Improving Decision Making in Crowdsourcing

DISSERTATION
of the University of St.Gallen,
School of Management,
Economics, Law, Social Sciences
and International Affairs
to obtain the title of
Doctor of Philosophy in Management

submitted by

Marcel Rhyn
from
Luzern

Approved on the application of

Prof. Dr. Jan Marco Leimeister
and
Prof. Dr. Christoph Riedl

Dissertation no. 5010

D-Druck Spescha, St.Gallen 2020

The University of St.Gallen, School of Management, Economics, Law, Social Sciences and International Affairs hereby consents to the printing of the present dissertation, without hereby expressing any opinion on the views herein expressed.

St.Gallen, May 18, 2020

The President:

Prof. Dr. Bernhard Ehrenzeller

ACKNOWLEDGEMENTS

This dissertation would have not been possible without the support of many people who were involved in this project and who have helped me on my journey. I thank these people and owe them my sincere acknowledgement.

I am extremely grateful to my supervisor Prof. Dr. Jan Marco Leimeister for giving me the opportunity to write this dissertation and for providing his guidance. His assistance, his expertise, and most importantly his support during the past years have been invaluable for the dissertation.

I wholeheartedly thank Prof. Dr. Ivo Blohm and my co-supervisor Prof. Dr. Christoph Riedl for all their help and advice. The success of my work was due to their great support and feedback. It was both a pleasure and privilege to learn from their knowledge and experience – especially during my time in Boston.

I am thankful to my colleagues at the IWI-HSG for all the fruitful discussions, for their insights, and for the friendship. Working with such a great team at the institute was a constant source of inspiration and motivation during the past years. I am especially thankful to Niklas Leicht, who not only served as a mentor and research partner, but who also became a close friend over the years.

I am deeply grateful to my family for their continuous encouragement, their patience, and their unconditional support. For all my life, I could have not been more fortunate to have my parents Silvia and Reto, my brother Sandro, and my grandparents by my side. In so many ways, I owe my work to them.

Last, I want to express my special thanks Emilie. It was her love and her patience that helped me get through rough patches, that always supported me, and that provided me with motivation and joy. I am extremely happy to have her by my side.

St.Gallen, 03.07.2019 Marcel Rhyn

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

List of Abbreviations

AUC	Area Under the Curve
CEO	Chief Executive Officer
cf.	confer
CIO	Chief Information Officer
CSS	Cascading Style Sheets
CST	Crowdsourced Software Testing
DF	Design Feature
DP	Design Principle
DR	Design Requirement
DSR	Design Science Research
DSS	Decision Support System
e.g.	exempli gratia
et al.	et alii
HTML	Hypertext Markup Language
i.e.	id est
ICT	Information and Communication Technology
IS	Information Systems
IT	Information Technology
LDA	Latent Dirichlet Allocation
MSEM	Multilevel Structural Equation Modeling
QA	Quality Assurance
ROC	Receiver Operating Characteristic
SNA	Social Network Analysis
TF-IDF	Term Frequency–Inverse Document Frequency

ABSTRACT

Crowdsourcing represents a powerful approach for organizations to systematically collect data from large networks of people. While research already made great strides in recent years to develop the technological foundations for processing large amounts of user-generated data, it remains mostly unclear how these new data sources and technologies affect decision making in organizations. The objectives of this dissertation are to (1) identify patterns of decision making that emerge in crowdsourcing, (2) understand how decision making in crowdsourcing can be improved with text mining and machine learning, and (3) capture the necessary design knowledge to develop decision support systems in crowdsourcing. To accomplish these objectives, the dissertation is organized in three research streams. The first research stream aims to describe common patterns of decision making in crowdsourcing. It is based on an exploratory interview study that seeks to offer a better understanding of how the structure of decision problems, the characteristics of the available data, and the way in which such data can be generated in crowdsourcing affect decision making. The second research stream aims to examine how decision making in crowdsourcing can be improved with text mining and machine learning. Statistical analyses are used to better understand how crowds create valuable contributions for organizations and how decision makers can identify and process these contributions more efficiently and effectively. Finally, the third research stream follows a design science research approach. It is concerned with integrating the previous findings and capturing design knowledge to develop decision support systems in crowdsourcing. Taken together, the dissertation provides a number of important theoretical contributions. First, it illustrates the limitations of traditional decision making models in data-driven environments, such as crowdsourcing, and describes four common patterns of decision making that emerge when decision makers have access to large-scale, user-generated data. Second, the dissertation provides the empirical foundations to increase the efficiency and effectiveness of decision making in crowdsourcing by offering a better understanding of how crowds generate valuable contributions and how decision makers may process these contributions with text mining and machine learning technologies. Third, the dissertation provides prescriptive design knowledge in the form of design requirements and design principles for the development of decision support systems in crowdsourcing. For practitioners, the dissertation offers recommendations on how to improve the efficiency and effectiveness of decision making in crowdsourcing, how to leverage text mining and machine learning technologies in this context, and how to instantiate the technologies.

ZUSAMMENFASSUNG

Crowdsourcing beschreibt ein Ansatz, mit dem Organisationen systematisch nutzergenerierte Daten von grossen Netzwerken an Personen sammeln können. Während in den vergangenen Jahren bereits die technologischen Voraussetzungen für die Analyse solcher Daten geschaffen wurden, ist noch weitestgehend unklar, wie sich Entscheidungsprozesse in Unternehmen durch den Zugang zu diesen Daten verändern. Das Ziel der Dissertation ist es, (1) eine Charakterisierung der im Crowdsourcing entstehenden Entscheidungsprozesse vorzunehmen, (2) ein besseres Verständnis dafür zu schaffen, wie diese Entscheidungsprozesse mithilfe von Text Mining und Machine Learning verbessert werden können und (3) Gestaltungswissen für Entscheidungsunterstützungssysteme im Crowdsourcing zu entwickeln. Dazu ist die Dissertation in drei Teilen organisiert. Ziel des ersten Teils ist es, typische Muster von Entscheidungsprozessen im Crowdsourcing zu beschreiben. Es wird eine explorative Interviewstudie durchgeführt, um besser zu verstehen, wie die Struktur von Entscheidungsproblemen, die Eigenschaften der verfügbaren Daten und die Art und Weise der Datengenerierung die Entscheidungsprozesse beeinflussen. Im zweiten Teil wird untersucht, wie Entscheidungsprozesse im Crowdsourcing durch Text Mining und Machine Learning verbessert werden können. Mithilfe von statistischen Analysen soll erklärt werden, wie Netzwerke von Personen wertvolle Beiträge für Unternehmen schaffen und wie diese Beiträge effizient und effektiv identifiziert werden können. Der dritte Teil folgt abschliessend einem gestaltungsorientierten Forschungsansatz. Ziel ist es, Gestaltungswissen für die Entwicklung von Entscheidungsunterstützungssystemen im Crowdsourcing auszuarbeiten. Zusammengefasst bietet die Dissertation damit eine Reihe von wichtigen theoretischen Beiträgen. Erstens zeigt die Dissertation die Grenzen der traditionellen Entscheidungstheorie in datengetriebenen Kontexten, wie beispielsweise im Crowdsourcing, auf und beschreibt vier Muster von Entscheidungsprozessen, die typischerweise auftreten. Zweitens bietet die Dissertation neue empirische Erkenntnisse, wie Netzwerke von Personen Wert für Unternehmen generieren und wie infolge dieser Erkenntnisse die Effizienz und Effektivität von Entscheidungsprozessen durch Text Mining und Machine Learning verbessert werden können. Drittens bietet die Dissertation präskriptives Gestaltungswissen in der Form von Gestaltungsanforderungen und -prinzipien, welche die Entwicklung von Entscheidungsunterstützungssystemen im Crowdsourcing leiten sollen. Für Praktiker enthält die Dissertation Empfehlungen, wie die Effizienz und Effektivität von Entscheidungsprozessen im Crowdsourcing gesteigert werden kann, wie Technologien im Bereich Text Mining und Machine Learning dabei eingesetzt werden können und wie deren Implementierung in der Praxis aussehen sollte.

1 INTRODUCTION

1.1 Motivation

In recent years, organizations have begun to systematically harness the collective knowledge and creativity of the “masses” (Zhao & Zhu, 2014). Driven by advances in the field of information and communication technologies (ICTs), organizations are now able to interact with large online crowds that may include their customers (e.g., Leimeister, Huber, Bretschneider, & Krcmar, 2009), their employees (e.g., Zuchowski, Posegga, Schlagwein, & Fischbach, 2016), or volunteers (e.g., Magnusson, 2009) to carry out value creation activities and solve organizational problems. *Crowdsourcing* is an umbrella term for approaches that aim to leverage the potential of such large networks of people (Geiger & Schader, 2014) and represents one of the most widespread illustrations of modern “technology-mediated mass collaboration” (Love & Hirschheim, 2017, p. 315). The fundamental principle of crowdsourcing revolves around the use of an open call through which an organization outsources a predefined task to a potentially large and diverse network of individuals (Blohm, Leimeister, & Krcmar, 2013). Instead of relying on only few inputs from dedicated employees or contractors, crowdsourcing deliberately seeks to span organizational boundaries and gather user-generated inputs from vast pools of independent contributors to resolve a given task (Afuah & Tucci, 2012; Chiu, Liang, & Turban, 2014). This allows organizations to mobilize knowledge, creativity, or workforce distributed amongst a large and diverse panel of people (Schenk & Guittard, 2011). The approach has found widespread adoption in the industry by well-renowned organizations such as Adobe, Best Buy, Dell, Google, Starbucks, or Salesforce (Huang, Singh, & Srinivasan, 2014). By 2015, up to 85 percent of the top 100 brands have been reported to use crowdsourcing (Owyang, 2015).

One of the major changes that crowdsourcing brings to organizations is the way in which organizations can generate and leverage data (Chiu et al., 2014). So far, organizations have primarily dealt with structured, enterprise-specific data that were retrieved through internal information systems (IS). Decisions were based upon data collected through systematic and purposeful processes that address specific information needs (Constantiou & Kallinikos, 2015). With crowdsourcing, organizations are now able to collect data from much larger and more diverse panels of contributors to uncover new behavioral trends (e.g., Brynjolfsson, Geva, & Reichman, 2015), design innovative products (e.g., Leimeister et al., 2009), or gain insights about user preferences (e.g., Blohm, Riedl, Füller, & Leimeister, 2016). With these new opportunities to source and evaluate user-generated data by crowds, three important issues emerge.

First, it is assumed that crowdsourcing will lead to changes in decision making (Chiu et al., 2014). Decision making describes the sequences of data-processing activities and evaluation patterns, by which focal actors in organizations (e.g., innovation managers) analyze data and choose courses of actions to solve a decision problem (e.g., decide what type of new product to develop). Given that crowdsourcing offers the opportunity to freely source and evaluate large amounts of user-generated data, many scholars expect a shift towards more open, data-driven decision making in crowdsourcing that draws upon actual information about people's behavior, opinions, or choices rather than subjective intuition and experience of the decision maker (e.g., Abbasi, Sarker, & Chiang, 2016; Lycett, 2013; Sharma, Mithas, & Kankanhalli, 2014). Bonabeau (2009), in particular, predicts that the use of crowdsourced data will mark "a paradigm shift in the way companies make decisions" (p. 46). However, Abbasi et al. (2016) note that further research is needed to understand how organizations and individuals "actually make decisions" (p. xii) in such environments. It is crucial to understand "how people perceive problems, use information, and analyze data in developing solutions, ideas, and knowledge" (Marchand & Peppard, 2013, p. 109). The characteristics and patterns of decision making that emerge when decision makers have access to large amounts of crowdsourced data are still largely unclear (Abbasi et al., 2016; Sharma et al., 2014).

Second, crowdsourcing requires new approaches to support decision makers in processing and evaluating data. With crowdsourcing, organizations now face unstructured data coming from diverse and fragmented sources (Barbier, Zafarani, Gao, Fung, & Liu, 2012). Decision makers in organizations are required to process much more data than beforehand and face larger and more diverse sets of options. As a result, processing and extracting relevant information from user-generated data are often described as the most time-consuming and cost-intensive activities in crowdsourcing (Blohm et al., 2013). Google, for example, required almost three years and 3'000 employees to analyze the 150'000 ideas that were submitted to its *Project 10^{100}* (Blohm et al., 2013). To date, existing research has mainly focused on manual techniques for the evaluation of crowdsourced contributions, such as peer reviews, agreement filters, rating scales, or expert panels (Allahbakhsh et al., 2013). However, the growing quantity and complexity of information in crowdsourcing make it nearly impossible to process all data manually. In order to cope with large amounts of user-generated contributions in crowdsourcing, research and practice are increasingly aiming to use text mining and machine learning to support the evaluation. The ability of these algorithms to recognize patterns and extract useful information in a fast, scalable, and repeatable way is argued to be a key factor for the (semi-)automated analysis of crowdsourced data (Chen, Chaing, & Storey, 2012). However, Constantiou and Kallinikos (2015) emphasize that it is still unclear what processes of data reduction and aggregation are required to support decision makers in crowdsourcing and how unstructured data generated by crowds can be made relevant for business purposes. Thus, there is need to gain a better understanding of how crowds create valuable contributions for organizations and how text mining and machine learning can be used to support decision makers in processing and identifying them.

Third, with the availability of large amounts of unstructured data and new technologies to process these data, crowdsourcing makes it necessary to reconsider and investigate adequate IS designs in such contexts. Adequate IS designs are not only crucial for the acceptance and adoption of new decision support technologies (W. Wang & Benbasat, 2005), they also affect how people are able to improve their decision making and make use of user-generated data (Sharma et al., 2014). In existing crowdsourcing research, studies have mostly focused on domain-specific instantiations of decision support technologies to demonstrate their technical capabilities (e.g., Barbier et al., 2012; Feng, Chen, Jones, Fang, & Xu, 2015; Hoornaert, Ballings, Malthouse, & Van den Poel, 2017; Nagar, De Boer, & Garcia, 2016; Walter & Back, 2013). They have focused less on design knowledge that guides the deployment and adoption of text mining and machine learning in decision support systems (Zhao & Zhu, 2014). Thus, while the technical development of text mining and machine learning algorithms to process unstructured data is already advanced, it is still unclear how decision support systems based on these algorithms should be designed when dealing with crowdsourced data (Abbasi et al., 2016). Scholars have thus called for research to "contribute guidelines for design artifacts" that support decision making in these contexts (Abbasi et al., 2016, p. xvii).

