (Runkler, 2020). It can be generally defined “[...] as the application of computer systems to the analysis of large data sets for the support of decisions.” (Runkler, 2020). This definition showcases the broadness of the topic, as most computer systems process some amount of data and theoretically allow for some kind of decision making. Due to this broad definition, data analytics can cover slightly different subject areas depending on the context it is discussed in. In this elaboration, data analytics refers to the processing of large amounts of data, also referred to as “big data”, through mathematical procedures or machine learning methods with the goal of creating new knowledge. In summary, processes that merely prepare or show data are not considered data analytics, but only processes that process data in such a way that new knowledge can be derived from it. This distinction is made to differentiate data analytics from traditional data processing areas like business intelligence. The goal of data analytics, as is discussed in this literature review, is to retrieve some kind of previously unknown knowledge from a set of data. This process can be generally described using the “information value chain” model. In their research, Abbasi et al. analyze this model in the context of big data in an effort to create an inclusive research agenda for big data in information system research (Abbasi et al., 2016).

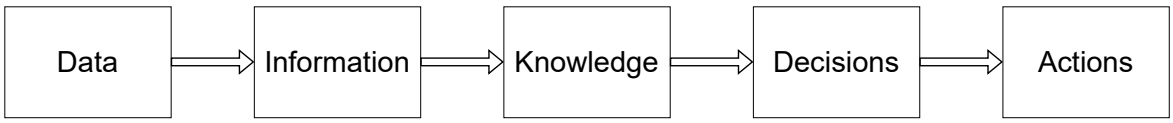

Figure 1: Information Value Chain

2.3 Information Value Chain

The information value chain (figure 1) is a set of phases that define the transformation of raw data to information and eventually into knowledge. "Data" describes raw facts without any structuring. Once organized, the processed data represents "information". This "information" is then used to find patterns and draw conclusions. At this time, the information becomes knowledge (Fayyad et al., 1996a), Fayyad et al., 1996b. This knowledge is then used to make "decisions" and take corresponding "actions" (Sharma et al., 2014). Each phase of the information value chain also includes a different set of technologies and methodologies. For example, the "data" phase contains technologies and actions regarding the basic storage of data like database systems or data warehouses (Abbasi et al., 2016). The conventional version of this information value chain represents an approach that generally explains the processing of data. The main steps of this information value chain are also applicable for big data (Abbasi et al., 2016). This general structure of processing data is also supported by literature from the data analytics field (Runkler, 2020). In addition, the information value chain contains the further phases "decisions" and "actions", which deal with the influence of the processed data. These phases reflect the impact of data analytics, since data analytics is primarily a technology for the decision-making process (Runkler, 2020). For this reason, the information value chain is a suitable model to structure different phases in the processing of data in the context of data analytics.

2.4 Boundaries and Conflicts in Organizations

This literature review uses the terms boundary and conflict interchangeably. In order to include as much literature as possible, the criteria for boundaries are kept very general. Prior to conducting the literature review, there was no formal definition of boundaries in the context of data analytics used for the selection of literature. Generally, boundaries are described as

"[...] a real or imagined line that marks the limits or edges of something and separates it from other things or places [...]" (Hornby, 2015). Based on this general description, the term boundary is defined in the context of this elaboration as any circumstance that leads to a reduction in the effectiveness or efficiency of an organization. Boundaries and conflicts are therefore used to describe any circumstance that hinders an organization from being perfectly productive. An example of such boundaries or conflicts would be communication issues between different departments, which lead to a reduction of productivity.

2.5 Design Science Research Methodology

2.6 Requirement Engineering

Die Methodik, um die Anforderungen der Endanwender in dieser Arbeit zu bestimmen, orientiert sich an dem sogenannten "Requirement Engineering" Institute of Electrical and Electronics Engineers (IEEE) Standard für die Analyse und Evaluation von Anforderungsspezifikationen. (Vgl. Alain Abran, James W. Moore, 2004, S.2) Der Begriff des "Requirement Engineering", frei übersetzt mit "Anforderungsentwicklung", ist dabei ein englischer Begriff aus der Systemanalyse und wird zur Analyse und Evaluierung von Endanwenderanforderungen genutzt (Vgl. Sommerville, 2011, S.82-111) Der IEEE Standard 610.12-1990 definiert eine Anforderung (englisch "Requirement") dabei folgendermaßen:

