Essays on Business Analytics, Digital Transformation, and Improved Literature Searches

Der Wirtschaftswissenschaftlichen Fakultät der

Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des akademischen Grades

Doktor der Wirtschaftswissenschaften

– Doktor rerum politicarum –

vorgelegte Dissertation

von

Jens Hendrik Paßlick, Master of Science

[redacted]

2020

Abstract

This cumulative dissertation describes research on digital transformation, business analytics, and research methods in information systems research. Within the research methods section, a method for the assessment of research projects is discussed and an article is presented, which examines literature search methods. Thus the article should stimulate a discussion on changing the current practice of literature search in information systems research. In the second part, participatory design in digital transformation is addressed, privacy concerns in bring your own device initiatives are examined, different aspects of chatbots are analyzed, and a taxonomy for predictive maintenance business models is developed. Part three covers different facets of self-service business intelligence, a hybrid machine learning approach and a process model for data science projects, as well as operations research models in industry 4.0. In addition, a research agenda is presented, in which examples for further research opportunities of the respective parts of the dissertation are presented.

Keywords: Literature Search, Industry 4.0, Digital Transformation, Chatbots, Business Analytics, Self-Service Business Intelligence, Data Science

Management Summary

Information systems research (ISR) is characterized, among other things, by its interdisciplinarity. Topics can be examined from different perspectives and with different methods. In this cumulative dissertation, different topics are investigated and different research designs are used. The topics described are all based on articles that have either been published or are in a review process. The dissertation is divided into three major parts. The parts cover research methods of ISR, digital transformation, and business analytics. After describing the three parts, a research agenda is presented, which presents opportunities for further research for each of the topics covered.

Part A presents two articles that deal with the assessment of research and the search for literature. A framework for the assessment of research is presented, which allows the planning for a research project to be presented in a structured way on one page (Passlick et al. 2018). Figure I shows the different boxes of this page.

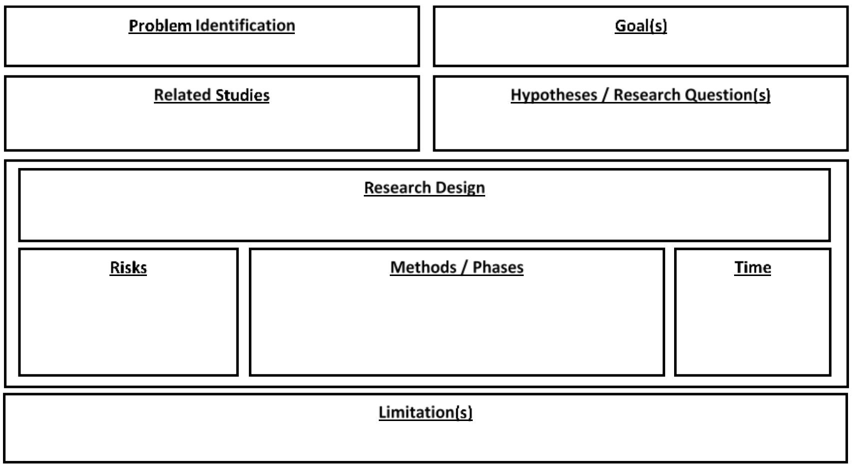

Figure I. The modified Version of the Framework (Passlick et al. 2018)

Roughly, the aspects can be divided into two parts. The boxes for problem identification, the objectives, related studies, and the hypotheses / research questions form the first part, which roughly represents the problem definition of the research project. The second part describes how the questions should be approached and what problems might occur.

The amount of literature is constantly growing, which makes finding relevant literature an increasing challenge. Accordingly, the use of literature search techniques must improve to meet this challenge. The second article of Part A is dedicated to this goal (Passlick et al. 2020d). First of all, it describes the goals of a literature search, the current state of knowledge in literature searches, and which search methods are currently available. Afterwards, literature reviews from high-quality journals of the past years are examined to analyze which methods are used and whether the search methods used have changed in recent years. It is shown that the methods used have changed and that methods with a higher degree of automation have been used in recent years. Afterwards, an exemplary search is conducted to show how the different literature search methods can be combined to use the advantages of each method in a meaningful way. Based on the findings of the previous steps, five theses are formed, which serve as a basis for further research. The first thesis describes that a further increase in the use of automated search methods can be assumed. The second thesis describes that more iterative search approaches are chosen. That Google Scholar has a legitimacy is expressed in the third thesis. The precision of a search becomes increasingly important is the statement of the fourth thesis and thesis five claims that repeatability should be more important than reproducibility in the future.