The dissertation aims to address these issues. The objectives of this dissertation are to (1) identify patterns of decision making that emerge in crowdsourcing, (2) understand how decision making in crowdsourcing can be improved with text mining and machine learning, and (3) capture the necessary design knowledge to develop decision support systems in crowdsourcing based on these technologies.

1.2 Research Questions and Research Methods

To accomplish these objectives, the dissertation follows three primary research streams with separate research questions and research methods. Each research stream consists of one or several studies with substantial standalone contributions. The first research stream uses a qualitative research approach to describe the patterns of decision making that emerge in crowdsourcing. It is based on an exploratory interview study that aims to explain how the structure of decision problems, the characteristics of the available data, and the way in which such data can be generated in crowdsourcing affect decision making. Building upon these findings, the second research stream aims to examine how to improve decision making in crowdsourcing. Three quantitative studies are conducted to better understand how crowds create valuable contributions for organizations and how decision makers can identify and process these contributions more efficiently and effectively. Finally, the third research stream integrates the previous findings and aims to capture the necessary design knowledge on how to build decision support systems in crowdsourcing. For this purpose, a design science research study is conducted. The following paragraphs describe each research question and the methodological approach used in the studies in more detail.

Research Question 1: What decision making patterns emerge in crowdsourcing?

The first research question seeks to investigate the characteristics of decision making in crowdsourcing and identify patterns of decision making that emerge in crowdsourcing. Following the well-established phase theorem of decision making (cf. Boonstra, 2003; Mintzberg, Raisinghani, & Théorêt, 1976; Simon, 1960), decision making is viewed as a sequence of data-processing activities and evaluation patterns, by which focal actors process data (e.g., screen ideas), assess different options (e.g., evaluate their projected costs and earnings), and commit to a particular action (e.g., implement an idea). This perspective makes it possible to study how decision makers typically source, process, and use crowdsourced data to inform decisions and allows to examine how different types of decision problems and modes of acquiring information induce patterns.

For this purpose, a qualitative research approach with semi-structured interviews is used. Qualitative research allows data to be collected in natural settings and ultimately offers rich and holistic insights through local groundedness (Miles, Huberman, & Saldaña, 2014). It is especially well-suited to capture events, processes, or structures experienced by decision makers in an explorative way and thus represents an adequate way of analyzing the characteristics of decision making in crowdsourcing (Miles et al., 2014). For the analysis of the interviews, a multi-staged, inductive coding approach based on Gioia et al. (2013) in combination with a temporal bracketing strategy proposed by Langley (1999) is used. This makes it possible to reconstruct and analyze the sequences of data-processing activities and evaluation patterns, by which decision makers typically generate and analyze user-generated data in crowdsourcing.

Answering the first research question of the dissertation is intended to yield two major contributions. On the one hand, the results aim to offer a better understanding of different types of decision making patterns that emerge in crowdsourcing. These patterns may explain how the structure of decision problems, the characteristics of the available data, and the way in which such data can be generated in crowdsourcing affect decision making. On the other hand, the results aim to contribute to a better understanding of how information systems may support decision making in crowdsourcing. In this way, they pave the way for the remainder of the dissertation to examine in more detail how decision makers can benefit from text mining and machine learning technologies.

Research Question 2: How can decision making in crowdsourcing be improved with text mining and machine learning?

The second research question aims to develop the empirical foundations to increase the efficiency and effectiveness of decision making in crowdsourcing. It intends to offer a better understanding of how crowds generate valuable contributions and how decision makers may process and identify these contributions with text mining and machine learning technologies. Three quantitative studies are conducted for this purpose.

The first study focuses on improvements to the efficiency of decision making. It examines the potential of text mining and machine learning to automate repetitive data-processing tasks for decision makers. The aim is to build a classifier capable of predicting the quality of crowdsourced contributions based on their textual characteristics to automatically filter them and reduce the manual workload for decision makers. To achieve this objective, a two-pronged approach is chosen. First, a regression analysis is conducted to identify the textual characteristics that are associated with contribution quality in crowdsourcing. Afterwards, these textual characteristics are used for predictive modeling with machine learning algorithms. Hence, this study provides a set of variables and models to explain and predict contribution quality in crowdsourcing. These models and variables can be used to assess textual contributions with machine learning algorithms and contribute to a partial automation of the evaluation process.

The second study is concerned with the effectiveness of decision making. Existing research shows that, when faced with a large number of contributions in crowdsourcing, decision makers often attend to only a subset of contributions due to their limited ability to process all available information (Piezunka & Dahlander, 2015). This ultimately hampers their ability to make effective decisions as it becomes increasingly difficult to identify useful contributions in the vast pools of data generated by crowds. Thus, the second study investigates the potential of network analysis and text mining to support decision makers in tracking the origin of contributions, analyzing their content, and ultimately spotting the most useful ones. For this purpose, the study uses statistical approaches from the fields of information retrieval, text mining, and network analysis in combination with logistic regression to identify distinctive characteristics of useful ideas. In this way, the study provides a set of variables that represent statistically significant predictors for useful contributions. These predictors may be used to develop algorithms that assist the identification of useful contributions on crowdsourcing platforms with related business analytics or decision support systems and ultimately make the evaluation of crowdsourced contributions more effective.

The third study extends these findings. It examines how collaboration in online crowds affects the manner in which individuals become creative and solve innovation problems. The study combines statistical approaches from the fields of network analysis and text mining to study 8 years of activity by a crowd who developed more than 200’000 ideas and comments to solve innovation problems for organizations. It this way, the study provides a more in-depth understanding of how crowds create valuable contributions for organization and how network analysis and text mining can be used to potentially support decision makers in processing and identifying these contributions.

Research Question 3: What design principles should guide the development of intelligent decision support systems in crowdsourcing?

The third research question aims to conclude the dissertation. It is concerned with the development of design principles that prescribe how to build decision support systems based on text mining and machine learning technologies. Design principles are one of the most widely used vehicles to “convey design knowledge that contribute beyond instantiations applicable in a limited use context” (Chandra, Seidel, & Gregor, 2015, p. 4039). They capture abstract design knowledge and prescribe “what and how to build an artifact in order to achieve a predefined design goal” (Chandra et al., 2015, p. 4040).

To develop these design principles, the study follows a design science research approach (Hevner, March, Park, & Ram, 2004). Design science research (DSR) represents a well-established approach in information systems research that is concerned with the creation of artifacts that seek to extend the boundaries of human and organizational capabilities (Hevner et al., 2004). These artifacts may range from specific instantiations in the form of implemented software or algorithms to more theoretical contributions in the form of abstract design principles (Gregor & Jones, 2007). The study follows the standard DSR process proposed by Peffers et al. (2007) to develop the design principles. This process synthesizes the common phases of design science research proposed in existing literature (e.g., Hevner et al., 2004; Kuechler & Vaishnavi, 2008; Walls, Widmeyer, & El

Sawy, 1992). It consists of the six phases of specifying the problem, defining the objectives of a solution, designing the solution, demonstrating the solution's feasibility, evaluating the solution, and communicating the results (Peffers et al., 2007). As design science research represents an inherently iterative and incremental approach (Hevner et al., 2004) and aims to bridge theory and practice (Holmström, Ketokivi, & Hameri, 2009), three design-and-evaluate cycles with a cross-industry research consortium are conducted (Österle & Otto, 2010). They aim to (1) define the design requirements, design principles, and design features for decision support systems in crowdsourcing, (2) develop software prototypes for a formative evaluation, and (3) instantiate these prototypes in a decision support system for a summative evaluation in organizations.

This study contributes to the dissertation in two ways. For researchers, it captures the design knowledge that has been gained in the design science research project. The design principles provide the theoretical foundation for developing decision support systems in crowdsourcing based on text mining and machine learning technologies. For practitioners, the study provides a set of design features for the actual implementation of these algorithms in crowdsourcing. Such decision support mechanisms may serve as additional value propositions for crowdsourcing platforms or as means to increase the efficiency and effectiveness of internal data processing and decision making.

1.3 Structure of the Dissertation

To address the research questions as defined in the previous section, the remainder of the dissertation is organized in twelve major chapters. Figure 1 (p. 9) provides an overview of the dissertation's structure and the content of each chapter.

Chapter 2 provides the theoretical background of the dissertation for all subsequent chapters. It reviews related work in the fields of search theory, crowdsourcing, decision making, and decision support. First, in section 2.1, search theory is introduced as a theoretical foundation for crowdsourcing. Based on this foundation, section 2.2 describes the concept of crowdsourcing and outlines the challenges associated with crowdsourced data. Afterwards, section 2.3 provides the theoretical background on decision making and describes the challenges of decision making in crowdsourcing. Finally, in section 2.4, research on decision support and decision support systems is reviewed.

Chapter 3 builds upon this theoretical background and addresses the first research question of the dissertation. It examines patterns of decision making in crowdsourcing. The chapter presents the results of a qualitative interview study. The results explain how the structure of decision problems, the characteristics of the available data, and the way in which such data can be generated in crowdsourcing affect decision making.

Chapter 4 addresses the second research question of the dissertation. It focuses on improvements to the efficiency of decision making in crowdsourcing and examines the potential of text mining to automate repetitive data-processing tasks for decision makers. This chapter presents the results of a study using multiple regression analysis and predictive modeling with machine learning algorithms.

Chapters 5 and 6 also address the second research question but focus on the effectiveness of decision making in crowdsourcing. The chapters are based on two studies. Chapter 5 investigates the potential of network analysis and text mining to support decision makers in tracking the origin of contributions, analyzing their content, and spotting the most useful ones. Chapter 6 extends these findings and investigates the effects of collaboration on an individual's ability in crowds to create valuable contributions.

Chapter 7 addresses the third research question of the dissertation. It investigates how to design decision support systems in crowdsourcing based on text mining and machine learning. The chapter reports the results of the DSR study that was conducted with a cross-industry research consortium. It describes design requirements, design principles and design features for the development of decision support systems in crowdsourcing.

Chapter 8 integrates the results of the previous chapters and synthesizes the main findings of the dissertation. The objective is to provide an overall discussion of the results. The findings of chapters 3 to 7 are discussed with regard to the three research gaps and research questions outlined previously in sections 1.1 and 1.2.

Chapters 9 and 10 describe the theoretical and practical contributions of the dissertation. Chapter 9 outlines the implications of the findings for research on decision making, research on crowdsourcing, and research on decision support systems. Chapter 10 illustrates the practical implications of the dissertation's findings. It shows how the findings may be used to improve crowdsourcing processes, build text mining and machine learning models, and design decision support systems in crowdsourcing.

Finally, chapters 11 and 12 conclude the dissertation. Chapter 11 acknowledges the limitations of the dissertation and describes potential avenues for future research to extend the findings of the dissertation. The chapter provides an outlook and potential research agenda for studies that intend to delve deeper into decision making in crowdsourcing. Chapter 12 offers a concluding summary of the dissertation.

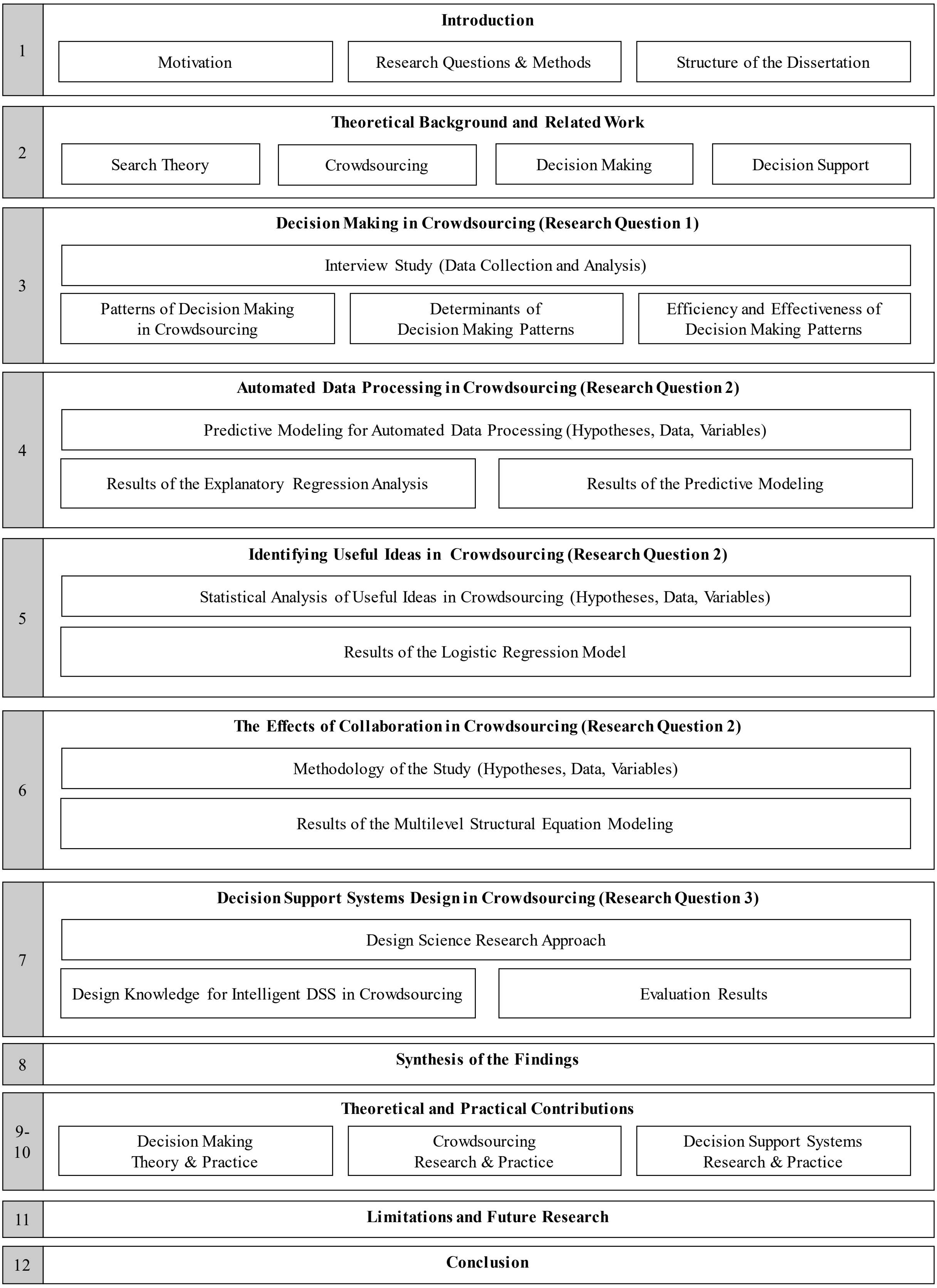


Figure 1: Structure of the Dissertation

Source: Own Illustration

1.4 Overview of Publications

Parts of this dissertation have been published in proceedings of peer-reviewed conferences or are under review for publication. Table 1 provides the list of the papers that have been published or are currently under review for publication. Table 1 also indicates in which chapters the content of these papers has been used.