> ***"(1) Eine Bedingung oder Fähigkeit, die ein Benutzer benötigt, um ein Problem zu lösen oder ein Ziel zu erreichen. (2) Eine Bedingung oder Fähigkeit, die ein System oder eine Systemkomponente erfüllen oder besitzen muss, um einen Vertrag, eine Norm, eine Spezifikation oder andere formell auferlegte Dokumente zu erfüllen. (3) Eine Dokumentendarstellung einer Bedingung oder Fähigkeit wie in 1 oder 2."*** (IEEE, 1990, S.62)

Die hier verwendete Vorgehensweise orientiert sich allerdings nur in Teilen an diesem Standard, da das "Requirement Engineering" grundsätzlich ein Teil des Systementwicklungsprozesses ist und sich diese Arbeit nur mit dem Analysieren und Evaluieren von Standardinhaltsanforderungen in einem existierenden System befasst. Deshalb orientiert sich die Definition

im Kontext dieser Arbeit an Punkt (1) des IEEE Standard 610.12-1990 mit dem Ziel, dem Benutzer eine verbesserte Usability zu gewährleisten. Dabei kann das "Requirement Engineering" in die vier Teilbereiche der Anforderungserhebung, Anforderungsanalyse, Anforderungsspezifikation und Anforderungsbewertung aufgeteilt werden. Bei der Anforderungserhebung werden zunächst die Anforderungen an ein neues Softwaresystem definiert, diese dann analysiert und ausgearbeitet, um sie dann zu dokumentieren und zu evaluieren. (Vgl. Sommerville, 2011, S.82-111) Die Anforderungserhebung der Endanwender an das System wird in dieser Arbeit mithilfe einer Umfrage durchgeführt. Diese Umfrage wird anonymisiert mit Teilnehmern des SAP Enable Now Round Tables stattfinden. In der Anforderungsanalyse wird die momentane Situation der SAP Marketing Cloud betrachtet und mithilfe der Umfrageergebnisse verglichen, um die Anforderungen der Endanwender abzuleiten. Diese Anforderungen werden dann in einem Proof-of-Concept beispielhaft umgesetzt und dazu verwendet, die Anforderungsbewertung durchzuführen. In der Anforderungsbewertung wird dabei der International Organization for Standardization (ISO) 9241 Standard(ISO Org., 2019) genutzt, um die Verbesserung der Usability zu verdeutlichen. Das Ziel der Umfrage und dieser Arbeit ist es also herauszufinden, wie der SAP Enable Now Webassistent von den Endanwendern genutzt wird und wie der dafür bereitgestellte Standardinhalt in der SAP Marketing Cloud verbessert werden kann.

3 Identification of the Problem

3.1 Previous Studies and Gaps in the Literature

3.1.1 Literature Review: Behavioral Research in Data Analytics (on the basis of the Information Value Chain)

3.2 Applications (Anwendungen) for Behavioral Research

4 Definition of Objectives for a solution

4.1 Literature Review Studies in Data Analytics and General

4.2 requirements elicitation

4.2.1 Functional and non-functional requirements

Damit eine Anwendungsmodernisierung durchgeführt und eine modernisierte Architektur konzeptioniert werden kann, ist die Bestimmung der Anforderungen wichtig.(Vgl. Seacord et al., 2003, Kapitel 3) Anforderungen können nach ISO/IEC 25000, beziehungsweise dem Qualitätsmodell aus ISO/IEC 25010, als Qualitätskriterien an Software und Systeme klassifiziert werden.(Vgl. ISO/IEC 25010, 2011) Das IEEE definiert Anforderungen als eine Bedingung oder Eigenschaft, welche von einem System oder einer Systemkomponente erfüllt werden muss, um eine Problemstellung oder Zielsetzung eines Nutzers oder formalen Dokuments zu erfüllen.(Vgl. IEEE, 1990, S.62) Anforderungen können nach diesen beiden Definitionen als zu erfüllende Eigenschaften oder Qualitätskriterien einer Software oder eines Systems definiert werden. Aus diesem Grund werden die Anforderungen an den SAP Lean Catalog aus den von der SAP beschriebenen Produkteigenschaften und Merkmalen des SAP Lean Catalogs abgeleitet. Zusätzlich werden allgemeine Anforderungen an eine Softwaremodernisierung berücksichtigt, welche erfüllt sein müssen, um eine erfolgreiche Migration durchzuführen. Diese Anforderungen an den SAP Lean Catalog und an eine Softwaremigration im Allgemeinen werden weiter in funktionale und nichtfunktionale Anforderungen eingeteilt.