Part B discusses five articles that deal with digital transformation. The first article deals with the problem that although there are many ways to analyze the visiting behavior of an organization's website, in practice these are far from being exhausted (Janssen et al. 2019). In a design science research approach a participatory design model is developed. The result is a process model, with which individual web analytics reports are developed with the involvement of the future users. The process model is shown in Figure II.

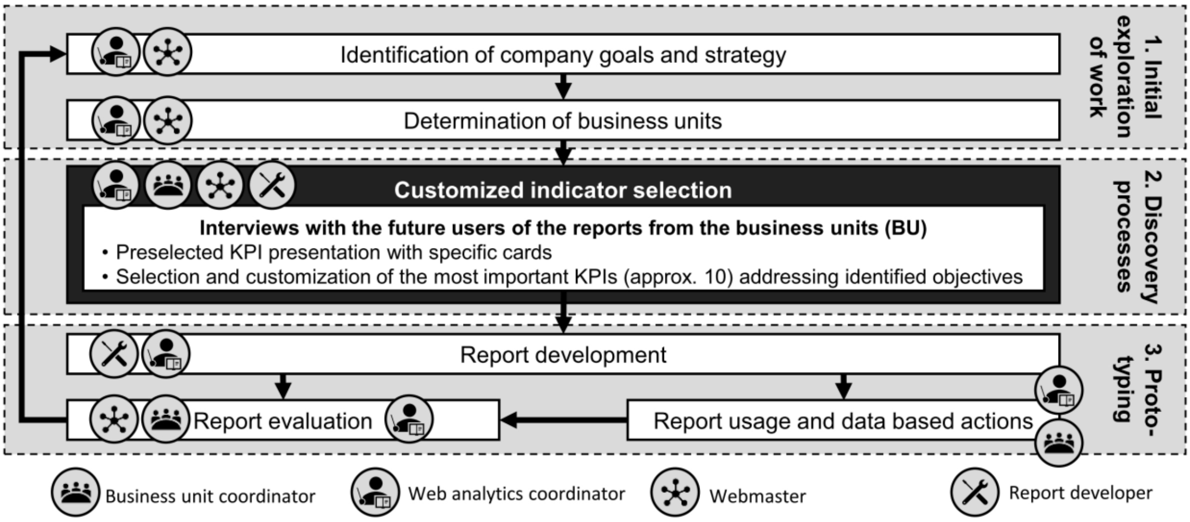

Figure II. The process model for web analytics report development (Janssen et al. 2019, p. 5)

The model shows the relevance of participatory design approaches in the digital transformation, because it can be assumed that the problem with the lower use of web analytics reports can be transferred to other application scenarios of the digital transformation. In the presented example of web analytics, an improved use of the collected analysis data could be achieved by involving the users.

Another article deals with privacy concerns when bring your own device (BYOD) is discussed (Degirmenci et al. 2019). In the article two case studies are conducted to examine the extent to which works councils take data protection concerns into account when evaluating BYOD concepts. In both of the companies studied, data protection is of great importance to the works council and is included in the agreements. In a second step, a structural equation model is used to investigate the relationship between data protection concerns and the use of BYOD. For this purpose, a survey is conducted in Germany, South Korea, and the USA. The investigated structural equation model is shown in Figure III.