No.	*Publication*	*Chapter*	*RQ*
1	Rhyn, M., & Blohm, I. (2019). Patterns of Data-Driven Decision-Making: How Decision-Makers Leverage Crowdsourced Data. Under Review: *Proceedings of the 40th International Conference on Information Systems (ICIS)*. Munich, Germany: AIS.	1, 2, 3, 8, 9, 10, 11	RQ1
2	Rhyn, M., & Blohm, I. (2017a). A Machine Learning Approach for Classifying Textual Data in Crowdsourcing. *Proceedings of the 13th International Conference on Wirtschaftsinformatik (WI)*. St. Gallen, Switzerland: AIS.	1, 2, 4, 8, 9, 10, 11	RQ2
3	Rhyn, M., Blohm, I., & Leimeister, J. M. (2017). Understanding the Emergence and Recombination of Distant Knowledge on Crowdsourcing Platforms. *Proceedings of the 38th International Conference on Information Systems (ICIS)*. Seoul, South Korea: AIS.	1, 2, 5, 6, 8, 9, 10, 11	RQ2
4	Rhyn, M., & Blohm, I. (2017b). Combining Collective and Artificial Intelligence: Towards a Design Theory for Decision Support in Crowdsourcing. *Proceedings of the 25th European Conference on Information Systems (ECIS)*. Guimarães: AIS.	1, 2, 7, 8, 9, 10, 11	RQ3
5	Rhyn, M., Leicht, N., Blohm, I., & Leimeister, J. M. (2020). Opening the Black Box: How to Design Intelligent Decision Support Systems in Crowdsourcing. Under Review: *Proceedings of the 15th International Conference on Wirtschaftsinformatik (WI)*. Potsdam, Germany: AIS.	1, 2, 7, 8, 9, 10, 11	RQ3

Table 1: Overview of Publications

Source: Own Illustration

2 THEORETICAL BACKGROUND AND RELATED WORK

The following sections provide the theoretical background of the dissertation.[1] They offer a review of related work on search theory (section 2.1), crowdsourcing (section 2.2), decision making (section 2.3), and decision support (section 2.4).

2.1 Search Theory

In organizational theory (e.g., Dosi, 1982; Nelson & Winter, 1982), search refers to "the controlled and proactive process of attending to, examining, and evaluating new knowledge and information" in order to solve organizational problems or drive innovation (Q. Li, Magiitti, Smith, Tesluk, & Katila, 2013, p. 893). Existing research generally distinguishes between two notions of search: local search and distant search (Katila & Ahuja, 2002; March, 1991). In local search, an organization relates to knowledge that is close to its existing knowledge base and addresses problems by building upon established capabilities and routines (Stuart & Podolny, 1996). Research indicates that this is the predominant search strategy used by organizations (Martin & Mitchell, 1998; Tripsas & Gavetti, 2000). Local solutions are familiar and can be found at relatively low costs or communication efforts (Carlile, 2002; Helfat, 1994; Rosenkopf & Almeida, 2003). This makes local search efficient and reliable for organizations. However, while local knowledge allows for exploitation and facilitates learning, it often suffers from bounded rationalities and lacks the required diversity for effective problem solving (Laursen, 2012; Rosenkopf & Nerkar, 2001; Rothaermel & Alexandre, 2009).

In distant search, organizations move away from predefined routines and reach beyond their boundaries to access unfamiliar knowledge and incorporate new information (Katila & Ahuja, 2002). A large body of literature suggests that gaining access to distant knowledge greatly benefits organizations in adapting, diversifying, or reinventing themselves (Katila & Ahuja, 2002; Katila, Chen, & Piezunka, 2012). In this sense, distant knowledge has been found to inherit a particularly high potential for developing breakthrough innovation (Fleming, 2001; Fleming & Sorenson, 2004). The search for distant knowledge can either span technological boundaries or organizational boundaries. Searching beyond organizational boundaries is argued to be especially impactful for exploration (Rosenkopf & Nerkar, 2001). There is a large body of literature suggesting

[1] Parts of this chapter have been published in proceedings of peer-reviewed conferences. Section 2.1 has been published in Rhyn et al. (2017). Section 2.2 contains parts published in Rhyn et al. (2017) and Rhyn and Blohm (2017a). Section 2.4 contains parts published in Rhyn and Blohm (2017b). The content has been reformatted, restructured, modified, and extended to provide a comprehensive theoretical background for this dissertation.

that the interaction with external sources of knowledge is essential to innovating or solving problems in organizations (Chesbrough, 2003; Powell, Koput, & Smith-Doerr, 1996; Von Hippel, 2005).

However, Katila and Ahuja (2002) outline that search efforts not only vary with regard to their scope (i.e., local versus distant search) but also with regard to their depth. Differences in depth of search can lead to varying degrees of familiarity with the acquired knowledge and, in turn, affect the organizations ability to generate new solutions from it (Katila & Ahuja, 2002). In this regard, much research emphasizes the role of knowledge recombination (e.g., Fleming, 2001; Fleming & Sorenson, 2004; Hargadon & Sutton, 1997). It is argued that, "by combining firm-specific accumulated understanding of certain knowledge elements (depth) with new solutions (scope), firms are more likely to create new, unique combinations that can be commercialized" (Katila & Ahuja, 2002, p. 1180; Winter, 1984).

As outlined by Piezunka and Dahlander (2015), organizations may rely on different means to access distant knowledge and combine it with local knowledge, for example, by hiring employees (e.g., Rosenkopf & Nerkar, 2001), by acquiring new organizational units (e.g., Ahuja & Lampert, 2001), or by forming alliances (e.g., Stuart & Podolny, 1996). In recent years, crowdsourcing has emerged as a powerful, IT-facilitated approach for organization to gain access to distant knowledge by broadcasting tasks or value creation activities to a large and diverse network of people (Afuah & Tucci, 2012).

2.2 Crowdsourcing

2.2.1 The Concept of Crowdsourcing

The fundamental principle of crowdsourcing revolves around the use of an open call through which an organization engages an independent network of people and leverages their collective knowledge, creativity, or workforce to resolve a predefined problem or task (Blohm et al., 2013; Zhao & Zhu, 2014). While crowdsourcing can be seen as an innovative way of organizing work (e.g., Durward, Blohm, & Leimeister, 2016) or engaging with potential customers (e.g., Schulten & Schaefer, 2015), it has gained particular interest in search theory as a potential solution to distant search in organizations (e.g., Afuah & Tucci, 2012; Piezunka & Dahlander, 2015). The approach specifically seeks to mobilize resources distributed amongst a large number of individuals (Schenk & Guittard, 2011). Compared to traditional search mechanisms that target only few dedicated employees or contractors, participation in crowdsourcing is generally non-discriminatory (Zogaj, Bretschneider, & Leimeister, 2014) and facilitates the self-selection of potential contributors to a problem (Afuah & Tucci, 2012).

This is based on the tenet that individuals who are not bound to the current thinking in the field of a particular problem are capable of offering "perspectives and heuristics that are novel and thus useful for generating solutions to these problems" (Jeppesen & Lakhani, 2010, p. 1019). While it is more difficult for these individuals to assess the feasibility of a solution or an idea (Poetz & Schreier, 2012), existing research provides empirical evidence that individuals distant to a domain are able produce more original and radical ideas than experts in the field (Kristensson, Gustafsson, & Archer, 2004; Magnusson, 2009). Crowdsourcing allows organizations to gain access to such distant knowledge and collect a high number of diverse solutions from outside their boundaries in a very efficient and effective way (Afuah & Tucci, 2012; Chesbrough, 2003).

2.2.2 Data Collection in Crowdsourcing

Given the decentralized nature of crowdsourcing, the interaction between the organizations and their crowds generally unfolds on IT-based platforms (Blohm, Zogaj, Bretschneider, & Leimeister, 2018; Doan, Ramakrishnan, & Halevy, 2011). On the one hand, these platforms enable organizations to allocate tasks to a crowd and coordinate their activities. On the other hand, the platforms act as focal points for organizations to aggregate and retrieve contributions. In this way, the platforms represent the interface between the organizations seeking to broadcast a task and a large number of contributors willing to perform the task. In general, literature distinguishes between two types of approaches to crowdsourcing on these platforms: competition-based crowdsourcing and collaboration-based crowdsourcing (Blohm et al., 2013; Zhao & Zhu, 2014).

Competition-based crowdsourcing seeks to efficiently match organizations facing a particular problem with individuals possessing the relevant knowledge for its resolution (Felin & Zenger, 2014). It is especially well suited for technical problems or design projects (Bourreau, Gensollen, & Moreau, 2012). However, as outlined by Majchrzak and Malhotra (2013), a problem with an approach to innovation that uses the crowd for the sole purpose of gathering isolated contributions "is the lack of collaborative discourse that leads to generative co-creation, a foundational requirement for innovation from diverse sources" (p. 263). Thus, while research on crowdsourcing initially focused on temporary ideas competitions, organizations are increasingly interested in issuing more long-term calls and using collaborative crowdsourcing platforms as an integral part of their search activity – both internally and externally (Schemmann, Herrmann, Chappin, & Heimeriks, 2016; Zuchowski et al., 2016).

Collaborative crowdsourcing focuses on the recombination of knowledge and works best when members of the crowd can share information freely and accumulate or alter ideas (Boudreau & Lakhani, 2013). In these collaborative settings with interactions unfolding on the platform, crowds can be regarded as connected networks of people that form around a focal organization to jointly generate new ideas or solutions (Simula & Ahola, 2014). Such networks of people may provide organizations with "access to diverse, and otherwise hidden knowledge, while at the same time providing in some circumstances support for rich forms of knowledge exchange" (Felin & Zenger, 2014, p. 922). In this way, platforms that foster interaction are especially powerful for sourcing new knowledge as they allow crowds to engage in a discourse and jointly develop alternatives, share ideas, or modify problem observations "to co-create solutions that would not have been suggested if only a single perspective had been represented" (Majchrzak & Malhotra, 2013, p. 263). A number of studies provide evidence that such co-created ideas in crowdsourcing are generally of higher quality than those autonomously submitted by individuals (e.g., Blohm, Bretschneider, Leimeister, & Krcmar, 2010; Majchrzak & Malhotra, 2013).

2.2.3 Characteristics of Crowdsourced Data

Although there are many benefits to using a large online crowd to solve tasks, opening up the participation to a decentralized network of individuals makes it more difficult to control the content and format of the data (Lukyanenko, Parsons, & Wiersma, 2014). This is especially challenging for the broad range of crowdsourcing settings that are based on contributions submitted in an free text format, such as ideas on open innovation platforms (Leimeister et al., 2009) or user feedback in crowdsourced software testing (Zogaj et al., 2014). These textual contributions represent an unstructured data format and come with several problematic characteristics regarding both their contextual and their representational quality (cf. R. Y. Wang & Strong, 1996).

First, there is no ground truth to contributions such as ideas, feedback, or reviews. Hence, for these types of textual contributions, it is inherently complex to assess and compare contextual characteristics such as the relevancy or the completeness of the information (Barbier et al., 2012; Blohm et al., 2016). Members of a crowd may have different perceptions of what is relevant or interesting for such a task and will typically cover a broad range of topics in their contributions (Lukyanenko et al., 2014). Some contributions may lack focus and specificity; others may even include contradictory or false information (Blohm et al., 2013).

Second, the representation of information in textual contributions is generally of high variance and diversity (Barbier et al., 2012). Depending on their background and their degree of expertise, members of a crowd may express themselves in very distinct ways, using different expressions for similar issues or similar expressions for different issues (Blohm et al., 2013). Hence, not only is there a wide range of potential topics that could be covered in crowdsourced contributions, but also a wide range of potential descriptions and terms for these topics. This is aggravated by the fact that textual data generated by a crowd typically entail a high amount of noise due to spelling mistakes, grammatical errors, excessive punctuation, or informal writing styles (Barbier et al., 2012).

2.2.4 Evaluation of Crowdsourced Data

Given the previously described characteristics of textual data in crowdsourcing, it is often difficult to use traditional approaches to quality control (Allahbakhsh et al., 2013). For example, it is not possible to employ gold standard data as there is typically no ground truth to which the contributions can be compared. Hence, companies have to rely mostly on a manual assessment of the contributions. That is, someone has to read the contributions, evaluate the quality of the content, compare it to the requirements of the task, and either accept or dismiss the input for further consideration by the company (Zogaj et al., 2014). Expert panels that review and select relevant inputs represent one of the most reliable yet impractical means for this step (Blohm et al., 2016). The volume of textual data and the rate at which they are created in crowdsourcing often exceed their information processing capacities (Blohm et al., 2013). Other approaches rely on the crowd itself for the evaluation of the contributions. However, multiple studies have shown that the design of ratings scales is highly challenging and can fail to produce reliable results (Riedl, Blohm, Leimeister, & Krcmar, 2013). For example, rating scales have been found to frequently face the problems of bimodal distributions or self-selection bias (Ghose, Ipeirotis, & Li, 2012).

In consequence, a number of studies have experimented with text mining and machine learning algorithms to support the evaluation of textual data in crowdsourcing. Walter and Back (2013) use text mining algorithms to cluster ideas submitted to innovation jams in an attempt to provide decision support for expert panels reviewing the contributions. Similarly in the domain of crowdsourced software testing, existing research has used text mining approaches to automatically cluster bug reports and prioritize them for the developers (Feng et al., 2015). In the humanitarian aid sector, Rogstadius et al. (2013) and Barbier et al. (2012) outline the use of text mining algorithms for clustering crowdsourced incident reports and extracting named entities (e.g., locations or names) in order to make the coordination of appropriate responses more efficient.

2.3 Decision Making

The dissertation follows a large stream of related IS literature that views crowdsourcing as a sourcing approach for information (i.e., an “information market”; Bonabeau, 2009), which allows decision makers in organizations to improve their decision making through a form of collective intelligence (e.g., Bonabeau, 2009; Chiu et al., 2014; Geiger & Schader, 2014). Traditional sourcing approaches intend to generate or acquire information based on specialized actors, such as dedicated employees (e.g., Rosenkopf & Nerkar, 2001), intraorganizational units (e.g., Ahuja & Lampert, 2001), or interorganizational partners (e.g., Stuart & Podolny, 1996), to address very specific information needs (cf. Piezunka & Dahlander, 2015). Information is typically delivered and retrieved through standardized and systematic processes in order to reduce complexity in decision making (Constantiou & Kallinikos, 2015). By opening the generation of information to large networks of people, crowdsourcing fundamentally differs from these approaches.