Eine funktionale Anforderung beschreibt eine Funktion oder Fähigkeit eines Systemes, die konkret von einem System oder einer Softwarekomponente durchgeführt werden können muss.(Vgl. IEEE, 1990, S.35) Ein Beispiel für eine funktionale Anforderung wäre die Berechnung des Bestellpreises in Euro und in Dollar. Nichtfunktionale Eigenschaften beschreiben hingegen Verhaltensweisen des Systems(Vgl. Seacord et al., 2003, Kapitel 3) und gehen damit über die funktionalen Eigenschaften hinaus. Damit beschreiben funktionale Anforderungen was ein System können muss und nichtfunktionale Anforderungen wie es funktionieren soll. Nichtfunktionale Anforderungen beschreiben außerdem häufig die Qual-

ität der Funktionen und können mehrere andere Anforderungen beeinflussen.(Vgl. Balzert, 2011, S.109ff) Ein Beispiel für nichtfunktionale Eigenschaften wäre, dass die Umrechnung von Euro in Dollar in “wenigen Sekunden” durchgeführt werden muss.

4.3 Requirements analysis

5 Design and Dev artefacts

5.1 System Architecture and Components

5.2 User Interface Design and Implementation

5.3 Prototype Development

6 Demonstration of the Artifact

7 Evaluation of the solution

7.1 Prototype Testing

7.2 Requirements validation

7.3 (App Performance and Usability / User Feedback and Satisfaction)

8 Conclusion

8.1 Summary of the Study

8.2 Contributions and Implications

8.3 Future Work and Recommendations

Bibliography

Abbasi, A., Sarker, S., & Chiang, R. H. L. (2016). Big data research in information systems: Toward an inclusive research agenda. *Journal of the Association for Information Systems*, *17*(2).

Alain Abran, James W. Moore. (2004). *Swebok: guide to the software engineering body of knowledge*. IEEE Computer Society.

Balzert, H. (Ed.). (2011). *Lehrbuch der softwaretechnik: Entwurf, implementierung, installation und betrieb*. Spektrum Akademischer Verlag. https://doi.org/10.1007/978-3-8274-2246-0

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996a). From data mining to knowledge discovery in databases. *AI Magazine*, *17*(3).

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996b). The kdd process for extracting useful knowledge from volumes of data. *Communications of the ACM*, *39*(11), 27–34. https://doi.org/10.1145/240455.240464

Hornby, A. S. (2015). *Oxford advanced learner's dictonary of current english* (9th). Cornelsen.

IEEE. (1990). Ieee standard glossary of software engineering terminology - ieee std 610.12-1990.

ISO Org. (2019). *Norm iso 9241*. Retrieved September 30, 2019, from https://www.iso.org/standard/77520.html

ISO/IEC 25010 (Ed.). (2011). *Systems and software engineering - systems and software quality requirements and evaluation (square)*. ISO/IEC. Retrieved March 23, 2021, from https://www.iso.org/standard/35733.html

Runkler, T. A. (2020). *Data analytics*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-29779-4

Seacord, R. C., Plakosh, D., & Lewis, G. A. (2003). *Modernizing legacy systems: Software technologies, engineering processes, and business practices*. Addison-Wesley. Retrieved February 21, 2021, from https://learning.oreilly.com/library/view/modernizing-legacy-systems/0321118847/

Sharma, R., Mithas, S., & Kankanhalli, A. (2014). Transforming decision-making processes: A research agenda for understanding the impact of business analytics on organisations. *European Journal of Information Systems*, *23*(4), 433–441. https://doi.org/10.1057/ejis.2014.17

Sommerville, I. (2011). *Software engineering* (9th ed.). Pearson.

Appendix

All tables, results, interview data, collected data, used in the report, could be presented here.

Survey	Construct	Item Used	Source
Job (Survey 1)	***		
	***		
General (Survey 2)	***		
	***		
	***		

Table 1: Items Used to Measure Each Construct

Affidavit

I hereby declare that I have developed and written the enclosed master thesis entirely on my own and have not used outside sources without declaration in the text. Any concepts or quotations applicable to these sources are clearly attributed to them. This master thesis has not been submitted in the same or a substantially similar version, not even in part, to any other authority for grading and has not been published elsewhere. This is to certify that the printed version is equivalent to the submitted electronic one. I am aware of the fact that a misstatement may have serious legal consequences.

I also agree that my thesis can be sent and stored anonymously for plagiarism purposes. I know that my thesis may not be corrected if the declaration is not issued.

Mannheim, June 28, 2023

Max Darmstadt