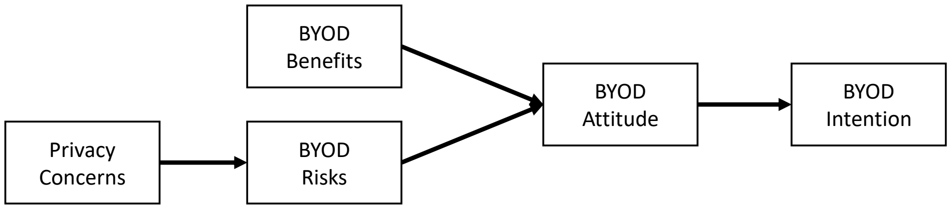

Figure III. The investigated hypotheses (Degirmenci et al. 2019)

For all hypotheses examined, a significant effect can be determined.

The next two articles, which are assigned to the part digital transformation, cover chatbots. Chatbots are a kind of conversational agent and they are discussed as a new form of interaction in very different application scenarios. In the first article, exactly these different application scenarios are examined to structure the different kinds of chatbots (Janssen et al. 2020a). A taxonomy is developed according to a process model of Nickerson et al. (2013), which differentiates domain-specific chatbots. The developed taxonomy consists of the three perspectives which are intelligence, interaction, and context. A total of 17 dimensions are assigned to these three perspectives. The taxonomy is based on a data set consisting of 103 different domain-specific chatbots (Janssen et al. 2020a). When analyzing the chatbots of the data set, it can be seen that many domain-specific chatbots do not yet exhaust all technical possibilities. In addition to this analysis, a cluster analysis is performed with the data set. Five groups can be identified which represent archetypes for domain-specific chatbots.

Based on this, in a second article a framework is developed, which supports the implementation of chatbots (Janssen et al. 2020b). The framework is developed with a design science research approach and is based on 15 expert interviews that are evaluated with grounded theory methods. The expert interviews are used to identify questions that should be asked during the implementation of a chatbot. These questions are structured according to development phases and the four perspectives of the PACT framework by Benyon (2014). Finally, the model is evaluated with a focus group discussion and seven expert interviews.

The last article of the digital transformation part deals with predictive maintenance (PdM) business models (Passlick et al. 2020a). Here, a taxonomy is developed according to the process model by Nickerson et al. (2013), too. In addition to a literature review, the taxonomy is based on a data set containing 113 PdM business models. The dataset was collected at the “Hannover Messe 2018”, the “crunchbase” database, and occasionally from lists on the internet. The finished taxonomy consists of seven dimensions, which includes both classical dimensions for the consideration of business models and one dimension that is based on an internet of things architecture model. In a further step, the frequencies of the respective characteristics of the data set are examined. Additionally, a cluster analysis is performed to identify PdM business model archetypes. Six archetypes can be identified which are shown in Table I.

Table I. PdM business model archetypes (Passlick et al. 2020a, p. 11)

	Archetype					
	1	**2**	**3**	**4**	**5**	**6**
Label	*Hardware development*	*Platform provider*	*All-in-one*	*Information manager*	*Consulting*	*Analytics provider*
Key activities	Hardware development	Provision of an application platform	Universal offer	Edge computer development	Consulting	Software development
Value promise	Condition monitoring	Forecasting	All-in-one solution	Condition monitoring	Condition monitoring	Forecasting
Payment model	One-time sales	Hybrid	Hybrid	Hybrid	Project	Time basis
Deployment channel	Physical	Physical + www (cloud)	Physical + www (cloud)	Physical + www (cloud)	Physical	www (cloud)
Customer segment	No industry focus	Manufacturing industry	No industry focus	Manufacturing Industry	No industry focus	No industry focus
Clients	B2B	B2B	B2B	B2B + B2B2B	B2B	B2B
Information layer	Object sensing and information gathering	Application and services	Multiple	Multiple & information delivering	Application and services	Application, services and information handling
Share in sample (113)*	21%	12%	27%	5%	13%	20%
Example company	Rockwell Automation	Test Motors	National Instruments	IXON	Hitachi Consulting	Senseye

It should be noted that there are new business models such as “information manager”, “platform” and “analytics provider”, as well as established business models such as “hardware development”. In addition, there are connecting business models such as “consulting” or “all-in-one” offers (Passlick et al. 2020a). After deducting the archetypes, an autoencoder procedure is applied to map the differences between the business models in a two-dimensional figure. In this way, it can be shown which archetypes are similar to each other and how homogeneous the business models are within an archetype. Figure IV shows the representation of the analyzed data set, with each symbol representing a business model. Business models of an archetype are represented by the same symbol. It is easy to see that e.g. the archetypes analytics provider and platform provider are similar.