Crowdsourcing seeks to deliberately increase the volume and diversity of the acquired information by gathering and processing data from large networks of individuals (Blohm et al., 2013). Such data may take the form of user-generated content (e.g., ideas, feedback) or automatically tracked data (e.g., click paths, session lengths). Information gained through crowdsourced data can either unfold its value for organizations in a non-emergent or an emergent way (Geiger & Schader, 2014). Value is non-emergent when it can be retrieved directly from individual contributions provided by the crowd. This is the case, for example, for ideas generated during innovation contests (e.g., Leimeister et al., 2009) or feedback in crowdsourced software testing (e.g., Leicht, Blohm, & Leimeister, 2017). Value is emergent, on the other hand, when it can be derived only indirectly from a collection of contributions that need to be transformed or aggregated. This is the case, for example, for crowdsourced votes (e.g., Blohm et al., 2016) or behavioral data to model user preferences (e.g., Brynjolfsson et al., 2015).

According to related research (e.g., Bonabeau, 2009; Chiu et al., 2014), crowdsourced contributions can potentially support all phases of decision making, ranging from an initial gathering or sharing of data for the identification of new opportunities or problems (e.g., mining behavioral data from crowds to uncover trends; Brynjolfsson et al., 2015), to the ideation and conceptualization of innovative products (e.g., designing a new product; Poetz & Schreier, 2012), to the final evaluation of alternatives (e.g., voting for the realization of a particular product design; Blohm et al., 2016). Bonabeau (2009) notes that, with crowdsourcing, “we now have access to more data — sometimes much more data — about customers, employees and other stakeholders so that, in principle, we can gain a more accurate and intimate understanding of our environment. But that's not enough; decisions still need to be made" (p. 45).

2.3.1 The Phase Theorem of Decision Making

Literature on decision making has an extensive background that can be broadly classified into three major streams: research on individual decision making (e.g., Simon, 1960; Todd & Benbasat, 1999), research on group decision making (e.g., Bettenhausen & Murnighan, 1985; De Dreu & West, 2001), and research on organizational decision making (e.g., Cyert & March, 1963; Maitlis & Ozcelik, 2004). This dissertation focuses on individual decision making and follows the first stream of research.

On an individual level, the most widely used conceptualization of decision making is the phase theorem of decision making (Arnott & Pervan, 2014). It describes decision making as a process that comprises three distinct phases: (1) a *processing* of informational cues, (2) an *assessment* of possible courses of actions, and (3) a *commitment* to action (March, 1994; Mintzberg et al., 1976; Simon, 1960). Simon (1960) refers to these phases as the "intelligence", the "design", and the "choice" phase. Mintzberg et al. (1976) termed them the "identification", the "development", and the "selection" phase. In crowdsourcing, such processes may involve an initial screening of user-generated ideas, an evaluation of their projected costs and market potential, and a final choice of implementation (Chiu et al., 2014). While early literature suggested a sequential relationship between these phases, recent studies provide a more fine-grained perspective on decision making and show that decision making processes often comprise multiple data-processing and evaluation sequences that can occur iteratively and recursively (Boonstra, 2003; Frisk, Lindgren, & Mathiassen, 2014; Mintzberg et al., 1976). Boonstra (2003), in particular, provides evidence that decision making is not always predetermined, linear, and explicit but rather exhibits different path configurations or "patterns" that can be explored.

2.3.2 Patterns of Decision Making

Decision making patterns are argued to be determined by an interplay between the structure of the decision problem and the mode of acquiring information to address the problem (Payne, Bettman, & Johnson, 1993; Simon, 1990).

The first important determinant for differences in decision making processes is the structure of the underlying decision problem. Decision problems are argued to exist on a continuum from structured to unstructured (Gorry & Scott Morton, 1971; Shim et al., 2002; Simon, 1960). Decision problems are structured to the extent that they are repetitive and routine so that a definite procedure has been worked out for handling them.

Decision problems are unstructured to the extent that they are non-trivial and novel so that no specific or predefined procedure has been worked out for handling them (Simon, 1960). Gorry and Scott Morton (1971) note that, in structured cases, much of the decision making process can be automated, whereas unstructured cases require adaptive judgement and problem-oriented action by the decision maker. However, studies show that even for unstructured problems, it is possible to observe patterns in decision making. Mintzberg et al. (1976) emphasize that "although the processes used are not predetermined and explicit, there is strong evidence that a basic logic or structure underlies what the decision maker does and that this structure can be described by systematic study of his behavior" (p. 247).

The second important determinant for differences in decision making is the mode of acquiring information. Information helps decision makers "establish options and select adequate courses of actions" (Vandenbosch & Huff, 1997, p. 82). The modes of acquiring information cover a broad spectrum, ranging from general and unintentional to specific and goal-oriented (Aguilar, 1967; Huber, 1991). The former describes "the behavior people exhibit when they browse through information without a particular problem to solve or question to answer", whereas the latter describes behavior people exhibit when they "are looking for something specific" and search particular information (Vandenbosch & Huff, 1997, p. 83). However, there are different perspectives with regard to how access to information and the mode of acquiring it affect decision outcomes. Some scholars (e.g., Anderson, 1983) follow a rational model of choice and argue that decision makers enter decision situations with known objectives that allow them to make optimal choices when given appropriate information sources. Others see decision makers constrained by cognitive limitations (Todd & Benbasat, 1999) and bounded rationalities (Simon, 1979) that impede optimal choices.

2.3.3 Challenges in Crowdsourcing

In crowdsourcing, contributions and data are often collected through IT platforms (Blohm et al., 2018; Doan et al., 2011). These information systems act as an interface between the crowd and the organization and facilitate the sourcing and aggregating of data at a focal point (Geiger & Schader, 2014). Individuals working at this interface (e.g., product owners, test managers) take an important boundary-spanning role as decision makers for the organizations (Tushman, 1977; Tushman & Katz, 1980). They are responsible for processing and interpreting the data to extract or select relevant information for the organization (Geiger & Schader, 2014; Schenk & Guittard, 2011). However, as emphasized by Sharma et al. (2014), information and insights for organizations do not emerge automatically out of raw data. They rather emerge out of active decision making processes by individuals working with crowdsourced data. Thus, much research in recent years has started to examine how such new opportunities to source or prospect data may affect decision making processes (e.g., Bonabeau, 2009; Chiu et al., 2014). Scholars have thus begun to question how the phase theorem of decision making translates to data-driven environments, such as crowdsourcing, and what patterns might emerge in this context (e.g., Abbasi et al., 2016; Constantiou & Kallinikos, 2015; Sharma et al., 2014). On the one hand, decision problems have drastically changed with regard to their structuredness, as both problem and solution spaces for decision makers have risen in quantity and complexity (Barbier et al., 2012). On the other hand, these developments have also paved the way for much more diverse and thorough modes of acquiring information. Decisions can now be informed much more thoroughly by actual data and insights rather than subjective judgement or intuition.

2.4 Decision Support

Decision support is the area of IS research that is concerned with supporting and improving decision making in organizations (Arnott & Pervan, 2014). The use of information technology in the form of *decision support systems* (DSS) is considered essential for this endeavor. On the one hand, DSS have evolved from theoretical studies offering insights on decision making processes (e.g., Mintzberg et al., 1976; Simon, 1960; Todd & Benbasat, 1999). On the other hand, the development of the DSS concept was influenced by technical work (e.g., Ariav & Ginzberg, 1985; Gerrity, 1971; Sprague, 1980) which provided the necessary frameworks and technologies to build and understand systems capable of supporting decision making processes (Shim et al., 2002). Today, there is an array of distinct types of DSS, which differ with regard to their dominant technology components or drivers of decision support. They include data-driven DSS, model-driven DSS, knowledge-driven DSS, document-driven DSS, and communication-driven DSS (Power, 2008). They may support individuals or groups in organizations (Arnott & Pervan, 2014). This dissertation focuses on personal DSS.

2.4.1 Decision Support Systems

As outlined previously, decision making is generally defined as a process comprising three distinct phases: (1) a processing of informational cues, (2) an assessment of possible courses of actions, and (3) a commitment to action (Simon, 1960). In crowdsourcing, such processes may involve an initial screening of user-generated ideas, an evaluation of their projected costs and market potential, and a final choice of implementation (Chiu et al., 2014). However, extant research suggests that cognitive limitations (Todd & Benbasat, 1999) and bounded rationalities (Simon, 1979) constrain decision makers in their assessment of information during such processes. With increasing information load (Eppler & Mengis, 2004), it becomes more difficult for decision makers to identify relevant information (Jacoby, 1977) or to recall prior information and set priorities (Schick, Gordon, & Haka, 1990). Studies also show that their search strategies through data sets become limited and less systematic (Cook, 1993; Swain & Haka, 2000).

Personal decision support systems (DSS) are designed to expand human information-processing capabilities and improve their decision making in such settings of high information load (Todd & Benbasat, 1999). A classic DSS design includes components for "(1) sophisticated database management capabilities with access to internal and external data, information, and knowledge, (2) powerful modeling functions accessed by a model management system, and (3) powerful, yet simple user interface designs that enable interactive queries, reporting, and graphing functions" (Shim et al., 2002, pp. 111–112). These three major subsystems represent the basic foundations which a DSS generally comprises (Sprague, 1980). Traditionally, research has focused on achieving two objectives with these DSS: increasing the efficiency and the effectiveness in decision making (Shim et al., 2002; Todd & Benbasat, 1999; W. Wang & Benbasat, 2009). That is, a DSS should support a decision maker in making higher-quality decisions with less effort. These objectives can be achieved by either automating resource-intensive and standardizable information processing tasks or by defining and ordering the necessary activities for decision making, i.e., structuring the process and providing decisional guidance (Häubl & Trifts, 2000; Silver, 1991). Decisional guidance in DSSs can take a purely informative form that includes pertinent information but no recommendations or a suggestive form with clear recommendations for the decision maker (Silver, 1991). Furthermore, it is possible to distinguish between predefined guidance where the designer of a DSS prepares the recommendations, dynamic guidance with adaptive mechanisms that let the system learn as it is used, and participative guidance where the decision maker defines the preferences (Parikh, Fazlollahi, & Verma, 2001; Silver, 1991). Inherently, when decision makers use a DSS, their decision making process is restricted to the processes or strategies supported by the DSS (Silver, 1988; W. Wang & Benbasat, 2009). Well-designed DSS have been found to help decision makers in analyzing problems in greater depth and, ultimately, making effective decisions in a more efficient fashion (Häubl & Trifts, 2000; Hoch & Schkade, 1996). By integrating information systems and decision making processes, organizations have experienced substantial improvements in performance over the past decades (Arnott & Pervan, 2012).

2.4.2 Text Mining and Machine Learning

While traditional DSS have mostly focused on structured data, DSS research has recently witnessed a "move toward dealing with massive collections of relatively unstructured data such as audio, video, clickstream, and text" (Holsapple, Lee-Post, & Pakath, 2014, p. 131). Thus, text mining and machine learning are gaining in importance for decision support. DSS based on text mining and machine learning technologies are often referred to as *intelligent DSS* (Arnott & Pervan, 2014).

Text mining denotes the process of extracting useful information from unstructured, textual data through the exploration of meaningful patterns (Feldman & Sanger, 2007). These patterns are extracted by combining algorithms and methods from the fields of natural language processing, statistics, and machine learning (Tan, 1999). The standard procedure for text mining consists of two basic steps. First, the unstructured, user-generated text has to be preprocessed into a format that is compatible for machines (e.g., through tokenization and stemming). Afterwards, complementary machine learning techniques provide the means to structure the data, recognize patterns, or extract useful information. A number of supervised and unsupervised approaches are available for this task. Supervised approaches (e.g., classification; see Sebastiani, 2002) provide the means to assign contributions to predefined classes while unsupervised approaches (e.g., clustering; see Jain, 2010) are capable of automatically finding relationships and structures in large sets of contributions without predefined classes.

A number of studies have already demonstrated the potential of text mining and machine learning algorithms for decision support in crowdsourcing. Walter and Back (2013) used text mining in combination with clustering algorithms to support decision makers in selecting novel ideas from more than 40’000 contributions. Similarly, Nagar et al. (2016) developed and tested models based on a large citizen-science platform to predict expert decisions about the submissions to accelerate the review process and reduce manual efforts. Feng et al. (2015) applied clustering algorithms in crowdsourced software testing to support developers in prioritizing test reports during defect management. Barbier et al. (2012) employed text mining to automatically extract named entities (e.g., locations) from crowdsourced incident reports to assist organizations in distributing relief supplies during natural disasters. While such instantiations demonstrate the technical capabilities of text mining and machine learning algorithms in crowdsourcing, there is a lack of prescriptive design knowledge to guide researchers and practitioners in systematically implementing them for decision support on crowdsourcing platforms.

3 Decision Making in Crowdsourcing

This chapter addresses the first research question of the dissertation and is concerned with understanding decision making in crowdsourcing. The chapter presents the results of an explanatory interview study[2] that investigated the patterns of decision making that emerge in crowdsourcing. This forms the foundation for all subsequent sections of the dissertation that delve deeper into potential mechanisms to support decision making in crowdsourcing (cf. chapters 4–6). In this chapter, section 3.1 first explains the motivation and objectives of the interview study in more detail. Afterwards, section 3.2 outlines the methodology of the study and provides a detailed description of the data collection and analysis. Third, section 3.3 reveals the results of the study and discuss their implications for both theory and practice. Finally, sections 3.4 and 3.5 conclude the chapter by acknowledging the study's limitations and offering an outlook for future research.

3.1 The Need to Study Decision Making in Crowdsourcing

Over the past years, much attention has been paid in both research and practice to the value that organizations could create through crowdsourced data (Abbasi et al., 2016; Barbier et al., 2012; Chen et al., 2012; Sharma et al., 2014). So far, decision makers in organizations mostly worked with enterprise-specific data collected through standardized and purposeful processes that address specific information needs (Constantiou & Kallinikos, 2015). Recently, with increased capabilities to collect large-scale, crowdsourced data, they have gained access to much more diverse and extensive data sources that allow them to uncover new behavioral trends (e.g., Brynjolfsson et al., 2015), derive insights about latent user preferences (e.g., Blohm et al., 2016), or design innovative products (e.g., Poetz & Schreier, 2012). Literature suggests that these developments enable a shift towards more open, "data-driven" decision making in organizations that draws upon actual information about people's elicited or observable behavior, opinions, and choices (Abbasi et al., 2016; Sharma et al., 2014). Bonabeau (2009), in particular, predicts that the use of crowdsourced data will mark "a paradigm shift in the way companies make decisions" (p. 46). Indeed, large companies are already beginning to embrace this shift. Microsoft, for example, used crowdsourced data from more than 48'000 Skype users in combination with EI Analytics, an analytics dashboard, to support decisions about investments in network and server infrastructure, emphasizing that such data

[2] Sections 3.1 to 3.5 of this chapter provide new data and insights on patterns of decision making in crowdsourcing for the dissertation. A modified version of the content is under review in: Rhyn, M. & Blohm, I. (2019). Patterns of Data-Driven Decision-Making: How Decision-Makers Leverage Crowdsourced Data. Proceedings of the 40th International Conference on Information Systems (ICIS), 1-16. Munich, Germany: AIS.

and tools have "a positive impact on the development process and decision making" (Musson et al., 2013, p. 43).

While research already made great strides in recent years to develop the technical foundations for processing large-scale data from crowds (Chen et al., 2012), decision making in organizations that builds upon these novel capabilities to source and analyze data on people's actual behavior, opinions, or choices is not well understood (Sharma et al., 2014). Decision making describes the sequences of data-processing activities and evaluation patterns, by which focal actors analyze data and choose courses of actions to solve an organizational problem (e.g., develop a new product based on ideas or behavioral data from a crowd). Marchand and Pepper (2013) note that scholars and practitioners have focused too much on technical facets of user-generated data and related analytics technologies and not enough on the people who work with them. They emphasize that it is crucial to understand "how people perceive problems, use information, and analyze data in developing solutions, ideas, and knowledge" (Marchand & Peppard, 2013, p. 109). Especially for analytics technologies, the logic behind many investments in them is that "giving managers more high-quality information more rapidly will improve their decisions and help them solve problems and gain valuable insights" (Marchand & Peppard, 2013, p. 106). However, much research suggests that investments in new technologies and approaches, such as crowdsourcing, provide little value per se if they are not well integrated with decision making processes in organizations (Brynjolfsson & Hitt, 1998; Willcocks & Lester, 1999). Thus, it is crucial to gain a better understanding of the structure and patterns of decision making that emerge when decision makers have access to large amounts of crowdsourced data (Abbasi et al., 2016). This would make it possible to better understand what decision making patterns may occur in different types of decision situations and how information systems can be adapted to provide decision support.

To address this gap, the objective of this study is to analyze and systematize decision making patterns in crowdsourcing. We answer the following research question: What decision making patterns emerge when decision makers have access to large-scale, crowdsourced data? Taking a process perspective (cf. Boonstra, 2003; Mintzberg et al., 1976; Simon, 1960), we examine decision making in crowdsourcing as a sequence of data-processing activities and evaluation patterns, by which focal actors process crowdsourced data and choose courses of actions to solve an organizational problem. To study the characteristics of such decision making processes and identify patterns, we conducted interviews with decision makers from 10 multinational corporations that regularly work with crowdsourced data. We use a multi-staged coding approach based on

Gioia et al. (2013) in combination with a temporal bracketing strategy (Langley, 1999) to conceptualize their decision making processes and identify decision making patterns.

With this study, we contribute to both research on crowdsourcing and research on decision making. For research on crowdsourcing, we outline the structure of decision making processes that emerge when decision makers have access to large amounts of crowdsourced data. We extend existing literature in this field, which has already focused on the technical foundations for processing and analyzing crowdsourced data (e.g., Barbier et al., 2012; Chen et al., 2012), by providing a better understanding on how decision makers leverage such data and derive decisions based on newly gained insights. More importantly, however, we show that decision making processes in crowdsourcing do not always represent a predetermined sequence of phases but that they rather follow four distinct patterns. For research on decision making, we answer the calls from various scholars to examine how decision making processes may change in data-driven environments (e.g., Abbasi et al., 2016; Sharma et al., 2014). Our findings from crowdsourcing show potential limitations of the traditional phase theorem of decision making in data-driven environments. We provide an integrated perspective on how the structure of the decision problems (Shim et al., 2002; Simon, 1960) and modes of acquiring information (Aguilar, 1967; Huber, 1991; Vandenbosch & Huff, 1997) evoke and affect patterns instead of a uniform, sequential process. Finally, for practice, the patterns identified in this study may help to better design information systems that provide decision support around crowdsourced data. We show that there is no "one-size-fits-all"-solution to information systems design and discuss how decision support mechanisms need to be adapted to different patterns in decision making.

3.2 Interview Study

The objective of our study is to analyze and systematize decision making patterns in crowdsourcing. We view decision making patterns as processes consisting of data-processing and evaluation phases, by which focal actors choose adequate courses of actions to solve an organizational problem. We aim to examine how the structure of the decision problem and the mode of acquiring and processing information evoke different patterns of decision making. To achieve this objective, we use a qualitative research approach with semi-structured interviews for our study. We follow Mintzberg et al. (1976) and aim at "eliciting the verbalizations of decision makers' thought processes", which can then be "analyzed to develop simulations of their decision processes" (p. 247). A qualitative research approach allows data to be collected in natural settings and ultimately offers rich and holistic insights through local groundedness (Miles et al., 2014). It is especially well-suited to capture events, processes, or structures experienced by decision makers and thus represents an adequate way of analyzing the characteristics and patterns of their decision making (Miles et al., 2014). We conducted semi-structured interviews with 16 decision makers across 10 organizations that regularly engage in crowdsourcing. For the analysis of the interviews, we employed a multi-staged, inductive coding approach based on Gioia et al. (2013) and a temporal bracketing strategy proposed by Langley (1999).

3.2.1 Data Collection

For the purpose of this study, we use semi-structured interviews as our primary source of data (Myers & Newman, 2007). With this type of interview, it is possible to gain insights from a sample of decision makers who frequently engage in crowdsourcing and study their decision making processes in detail. The semi-structured format of the interviews ensures that we collect comparable information from all decision makers but still allows us to engage in further enquiries as the discussion unfolds.

For the selection of the interview partners, we followed a purposive sampling strategy, which is the most commonly used form of non-probabilistic sampling (Guest, Bunce, & Johnson, 2006). "Purposive sampling strategies are non-random ways of ensuring that particular categories of cases within a sampling universe are represented in the final sample of a project" (Robinson, 2014, p. 32). In our study, we aimed to ensure that both structured and unstructured decision problems in crowdsourcing are represented, as related literature suggests that the structure of the decision problem greatly affects decision making processes (Payne et al., 1993). To cover structured decision problems, we interviewed decisions makers that use crowdsourced data in technical contexts with well-defined evaluation criteria and decisions (e.g., verifying and accepting defects in software testing). To cover unstructured decision problems, we interviewed decision makers that use crowdsourcing in creative contexts that typically have no clear solution but require adaptive judgement and choice for the final decision (e.g., identifying promising ideas for product development). Given that crowdsourcing is mostly organized in campaigns or projects, the responsible decision makers are often product owners or project managers in the organizations. They are in charge of defining the problem, specifying an appropriate crowd, and collecting the data. They are also the primary decision makers when it comes to retrieving the data, processing them, and making a decision to incorporate changes in the software or start a project based on an ideation campaign. To avoid biases, we interviewed decision makers from different industries (8), different organizations (10) and departments (15), and with varying experience on different crowdsourcing projects.

Regarding the sample size, Guest et al. (2006) found that basic elements for meta-themes typically become present within the first six interviews of a study while saturation usually occurs within twelve interviews. Similarly, Bertaux (1981) recommends a minimum sample size of fifteen. Kuzel (1992) suggests a sample size of six to eight interviews for homogeneous sources and twelve to twenty interviews for more heterogeneous sources. We follow Guest et al. (2006) and refer to saturation "as the point in data collection and analysis when new information produces little or no change to the codebook" (p. 65). This means that the interviews cease to reveal fundamentally new or different insights for the development of properties of a given category (e.g., phases in decision making), so that we become "empirically confident that a category is saturated" (Glaser & Strauss, 1967, p. 65). Miles et al. (2014) emphasize that sampling often has an "iterative or 'rolling' quality, working in progressive waves as the study progresses" (p. 33). Thus, we followed Miles et al. (2014) and conducted the interviews iteratively from September 2016 to March 2018 until information provided by the decision makers became repetitive and started to indicate an onset of saturation. We concluded the interview phase by 16 interviews. Table 2 lists the interview partners.

No.	*Position*	*Firm Type*	*Projects*	*Problem*[1]
1	Test Manager	Bank	18 projects	S
2	Senior Credit Risk Officer	Bank	1 project	S
3	Test Manager	Bank	12 projects	S
4	Test Manager	Bank	11 projects	S
5	Project Manager	Research	7 projects	S/U
6	Test Manager	Bank	5 projects	S/U
7	Chief Executive Officer	Intermediary	> 100 projects	S/U
8	Application Manager	Insurance	4 projects	S/U
9	Test Manager	Insurance	4 projects	S/U
10	Project Manager	Intermediary	> 100 projects	S/U
11	Innovation Manager	IT Service	1 project	S/U
12	Community Manager	Analytics	> 100 projects	S/U
13	Project Leader	Retail	67 projects	S/U
14	Consultant	Intermediary	10 projects	S/U
15	QA Manager	Insurance	10 projects	S/U
16	Innovation Manager	Logistics	20 projects	U

[1] Type of Decision Problem: U = Unstructured; S = Structured

Table 2: Interview Partners

Source: Own Illustration

The questions in our interview guideline aimed to uncover patterns of decision making that individuals exhibit when working with crowdsourced data. The structure of the interviews followed three parts: In the first part, we asked the decision makers to introduce themselves, explain their function and experience in the organization, and describe typical decision problems that they face in their organizations. This part aimed at gaining an overview of the types of decision problems and their structuredness. In the second part, we asked the decision makers to outline the sequences of data-processing activities and evaluation steps, by which they source, analyze, and use crowdsourced data to address decision problems. As outlined by Mintzberg et al. (1976), this part aimed at "eliciting the verbalizations of decision makers' thought processes", which can then be "analyzed to develop simulations of their decision processes" (p. 247). In the third part of the interviews, we were interested in the type information systems used during this process and their assessment of the efficiency and effectiveness of decision making based on crowdsourced data. In the end, we also gave our interview partners the possibility to further explain or discuss aspects that they deem important for their decision making but were not explicitly asked by us. The duration of the interviews ranged from 30-90 minutes. We recorded the interviews and took notes during the sessions.

3.2.2 Data Analysis

To systematically extract patterns of decision making and analyze how they relate to the structure of the underlying decision problems and the mode of acquiring and processing information, we coded the interviews. Codes "are labels that assign symbolic meaning to the descriptive or inferential information compiled during a study" (Miles et al., 2014, p. 71). They can be used to retrieve and categorize chunks of information in interview transcripts to cluster segments that relate to a particular construct or theme (Miles et al., 2014). In our case, the codes serve to structure the verbalizations of the decision making processes from the interviews. That is, we use the codes to derive distinct phases of decision making in crowdsourcing as described by decision makers and analyze different patterns based on how these phases are aligned.

We followed the inductive data analysis and coding approach proposed by Gioia et al. (2013), which is well-established in related literature on decision making and process research (e.g., Langley, Smallman, Tsoukas, & Van De Ven, 2013; Smith, 2014). This approach is based on a multi-staged coding scheme with first-order codes, second-order concepts, and aggregated dimensions (Gioia et al., 2013).

1st Order Codes (Examples)	*2nd Order Codes (Examples)*	*Phases*
Defining the problem Outlining the task Determining data formats Specifying labels	Identification of problem Definition of required data Collection of data	Sourcing
Assessing fit to task Removing duplicates Removing low-quality reports Verifying labels Adding context	Verification of information Omission of information Revision of information Extension of information	Validating
Clustering similar contributions Reducing clusters to their core Selecting unique contributions Summarizing results	Aggregation of information Integration of information Selection of information	Consolidating
Sorting the contributions Discussing the content Predicting the impact Assessing the severity Assessing popularity	Evaluation of feasibility Prediction of impact Estimation of efforts Determination of importance	Evaluating
Accepting contributions Starting a project Passing results to department Fixing an issue	Choice of an alternative Assignment of tasks Allocation of resources	Choosing

Table 3: Extract of Coding Scheme

Source: Own Illustration based on Gioia et al. (2013)

First-order codes represent informant-centric terms that emerge during the interviews. For these codes, we adhered to words that were used by the decision makers during the interviews to describe the processes and activities when engaging in crowdsourcing. Based on similarities and differences in these codes, it is possible to derive second-order concepts that represent germane themes and categories described during the interviews (Gioia et al., 2013). Deriving the second-order concepts was an iterative process of "constant comparison" (Corbin & Strauss, 1990). We developed potential concepts and dismissed, changed, or retained them based on comparisons across the interviews to achieve a coherent synthesis. To increase confidence in our analysis, we discussed preliminary results and variations and gave our raw data to independent students for analysis (cf. Lehrig, Krancher, & Dibbern, 2017). We adapted the concepts whenever suitable or necessary. We repeated this process to achieve aggregated phases in decision making in crowdsourcing. Table 3 provides an extract of the coding scheme.

As outlined earlier, distinct phases can occur iteratively and recursively during a decision making process and form "patterns" (Boonstra, 2003; Frisk et al., 2014; Mintzberg et al., 1976). To examine such patterns based on the codes, we followed a temporal bracketing strategy proposed by Langley (1999). It represents a standard approach for analyzing process data and is especially well-suited for an "open-ended inductive approach that most researchers use in process research" (Langley et al., 2013, p. 693). At its core, temporal bracketing refers to the "decomposition of data into successive adjacent periods [which] enables the explicit examination of how actions of one period lead to changes in the context that will affect action in the subsequent periods" (Langley, 1999, p. 703). That is, based on the codes (aggregated phases), we reconstructed the decision making processes, by which decision makers typically source and process crowdsourced data to derive decisions for the underlying projects. These processes can then be grouped based on the number of transitions between phases and similarities in their alignment to describe the processes as "evolving patterns" (Langley, 1999). The results of the temporal bracketing of the codes are depicted in Figure 2.