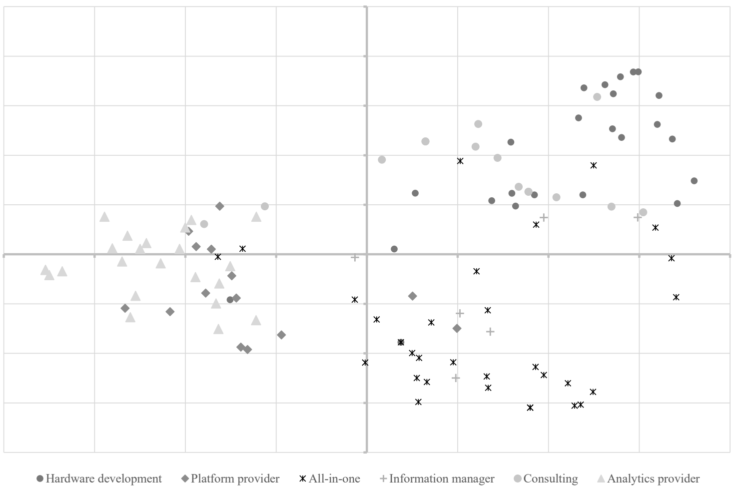

Figure IV. The visualization method with an autoencoder procedure (Passlick et al. 2020a, p. 12)

Part C covers topics of business analytics. The articles can be roughly divided into three streams: Self-service business intelligence (SSBI), operations research in Industry 4.0, and data science. The first article uses a design science research approach to develop an architecture that supports SSBI (Passlick et al. 2017). SSBI describes the ability of employees in a department to create analyses and reports independently of the IT department (Alpar & Schulz 2016). The developed architecture is based on a literature review (Webster & Watson 2002) and semi-structured expert interviews which were evaluated according to Mayring (2002). After completion of the development of the architecture, an applicability check was conducted with a focus group discussion (Morgan 1993). The architecture shows that different user groups exist and their different needs should be addressed with different components. However, components are also needed that promote the collaboration of these user groups.

A further article on SSBI examines which factors increase or decrease the intention of users and potential users to use SSBI (Passlick et al. 2020c). The question is addressed with a structural equation model, which is validated with a survey. It is shown that flexibility, expected time savings, and the importance of data quality lead to a utilitarian value of SSBI, which then increases the expected contribution to information needs. This contribution in turn has a positive influence on the intention to use SSBI. A direct influence of the utilitarian value on the intention to use SSBI was not found. In contrast, the experience with BI applications has a positive influence on the expected contribution to the information needs as well as on the intention to use SSBI. Furthermore, a negative effect of the perceived attention of the company on data quality on the intention to use SSBI was found. Figure V shows the described influences and their levels of significance.