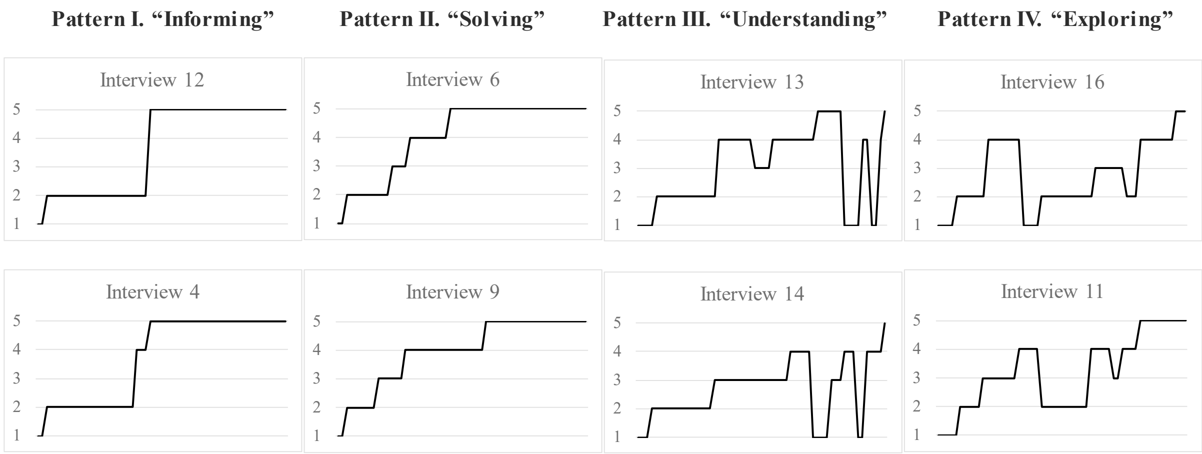
3.3 Patterns of Decision Making in Crowdsourcing

The interviews provide interesting insights into the characteristics of decision making in crowdsourcing. They revealed five distinct phases that typically occur during decision making in crowdsourcing. These five phases represent germane episodes of acquiring and processing newly gained data through crowdsourcing to derive decisions for the underlying projects. They provide a data-centric perspective on decision making. The sourcing phase comprises the acquisition of data to either identify new or address existing problems and opportunities. In crowdsourcing, decision makers may source new data at potentially all stages of decision making. In the validation phase, decision makers assess the appropriateness of the data to address the underlying decision problem. This phase is essential for decision makers working with crowdsourced data, as they can no longer rely on the trustworthiness of the sources nor the quality of their input. The consolidation phase describes the extraction of meaningful information from crowdsourced data. It revolves around aggregating data, integrating data, and selecting data provided by the crowd to derive valuable insights. In the evaluation phase, decision makers determine the value of the insights extracted from the crowdsourced data. That is, they have to estimate the required efforts for implementation (e.g., in terms of costs and time), assess the feasibility, and predict the impact for their organizations. The choice phase describes the commitment of resources to realize a project or implement changes in the organization.

However, we find that decision making processes do not always adhere to a predetermined sequence of these phases. They rather follow different patterns that range from sequential and goal-oriented to dynamic and data-driven. The patterns depend upon the structure of the decision problem and the mode of acquiring and processing information. In the following sections, we first describe the decision making patterns that emerged during the interviews based on how the previously outlined phases are aligned and how they typically (re-)occur. Second, we explain how the patterns relate to the problem structure and the mode of acquiring information described by our interview partners. Finally, we consider the efficiency and effectiveness of the patterns based on the interviews.

3.3.1 Description of the Patterns

Figure 2 below depicts the results of the temporal bracketing of the codes and illustrates exemplary sequences for each decision making pattern. Each step represents a phase of sourcing, validating, consolidating, or evaluating data, and choosing adequate courses of actions for the underlying project as described by the decision makers during the interviews. The patterns reveal great differences with regard to how many times these phases occur and how they are aligned during decision making.



Note: 1 = Sourcing; 2 = Validating; 3 = Consolidating; 4 = Evaluating; 5 = Choosing

Figure 2: Overview of Decision Making Patterns with Exemplary Sequences

Source: Own Illustration

Informing: The first pattern of decision making reconstructed from the interviews takes the form of a sequential process that is characterized by only few transitions between phases of gathering data and making a choice. That is, data are sourced once and then processed in a standardized and goal-oriented way to address specific information needs of the decision maker. We termed this pattern "informing". Interview partner 12, a manager for an analytics provider, described an exemplary case for such a pattern in retail audits for monitoring stores. She explained that she uses crowdsourcing to collect clearly specified data on mobile devices of customers for point of sales benchmarks (e.g., geo data). Crowdsourced geo data are collected, validated, and then displayed on analytics dashboards for managers to decide, for example, whether the positioning of certain products on shelves need to be changed. As both the problem and the required data are known, the process is standardized and aims for high efficiency in decision making. A similar process was described by interview partner 4, a test manager for a retail bank. He uses crowdsourcing for standardized regression testing in software development and explained: "*The most important aspect for me at the moment is to have the test cases that were executed by the crowd at the status 'okay'. The data are validated and forwarded during the actual crowdtesting session. Our test managers just synchronize the defect by the crowd from TFS [Microsoft Team Foundation Server] to HP QC [HP Quality Center]*."

Solving: The second decision making pattern also takes the form of a sequential process but progresses stepwise. In this case, data are sourced and evaluated in a more gradual and analytical manner to address problems that are less structured and require thorough examination. We termed this pattern "solving". Interview partner 6, a test manager for a retail bank, provided an example for such a pattern. He explained that he uses crowdsourced data to identify and fix defects in a large-scale banking software. He notes that "*understanding such defects from a technical point of view is very difficult*." For this purpose, he systematically sources, validates, consolidates, and evaluates defect reports and decides whether and how to fix them. He describes the process as a "*thorough analysis*" that is very "*tedious and time-consuming*". However, he also emphasizes that this allows him to get "*an extremely good picture of how well the application works*." Interview partner 9 referred to such decision making processes as a "*multi-step analysis*".

Understanding: The third decision making pattern is the first to deviate from the sequential structure. Early stages of the decision making process aim at incrementally developing potential options to address a problem. At later stages, however, we see decision makers iteratively source and compare new data to get a better understanding of these options and make adequate decisions. We termed this pattern "understanding". Interview partner 13 offered an exemplary description for such a pattern. As a project leader for a large retail company, he is responsible for developing new products. At early stages, he uses crowdsourcing to gather ideas on different types of products that could be included. At later stages, he typically has to source new data to better estimate market impacts and customer preferences of ideas (e.g., through votes) that he did not previously know. He compares the decision making process to a "*stage-gate process*". These phases are repeated until the available options and final decisions are backed by enough data.

Exploring: The fourth pattern of decision making takes the form of a highly dynamic process of sourcing and scanning data. It exhibits a rather undirected structure that revolves around iteratively probing data to uncover new solution spaces that were not hitherto known. We termed this pattern "exploring". Interview partner 16, an innovation manager in a large logistics group, offered an exemplary description for such a pattern. He explained: "*We use crowdsourcing as part of our search strategy. It represents a trend monitor in which we identify trends*". Data are sourced in an open and unrestricted manner to find potential "*search streams*" that are then further investigated and enriched with novel data as promising "*business cases*" unfold. In this pattern, crowdsourced data act as guidance in exploratory decision making process.

Pattern	*Description*
I. Informing	"Informing" describes a pattern of focused and directed search for specific information through crowdsourced data. The goal is to efficiently process crowdsourced data and quickly inform a decision that has clear requirements.
II. Solving	"Solving" describes a pattern of stepwise analysis of information in crowdsourced data. The focus lies on gradually accumulating required information and systematically approaching a decision.
III. Understanding	"Understanding" describes a pattern of iterative comparison of information gained through crowdsourced data. The focus lies on recursively sourcing data to evaluate different options for a decision.
IV. Exploring	"Exploring" describes a pattern of open and dynamic scanning of information in crowdsourced data. The focus lies on probing data to uncover new and hitherto unknown solution spaces for decisions.

Table 4: Summary of Decision Making Patterns

Source: Own Illustration

3.3.2 The Problem Structure and the Modes of Acquiring Information

The interviews not only reveal different patterns of decision making but also offer insights with regard to why and in which contexts these patterns occur. We find the four decision making patterns to be a consequence of how the structure of the decision problem and the decision maker's mode of acquiring information interact. Interview partner 14, a consultant for a crowdsourcing intermediary, explained: "*Decision problems are very diverse. Some organizations have a clear problem, such as choosing a name for a new product, and ask the crowd for their preference. Others face much more complex problems and need to understand, for example, media consumption of their customers*." However, not all decision makers address these different types of problems in the same way. Some decision makers use crowdsourced data for a goal-oriented search of specific information that addresses their problem. Others scan crowdsourced data more openly to develop and select potential options. This explains differences in how the phases of decision making are aligned and how often they (re-)occur (see Figure 2).

Decision making patterns in crowdsourcing that follow a rather sequential order (Pattern I and II) are typically driven by a decision maker's goal-oriented search for information. We find them to occur when decision makers put strong emphasis on their own expertise and experience and argue to have a good understanding on how to source, process, and analyze adequate data. Interview partner 11, for example, described: "*I examined the data myself. […] I know the business pretty well by now and can decide by myself whether an option is promising or not*." Thus, he looks for specific information that addresses predefined requirements. In such cases, crowdsourced data is primarily used to efficiently "inform" decisions or "solve" problems. "Informing" decisions (Pattern I) through a goal-oriented search usually occurs when decision makers face structured decision problems (e.g., in technical contexts). In these cases, decision makers have often worked out very efficient routines to process and evaluate crowdsourced data. For example, interview partner 12, a manager for an analytics provider, explained: "*Some projects start in the morning and can be finished by noon. In these cases, the validation of the data is actually the most time-consuming part, because they are all collected and retrieved at once*." Similarly, interview partner 11 notes that he often uses crowdsourced data in a standardized way to reassure his decision whether to invest in the development of a product or not. When a goal-oriented search is applied to a more complex, unstructured decision problem, the decision making pattern represents much of a "solving" process (Pattern II). In these cases, our interview partners reported to fragment the decision problems. That is, they analyze the crowdsourced data stepwise, accumulate information, and develop decisions progressively.

Decision making patterns in crowdsourcing that are rather recursive and iterative in nature (Pattern III and IV) are typically related to decision makers employing a more dynamic and open scanning for information. Such patterns often occur when decision makers put less emphasis on their own expertise and experience. In these cases, it is more likely that data – not predefined objectives – drive the decision making process and guide decision makers. That is, decision makers rely more thoroughly on data to openly "explore" new options or better "understand" them. "Exploring" new options through an open and dynamic scanning for information (Pattern IV) occurs when decision makers face highly unstructured decision problems. Interview partner 12, a manager for an analytics provider, explained: "*There are projects that are extremely complex and progress very slow. Even the definition of project is time-consuming*." Similarly, interview partner 6, argued: "*The most time-consuming part here is understanding the problem*." In these cases, decision makers face novel and non-trivial problems for which they have not yet found an efficient routine. Decisions develop iteratively and require decision makers to source new data multiple times to address novel information needs at particular stages of their decision making. If an open scanning for information is applied to a more structured decision problem, the process typically aims to better "understand" options by recursively sourcing and comparing crowdsourcing data (Pattern III).

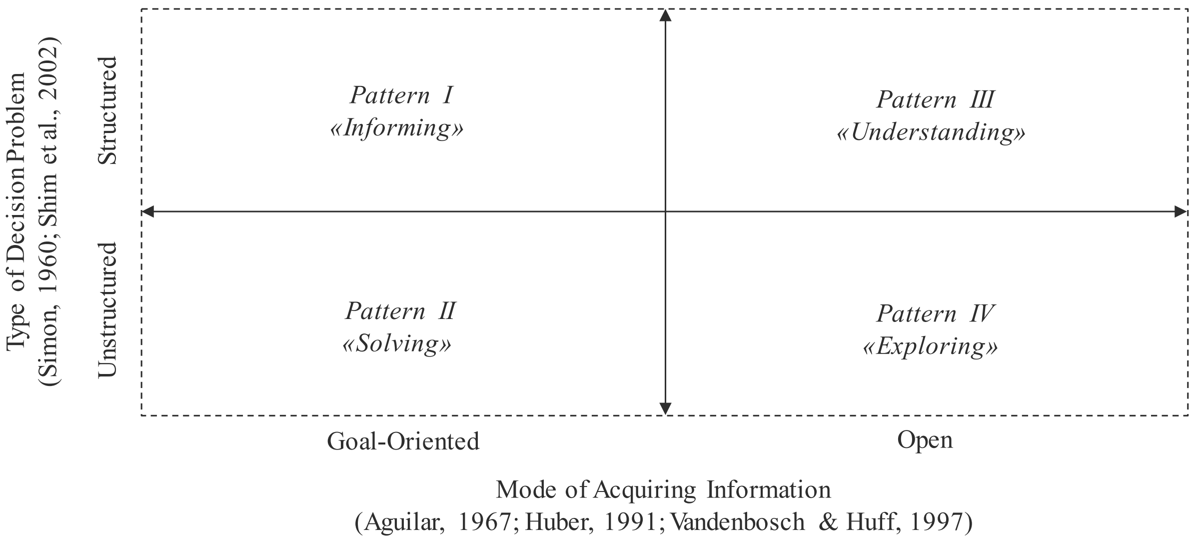


Figure 3: Framework for Decision Making Patterns in Crowdsourcing

Source: Own Illustration

3.3.3 Efficiency and Effectiveness of the Decision Making Patterns

Given that there are distinct patterns of decision making that depend upon the structure of the decision problem and the mode of acquiring information, there are also differences with regard to how the phases during these patterns relate to the efficiency and effectiveness of decision making and how information systems may provide adequate support for the decision makers.

Decision makers that employ a goal-oriented search to "inform" their decisions or "solve" a problem typically view the "validation" and "consolidation" phase as critical for their decision making. The decision makers reported that they often validate and consolidate data manually, requiring substantial amounts of time and effort. Project leader 13 emphasized: "*There is a lot of manual effort involved because most crowdsourced data are highly unstructured*." Test manager 6 underlined that this requires "*way too much time*". Thus, across the interviews, multiple decision makers suggested that there is vast potential for automated data-processing mechanisms to support decision making patterns I ("informing") and II ("solving"), as both are goal-oriented and primarily target a high efficiency in decision making. Surprisingly, however, we found that decision makers are still hesitant to rely on such technologies. Test manager 9 explained: "*Currently, I would not blindly trust automated reports. I always want to know what is going on. I want to have enough control to be able to intervene*." Similarly, innovation manager 11 stated: "*At a certain point, it is not possible to manually process and evaluate all data. However, when it comes to automation, I'm always concerned about missing high-potential ideas.*" This also resonates with earlier findings that pattern I and II often occur when decision makers put strong emphasis on their own expertise and believe to have a good understanding on how to source, process, and analyze data.