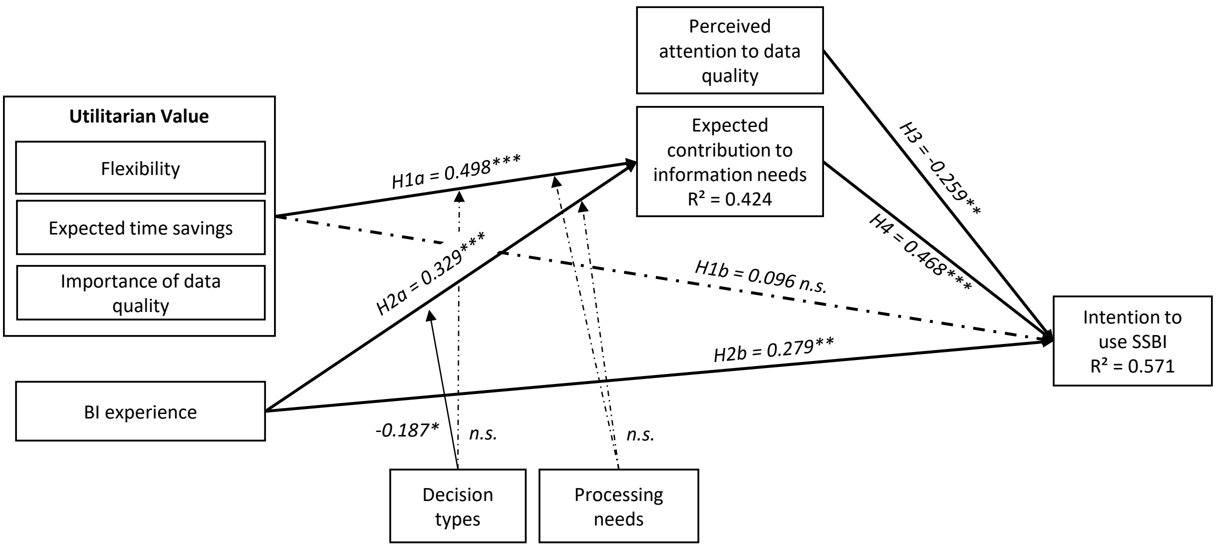

Figure V. Effects on the intention to use SSBI (Passlick et al. 2020c, p. 13)

The third article on SSBI deals with the problem that there is no clear differentiation between different SSBI application scenarios yet (Passlick et al. 2020b). To enable a better differentiation, a taxonomy according to Nickerson et al. (2013) is developed in this research. Based on a literature review, an analysis of SSBI tools, and a case study, the taxonomy development is performed. Afterwards, the taxonomy is evaluated with an illustrative scenario according to Szopinski et al. (2019). The final taxonomy consists of nine dimensions, of which two dimensions were not developed using the analysis of SSBI tools, but based on the literature review and the case study. Within the evaluation, a cluster analysis is performed in which three archetypes are identified. These do not allow conclusions about typical application scenario archetypes in practice, but they show which application scenarios are mainly addressed by the SSBI tool providers.

Two articles deal with operations research models in Industry 4.0, both of which are based on the idea that sensor data can be used to calculate the state of health of a machine. In the models, it is assumed that the state of health can be included in the calculations with a number between 0 and 1. In the first model, the information about the different health states of the machines is then used to calculate how many spare parts should be kept in stock (Dreyer et al. 2018). In addition, the costs of a machine breakdown and the costs for keeping a spare part in stock are also included in the calculation. It is also important to note that the model is based on the assumption that the number of spare parts can be varied at will, which may not be possible in reality, as parts once in stock cannot simply be reduced. However, in new business models that operate within advanced value networks, this would be conceivable.

The second model deals with the question of when maintenance should take place (Olivotti et al. 2018b). It is assumed that there is a set of machines of one type, which should be maintained in groups. The model optimizes how many maintenance groups should be formed, taking into account the expected downtime costs, to find the optimal balance between expected downtime costs and costs for setting up a maintenance event.

Both models have the problem that a large number of possibilities are considered to calculate the correct expected failure costs. As a result, the number of components or machines that can be calculated in one run of the model is limited. Further research is needed here, which could, e.g., by means of a prior grouping of machines or components, enable significantly larger quantities of elements to be optimized.

The last two articles covered in this dissertation can be roughly summarized under the term data science. Data science describes a special form of business analytics, in which more complex statistical, mathematical methods or advanced algorithms are used to analyze data. In the first article, an approach is presented that shows how machine learning algorithms can be combined with human expertise to create robust monitoring systems (Olivotti et al. 2018a). The approach consists of three components, some of which have been prototypically implemented. First an anomaly detection component supervises the sensor data of a machine and detects abnormal data progression. If this occurs, the data is transferred to a monitor component. In this component, a human expert assesses whether the abnormal behavior is really due to a critical wear or other defect and if so, what the cause could be. The result is passed to the classification component, where the data is used to improve the parameters of the anomaly detection component and to improve suggestions that the monitor component makes to the human expert. In this way, the forecast quality can be further improved during operation and the system can be used at the beginning even without extremely large amounts of data.

The second article develops a detailed definition of data science and presents a process model that describes the requirements that exist in the different phases of a data science project (Schulz et al. 2020). The article describes *Data Science as an interdisciplinary field in which, with the help of a scientific, semi-automatic approach, and by applying existing or future analysis methods, knowledge is extracted from partly complex data and made usable under consideration of social effects* (Schulz et al. 2020).

The process in a data science project consists of the phases, project order, data provision, analysis, provision of the analysis model, and usage. The implementation should be carried out according to a scientific procedure. The activities are embedded in the respective domain, which can have a strong influence on the analysis techniques. A large part of the activities is dependent on the organizational and technical infrastructure. Figure VI shows the described elements and their interrelations.

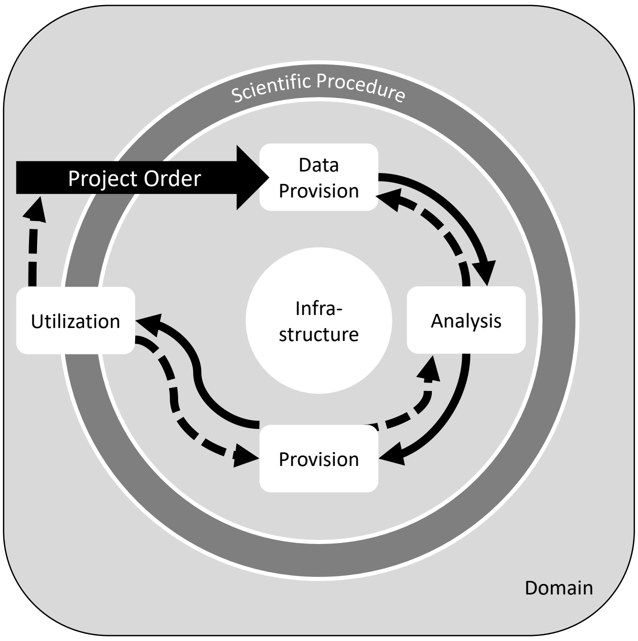

Figure VI. The presented data science process model (Schulz et al. 2020)

The last section of the thesis describes a research agenda for the three major parts of the thesis. For each research area opportunities for further research are presented. These opportunities are distinguished according to the three information systems research traditions. Design oriented research, behavioral research, and economics of information systems are these research traditions. It is shown that there are some connections, not only between the research streams, but also between the major parts of the dissertation. E.g., similar to our research on chatbots, language-based queries of data are also discussed in the context of business analytics. Here, it could be investigated to what extent this new interaction possibility has an impact on a stronger use of business analytics applications.

Table of Contents

Appendix

The following table shows the articles of this dissertation sorted by their occurrence. The appendix contains the articles in the same order. Pages for articles that have been published contain the respective link to the publication. Unpublished articles are attached.

#	Title	Section
1	Assessing Research Projects: A Framework	2.2
2	Towards an Improved Literature Search: A Call for Higher Precession	2.3
3	Using Web Analytics Data: A Participatory Design Model For Individual Web Traffic Report Development	3.2
4	Future of Flexible Work in the Digital Age: Bring Your Own Device Challenges of Privacy Protection	3.3
5	A Self-Service Supporting Business Intelligence and Big Data Analytics Architecture	4.2
6	Virtual Assistance in Any Context - A Taxonomy of Design Elements for Domain -Specific Chatbots	3.4
7	Encouraging the use of self-service business intelligence – an examination of employee-related influencing factors	4.3
8	Self-Service Business Intelligence Application Scenarios - A Taxonomy for Differentiation	4.4
9	A Framework for User-Centered Implementation of Chatbots	3.5
10	Predictive Maintenance as an Internet of Things enabled Business Model: A Taxonomy	3.6
11	Optimizing Machine Spare Parts Inventory Using Condition Monitoring Data	4.5
12	Maintenance Planning Using Condition Monitoring Data	4.6
13	Combining Machine Learning and Domain Experience: A Hybrid-Learning Monitor Approach for Industrial Machines *(Best Paper Nominee)*	4.7
14	DASC-PM v1.0 - Ein Vorgehensmodell für Data-Science-Projekte	4.8