Decision makers that employ a more open and dynamic scanning of crowdsourced data to "explore" new options or better "understand" them are often primarily concerned with the effectiveness of their decision making. Project leader 13, for example, explained: "*In retail, there is a 'one-in-one-out' rule. It is only possible to introduce a new product in a store if another product is removed at the same time*." He explains that crowdsourced data are used to make the "*right*" decisions with regard to their product assortment in stores. The most important aspect for him is the outcome: "*We have conducted 67 crowdsourcing campaigns so far where we developed products – from the initial problem definition to the final product launch. These products yielded a gross revenue of 100 Mio. Euros*". Hence, our interviews suggest that in pattern III ("understanding") and IV ("exploring"), the "sourcing" and "evaluation" phase are critical, as decision makers rely more on data and their decisions become more data-driven. The most important threats to such data-driven patterns revealed during the interviews are subjective judgement and biases by the decision makers (i.e., how data are interpreted). Sometimes, decision makers even reported to deliberately dismiss data. Innovation manager 11 said: "*We received a lot of votes from the crowd but did not really take these preferences into account for our final decision. […] We were able to do this based on our experience*." Project leader 13 admitted: "*We try to be as neutral and objective as* *possible, but there is definitely a lot of subjectivity involved in the process*". Multiple decision makers argued that there is high potential for information systems to support decision making patterns III and IV by offering pattern recognition and visualization capabilities that allow for an easier analysis of crowdsourced data during the "evaluation" phase. Project leader 13 said: "*It would be very interesting to have a pattern recognition system of some sort, to extract patterns that are otherwise not easily detectable. This would surely help to identify trends in data or classify customer segments. Artificial intelligence or machine learning are promising in this regard – especially when crowdsourced data are combined with internal data, such as sales data*."

3.4 Discussion of Decision Making Patterns in Crowdsourcing

Taken together, the results from our interviews offer a number of interesting insights into the characteristics of decision making in crowdsourcing. Existing literature on decision making is mostly grounded on the traditional phase theorem (Mintzberg et al., 1976; Simon, 1960) and builds upon the basic tenet that decisions are made by systematically gathering new information, assessing potential courses of actions based on this information, and then committing to one or more alternatives. In crowdsourcing, with new opportunities to source and analyze data, we found the characteristics of decision making processes to change.

Based on the interviews, we identified five distinct phases that occur when decision makers work with crowdsourced data. These five phases represent germane episodes of acquiring and processing data in order to make adequate decisions for the underlying projects. More importantly, however, it became clear during our analysis that decision making processes in crowdsourcing do not always represent a predetermined sequence of these phases. In related studies, Mintzberg et al. (1976) already noted that while it is possible to delineate distinct phases in decision making processes, such phases do not necessarily follow "a simple sequential relationship" (p. 250). Similarly, Boonstra (2003) argued that decision making processes "are not always predetermined, linear and explicit" and that it is instead possible to "identify general patterns" in decision making (p. 197). Our findings extend this perspective. In crowdsourcing, where decision makers have the opportunity to freely source and examine user-generated data, we find four different patterns of decision making to emerge that range from sequential and goal-oriented to dynamic and data-driven.

Decision makers that exhibit sequential patterns typically use crowdsourced data to efficiently "inform" decisions or gradually "solve" decision problems. These patterns resonate strongly with the traditional perspective on decision making, where decisions are based on data collected through systematic and purposeful processes that address specific information needs of the decision makers (Constantiou & Kallinikos, 2015). We find these patterns to still occur in crowdsourcing when decision makers employ a goal-oriented search and have a good understanding of how to source and analyze data. This allows them to develop and follow predefined routines. When facing structured decision problems, these routines are characterized by only few transitions between phases of sourcing data and making a choice, since the problem and required data are clear. As the decision problems become more unstructured, we found decision makers to fragment the decision problems and develop solutions stepwise by systematically sourcing, validating, consolidating, and evaluating data before making a choice. Mintzberg et al. (1976) describe similar behavior and note that, to reduce complexity, decision makers deal with unstructured problems by factoring them into "familiar, structurable elements" and following "interchangeable sets of procedures or routines" (i.e., phases).

Decision makers that exhibit more recursive and iterative patterns use crowdsourcing to "explore" new options for decision problems or better "understand" them. These patterns resonate strongly with a more recent perception of decision making and reflect Abbasi et al.'s (2016) notion of a "data-driven decision making process". In these patterns, data act as principal drivers or guidance for the decision making process and not clearly defined objectives by the decision maker. That is, decision makers rely more thoroughly on data to make a decision and employ a dynamic and open scanning for information. For structured decision problems, our interview partners described such data-driven decision making processes as "stage-gate processes", in which decisions are iteratively developed and backed by data. For unstructured decision problems, decision making processes become more explorative and increasingly revolve around probing data to uncover new, hitherto unknown solution spaces. Lycett (2013) describe such processes as "information technology driven sense-making processes", which revolve around organizing data to identify and regularize patterns into plausible explanations and actionable decisions.

3.4.1 Implications for Theory

From a theoretical perspective, we contribute novel insights to both research on crowdsourcing and research on decision making. In research on crowdsourcing, scholars have been increasingly interested in the value that organizations could create through crowdsourced data (e.g., Barbier et al., 2012; Bonabeau, 2009; Chiu et al., 2014). While related studies already made great strides to develop the technical foundations for processing large-scale data from crowds (e.g., Barbier et al., 2012; Chen et al., 2012), the structure and patterns of decision making that emerge when decision makers have access to crowdsourced data were mostly unclear (Abbasi et al., 2016). The findings of our study address this gap in two ways. First, we reveal five distinct phases that form decision making processes in crowdsourcing. With this data-centric perspective on decision making, it possible to better understand how decision makers in crowdsourcing source, validate, consolidate, and evaluate data from a crowd to choose adequate courses of actions for solving an organizational problem. Second, and more importantly, we show that decision making processes in crowdsourcing do not always represent a predetermined sequence of these phases as assumed by traditional decision making models used in crowdsourcing literature (e.g., Chiu et al., 2014). Instead, we identify four common patterns of decision making in crowdsourcing that range from very structured and goal-oriented to highly dynamic and data-driven. They systematize how the decision making phases are typically aligned.

Second, for research on decision making, we answer the calls from various scholars to examine how decision making processes may change in data-driven environments (e.g., Abbasi et al., 2016; Sharma et al., 2014). Existing literature in this field is mostly grounded on the traditional phase theorem (Mintzberg et al., 1976; Simon, 1960) and builds upon the basic tenet that decisions are made by systematically gathering new information, assessing potential courses of actions based on this information, and then committing to one or more alternatives (see Arnott & Pervan, 2014). However, in data-driven environments, such as crowdsourcing, decision problems have drastically changed with regard to their structure. Furthermore, decision makers now have access to much more diverse sources and more thorough means to acquire information. Thus, studies have emphasized that “we need to understand if and how we should revise existing decision making models” (Abbasi et al., 2016, p. 11). Our findings from crowdsourcing show the limitations of the traditional phase theorem for studying decision making in more data-driven environments. They indicate that, when given the opportunity to freely source and prospect data, decision makers tend to adopt different patterns of gathering and analyzing data. We find sequential patterns to still occur when decision makers employ a goal-oriented search and have a good understanding of how to source and analyze data. However, when decision makers rely more thoroughly on data to make a decision and employ a dynamic and open scanning for information, patterns become increasingly recursive and iterative in nature and reflect “sense-making processes” (Lycett, 2013). We provide an integrated perspective on how the structure of the decision problems (Shim et al., 2002; Simon, 1960) and modes of acquiring information (Aguilar, 1967; Huber, 1991; Vandenbosch & Huff, 1997) in data-driven environments may evoke and affect such patterns. Thus, our findings extend earlier works by Boonstra (2003) and show how the traditional phase theorem of decision making (e.g., Mintzberg et al., 1976; Simon, 1960) can be extended to data-driven contexts, by describing different path configurations of decision making phases rather than a uniform process.

Finally, the four distinct patterns show that it is imperative to account for differences in decision making when using crowdsourced data and designing decision support mechanisms. Our findings emphasize that, depending on the decision making pattern, there is different potential for information systems to support phases in such processes. For structured problems, we see great potential to increase the efficiency of decision making by automating highly repetitive or standardized data-processing tasks. For unstructured problems, we see great potential to increase the effectiveness of decision making by leveraging pattern recognition and visualization techniques to support managers in making sense of crowdsourced data.

3.4.2 Implications for Practice

From our findings, we are also able to derive a number of practical implications for organizations that are aiming to leverage crowdsourced data. The four distinct patterns identified in our study show that it is imperative to account for differences in decision making when using crowdsourced data and designing decision support mechanisms. They underline that there is no "one-size-fits-all"-solution for decision support. Instead, we urge organizations to pay close attention to the structure of the decision problems and the modes by which decision makers acquire information and adjust the mechanisms for decision support to fit different decision making patterns. In cases where decision making patterns resemble a goal-oriented search for information (e.g., when analyzing and verifying defect reports), we find it imperative to provide data in a manner that lets decision makers efficiently "inform" decisions or gradually "solve" decision problems. Data should be automatically preprocessed in the "validation" phase to reduce manual efforts and increase the efficiency of decision making. In cases where decision making patterns reflect a more dynamic exploration of data (e.g., when deciding on new products during innovation campaigns), our findings suggest that decision support systems should offer more options to experiment with data and visualize the results, e.g., to support the "consolidation" and "evaluation" phase.

Finally, in a broader sense, we recommend organizations to take advantage of crowdsourcing to source and analyze data for improved decision making. In crowdsourcing, decision makers have the opportunity to freely specify the type of data as well as the amount of data that should be generated at potentially all stages during their decision making. Especially for unstructured problems, our findings show that decision makers often repeatedly return to data "sourcing" phases to back their decisions with insights from crowds. This underlines the importance of access to adequate data during decision making. Decision making in such settings often revolves around the exploration of data in order to discover new options and incrementally develop better decisions. We believe that crowdsourcing represents a very powerful approach to provide access to such data and foster data-driven decision making processes in organizations.

3.4.3 Limitations and Outlook

As with all research, the findings presented in this study should be regarded in light of its limitations. First, we analyzed decision making in the domain of crowdsourcing. Although we do believe that crowdsourcing can be regarded as an exemplary context to gain insights into the nature of decision making when dealing with large amounts of user-generated data, we cannot claim that our findings apply to other domains or applications to the same extent. They may apply to contexts where decision makers are free to source and analyze data from large networks of people or information markets. Beyond that, we urge future research to extend our work and investigate different industries and other crowdsourcing contexts.

Second, we used semi-structured interviews as our primary source of data. While we made sure to select a broad sample of decision makers across 10 different organizations and departments with varying degrees of expertise, our results are still bound to the participants and the discussions with them. By conducting interviews, we aimed to gain in-depth insights into decision making as experienced and described by actual decision makers. We see great potential in future research to also consider quantitative data and triangulate our findings with insights from crowdsourcing platforms or other information systems that are used by decision makers. This would not only help to provide a more comprehensive picture of decision making with behavioral data, it would also allow for further insights into adequate designs of decision support systems for certain types of decision making patterns.

Third, although we studied decision making in an organizational context, we focused on individual decision making processes by managers responsible for engaging in crowdsourcing and evaluating the results. Another interesting avenue for future research is to further study the role of organizational structures, hierarchies, or teams for decision making and analyze how these aspects may change in contexts facing large amounts of user-generated data. We agree with Sharma et al. (2014) that individual decision making is only the first step in understanding how organizations may benefit from large-scale data and novel information systems, as individual decision making processes are typically embedded in organizational settings.

3.5 Conclusion of Chapter

Crowdsourcing represents a powerful approach for organizations to span their boundaries and systematically collect data from large and diverse networks of people. So far, however, literature gave little insights into the structure and patterns of decision making processes that emerge when decision makers have access to such large amounts of user-generated data. In this study, we addressed this gap and analyzed the characteristics of decision making in crowdsourcing based on interviews with decision makers from 10 corporations. Depending on the type of decision problem and the mode of acquiring information, we saw four patterns emerge with regard to how different phases in decision making are aligned and (re-)occur. Thus, for research, we showed how the traditional phase theorem of decision making can be extended to more data-driven contexts, such as crowdsourcing, by looking at different path configurations of such phases. For practitioners, we discussed how information systems can be adapted to support these patterns. In this way, our study may contribute to a better understanding of decision making in crowdsourcing.

4 AUTOMATED DATA PROCESSING IN CROWDSOURCING

This chapter addresses the second research question of the dissertation and is concerned with studying mechanisms to support decision making in crowdsourcing. The focus of the chapter lies on improvements to the efficiency of decision making. It presents the results of a quantitative study[3] that examined the potential of text mining and machine learning to automatically filter crowdsourced contributions and reduce the manual workload for decision makers. In this chapter, section 4.1 first explains the motivation and objectives of the study in more detail. Afterwards, section 4.2 explains the study's underlying hypotheses regarding the relationship between textual characteristics and contribution quality in crowdsourcing. It also describes the methodology used to test these hypotheses with a regression analysis and outlines the approach for predictive modeling with machine learning algorithms. Sections 4.3 and 4.4 analyze the results and illustrate their implications for both researchers and practitioners. Finally, section 4.5 concludes the chapter with a summary of the main findings.

4.1 The Potential of Automated Data Processing

In recent years, crowdsourcing has increasingly gained attention as an innovative approach to harness the collective resources of a broad and diverse network of people over the internet. The fundamental idea of crowdsourcing is that an organization proposes the voluntary undertaking of a task to an independent group of contributors in an open call (Blohm et al., 2013; Howe, 2006). It seeks to mobilize the creativity, knowledge, or distributed workforce of a large panel of people who perform value creation activities that have previously been carried out by designated agents, such as employees or third-party contractors. The approach grants scalable access to remote resources and allows tasks to be completed in a parallelized fashion regardless of time and location. In this vein, crowdsourcing has been found to greatly improve the efficiency and effectiveness of problem-solving in organizations (Afuah & Tucci, 2012; Jeppesen & Lakhani, 2010).

However, the potential that arises from the decentralized contributions provided by a crowd comes with a critical challenge. The quantity and complexity of information that needs to be processed and evaluated in crowdsourcing are high – especially when the contributions are submitted in a raw, textual format. In 2006, for example, more than 140'000 international participants joined the IBM Innovation Jam and submitted over 46'000 ideas in a single crowdsourcing contest (Leimeister et al., 2009). Similarly, the devastating earthquake in Haiti during January 2010 generated over 13'500

[3] A previous version of this chapter (sections 4.1 to 4.5) has been published in Rhyn and Blohm (2017a). The content has been reformatted, modified, and updated for this dissertation.

crowdsourced messages on online maps that were used to locate emergencies and distribute relief supplies (Barbier et al., 2012). As these contributions are submitted by a diverse network of people with different backgrounds and degrees of expertise, textual data in crowdsourcing usually entail a high amount of noise and ambiguity. Thus, the process of manually evaluating the data and filtering out low quality contributions is arduous and lengthy (Blohm et al., 2013). It generally accounts for one of the most time-consuming and cost-intensive steps in crowdsourcing (Zogaj et al., 2014). For example, it took Google almost three years and 3'000 employees to condense the 150'000 proposals submitted to its *Project 10^{100}* (Blohm et al., 2013).

Text mining and machine learning algorithms represent promising solutions to cope with the vast amount of contributions in crowdsourcing (Chen et al., 2012). They provide the means to discover patterns and extract useful information from textual data in a fast, scalable, and repeatable way (Debortoli, Müller, Junglas, & vom Brocke, 2016). In this vein, they offer the potential to automatically evaluate and filter contributions in crowdsourcing. Although multiple studies have asked for such automated approaches, research on crowdsourcing is still lacking feasible models for this task (Kittur et al., 2013; Zogaj et al., 2014). Our study aims to close this gap by addressing the following research question: "What textual characteristics can be used to assess and automatically predict the quality of contributions in crowdsourcing?" To answer this question, we choose a two-pronged approach that has already been used similarly in related studies (Ghose & Ipeirotis, 2011). First, we apply an explanatory regression analysis to examine textual characteristics that are associated with contribution quality in crowdsourcing. Then, we use these textual characteristics for predictive modeling with machine learning algorithms. That is, we build a classifier capable of predicting the quality of the contributions based on their textual characteristics.

Hence, the contribution of our study is twofold. For researchers, we provide a set of variables and models to explain and predict contribution quality in crowdsourcing. These models and variables can be used to assess textual contributions with machine learning algorithms and, thus, contribute to a partial automation of the evaluation process. For practitioners, we build a classifier based on the Random Forest algorithm that incorporates these variables. It is capable of automatically filtering high quality and low quality contributions submitted by a crowd and makes the process of reviewing large volumes of textual feedback more efficient.

4.2 Predictive Modeling for Automated Data Processing

4.2.1 Development of Hypotheses

For developing our model, we draw upon well-established textual features discussed in related literature (Jeon, Croft, Lee, & Park, 2006; Kim, Pantel, Chklovski, & Pennacchiotti, 2006; Liu, Cao, Lin, Huang, & Zhou, 2007; Otterbacher, 2009; Weimer & Gurevych, 2007) to operationalize contextual and representational characteristics of crowdsourced data (cf. section 2.2.3) and examine how these features are associated with contribution quality in crowdsourcing. Contextual characteristics account for the amount and the relevancy of the information provided in textual data. Representational characteristics account for the extent to which the text is presented in a clear and intelligible manner (Otterbacher, 2009).

First, the amount of information in a textual contribution has frequently been discussed as one of its most important features by related literature (Jeon et al., 2006; Kim et al., 2006; Weimer & Gurevych, 2007). Longer contributions contain more information that could potentially be relevant for the company than shorter ones (Otterbacher, 2009). It is also easier for companies to act on feedback that is well elaborated (Blohm et al., 2013), as it allows them to build a more comprehensive and coherent representation of the information in the text (Blohm et al., 2016). For example, Riedl et al. (2013) note that "more accurate, understandable, and comprehensive information enables decision makers to perform better" (p. 12). On the other hand, they emphasize that contributions that are short and less elaborated tend to deliver less information that could be required for an accurate understanding of the contributions and appropriate decision making (Riedl et al., 2013). Second, related literature also emphasizes the need to consider the relevancy of the information in a contribution (Otterbacher, 2009; Weimer & Gurevych, 2007). Otterbacher (2009) quantifies the extent to which a product review contains terms that are statistically important across other reviews. Similarly, Weimer and Gurevych (2007) use similarity features to measure the relatedness of a post to a forum topic. For crowdsourcing in particular, relevant contributions are typically characterized as containing clear and specific information for the companies to act on (Blohm et al., 2013), while vague and blurry descriptions have been found to be detrimental to contribution quality (Riedl et al., 2013). Hence, we hypothesize as follows:

Hypothesis 1. The length of a textual contribution is positively associated with the quality of the contribution.

Hypothesis 2. The specificity of the terms used in a textual contribution is positively associated with the quality of the contribution.

Besides contextual characteristics accounting for the amount and the relevancy of the information, a second layer of analysis is concerned with the representational characteristics of a contribution (Agichtein, Castillo, Donato, Gionis, & Mishne, 2008; Otterbacher, 2009). On the one hand, representational characteristics can be used as means to measure the sophistication of a contribution (Otterbacher, 2009). For example, the readability (Coleman & Liau, 1975) is frequently used to analyze the syntactic and semantic complexity of a text (Agichtein et al., 2008). In crowdsourcing, a higher readability of a contribution should enable companies to better understand the submitted content and extract relevant cues or information more easily (Blohm et al., 2016). On the other hand, representational characteristics can be broken down to purely superficial aspects, such as the extent to which a contribution respects common writing standards or reveals irregularities (Ghose & Ipeirotis, 2011; Weimer & Gurevych, 2007). Poorly written contributions containing spelling errors and grammatical mistakes increase the noise and ambiguity in the data (Agichtein et al., 2008). Such irregularities impose a higher cognitive load on the recipient in the company and make the contributions prone to misinterpretation (Blohm et al., 2016). Hence, they are likely to be detrimental to the interpretability or clarity of crowdsourced contributions and may render the acquisition of the embedded information more difficult for companies. Thus, we define the second set of our hypotheses as follows:

Hypothesis 3. The readability of a textual contribution is positively associated with the quality of the contribution.

Hypothesis 4. The number of spelling mistakes in a textual contribution is negatively associated with the quality of the contribution.

4.2.2 Data Collection

For our study, we retrieved textual data from a crowdsourcing project in the field of software testing. We conducted a crowdsourced software test in cooperation with a German-based intermediary that ranks amongst Europe’s leading platforms in this domain and manages a crowd of more than 100’000 international software testers. The test was designed as a user acceptance test for a website and has been carried out in August 2015 over the course of 5 days. It consisted of open tasks that asked the testers about their opinion on positive and negative aspects of the website as well as suggestions for further improvement. This setting was chosen for several reasons. First, user acceptance tests for websites represent one of the most frequently performed types of software tests by crowdtesting platforms, as they allow companies to gather feedback from real end users of the software (Zogaj et al., 2014). Second, user acceptance tests typically lead to a large amount of textual data which are especially time-consuming to evaluate by experts or developers. Third, the feedback retrieved during user acceptance tests resemble contributions in other domains, such as ideas in innovation management or reviews in product development. This allows the results of our study to be transferred to other crowdsourcing contexts and ensures their generalizability.

We received 309 contributions in a raw textual format from 104 testers who represent the target demographic of the website and who were randomly assigned to the software test by the intermediary. On average, the contributions contained 41 words with a standard deviation of 38 words. All contributions were written in English.

4.2.3 Expert Evaluation of Contribution Quality

As discussed previously, there is no ground truth to contributions such as ideas, feedback, or reviews. In the absence of objective measures, it is necessary to employ an expert-based baseline measure for contribution quality (Blohm et al., 2016). Therefore, we adapted the Consensual Assessment Technique for our study (Amabile, 1982). We asked two software experts to manually review the feedback. Both experts are involved in the development of the website for which the user acceptance test has been conducted. Thus, they are qualified to evaluate the contributions of the crowd. They independently reviewed all test reports by using the same evaluation scheme. The evaluation scheme is based on the framework proposed by Blohm et al. (2016) for crowdsourcing and includes four criteria: relevance, elaboration, feasibility, and novelty. To cover these criteria, we used questions developed by Nørgaard and Hornbæk (2009) who applied them analogously for assessing usability feedback in software testing. Hence, they are suitable for our study, which is concerned with similar feedback to user acceptance tests. Each criterion was rated on a 5-point Likert scale. To validate the ratings of the experts, we calculated the weighted Cohen's Kappa for each criterion (Cohen, 1968).

Relevance	*Elaboration*	*Feasibility*	*Novelty*
0.78**	0.76**	0.77**	0.73**

Note: ** Substantial agreement, see Landis and Koch (1977)

Table 5: Cohen's Kappa Statistics

Source: Own Illustration

The strength of agreement as listed in Table 5 is substantial (Landis & Koch, 1977) for all criteria, indicating that we have reliable quality measures. We used the mean to aggregate the expert ratings. Since we analyze contribution quality as a multidimensional construct (Blohm et al., 2016), we followed past research (Barki & Pinsonneault, 2001;

Blohm et al., 2010; Gallupe et al., 1992) and calculated a composite score for contribution quality by averaging the ratings.

4.2.4 Variables and Measurements

We draw upon related literature (Jeon et al., 2006; Kim et al., 2006; Liu et al., 2007; Otterbacher, 2009; Weimer & Gurevych, 2007) and use the textual features derived in section 4.2 as variables to explain and predict the quality of the crowdsourced contributions. We use two variables (i.e., length and specificity) to account for their contextual characteristics and two variables (i.e., readability and spelling) to account for their representational characteristics.

Length. We measure the length of a contribution by counting the total number of words per contribution.

Specificity. We measure the specificity by building the sum of all TF.IDF-indices for a contribution. The TF.IDF-index represents a term weighting scheme that accounts for the importance of a particular term in the data set based on the term frequency and the inverse document frequency (Hotho, Nürnberger, & Paaß, 2005). Generally speaking, broad and frequently used terms by the crowd (e.g., “bad” or “design”) will receive lower values than more specific terms (e.g., “unintuitive” or “navigation”). For calculating these TF.IDF-indices, we follow the commonly used bag-of-words approach with a vector space model and apply standard preprocessing steps (Feldman & Sanger, 2007). More specifically, we tokenize the contributions by breaking them up into individual terms. We apply standard transformations to the single terms, including normalization (i.e., transforming all characters to lower-case), stop word filtering (i.e., removing terms such as articles or prepositions that bear no value for the analysis) and stemming (i.e., reducing terms to their root form to avoid duplications) with the Porter stemmer (Porter, 1980).

Readability. We follow Ghose and Ipeirotis (2011) as well as Blohm et al. (2016) and measure the readability of the text by calculating the Coleman-Liau index (Coleman & Liau, 1975) for each contribution. This index captures the complexity of the contributions by analyzing part-of-speech tags and measuring the average length of their terms and sentences. A higher index indicates a better readability for the text.

Spelling. Finally, we measure irregularities and non-conformance to writing standards by counting the number of spelling errors per contribution. In order to ensure that the spelling errors were accurately captured, we manually reviewed all 309 contributions.

4.3 Models and Results

4.3.1 Explanatory Regression Analysis

In this section, we use regression modeling to analyze whether the textual features of the contributions are associated with their quality. The length of a contribution, the specificity of the terms, the readability of the text, and the number of spelling errors represent the independent variables. The contribution quality as rated by the experts represents the dependent variable. The results are depicted in Table 6.

Coefficient	*Estimate*	*Std. Err.*	*t-value*	*p-value*	
(Intercept)	2.890	0.039	74.395	< 2.2e-16	***
Length	12.091	0.783	15.451	< 2.2e-16	***
Length (poly 2)	-4.730	0.694	-6.813	5.18e-11	***
Length (poly 3)	2.930	0.710	4.124	4.82e-05	***
Specificity	1.752	0.721	2.429	0.016	*
Readability	2.333	0.708	3.297	0.001	**
Spelling	-1.847	0.814	-2.269	0.024	*

Note: ***p < 0.001; **p < 0.01; *p < 0.05

Residual Standard Error: 0.683; R-Sq. (adj.): 0.554; F(6,302): 64.8; p-value: < 2.2e-16

Table 6: Regression Analysis

Source: Own Illustration

It shows that the length ($t = 15.451$; $p = < 2.2e\text{-}16$) and the readability ($t = 3.297$; $p = 0.001$) of a contribution are highly significant indicators for its quality. Both features are positively correlated to the quality of the contribution. Interestingly, as indicated by the polynomials, we observe a diminishing marginal utility effect associated with number of words in a contribution, which seems conceptually reasonable. Writing 55 instead of 5 words benefits the contribution more than extending it from 150 to 200 words. Regardless of this effect, our results still support H1 which states that the length of a textual contribution is positively associated with the quality of the contribution. The model also supports H3 and shows that the readability of a textual contribution is positively associated with the quality of the contribution. Similarly, the specificity of the terms ($t = 2.429$; $p = 0.016$) and the number of spelling mistakes ($t = -2.269$; $p = 0.024$) in a contribution are significant indicators for its quality. The former is positively correlated to the quality of the contribution. The latter is negatively correlated to the quality of the contribution. These results support H2 and H4. The model reveals a high value for R^2 and explains the quality of the contributions significantly well. We found no evidence that potential effects between the individual contributors and their contributions affect our results. We also examined the residuals and found our model to be sound. There are no signs of heteroscedasticity nor autocorrelation. The residual show to be normally distributed. We can conclude that the proposed variables explain the quality of crowdsourced contributions at statistically significant levels. We find support for our four hypotheses and will use these findings as the foundation for predictive modeling.

4.3.2 Predictive Modeling

Based on the previously analyzed variables, we train and evaluate a classifier that is capable of predicting the quality of the contributions and automatically filter them. A binary classification allows for a clear selection rule (Blohm et al., 2013) that decides on whether the contributions fulfill the quality requirements and are thus eligible to be forwarded to the organization for further consideration or whether they are of poor quality and should be filtered out to not induce unnecessary workload. Hence, we represent the evaluation of the contributions as a classification problem. We set the threshold for separating high quality from low quality feedback to 3.5, which is comparable to previous studies conducted for product reviews (Ghose & Ipeirotis, 2011), and labeled the contributions. As a result, 83 contributions were classified as high-quality contributions, whereas 226 contributions were classified as low-quality contributions. This distribution is consistent with findings documented in previous studies on the quality of crowdsourced contributions (Blohm et al., 2013; Ghose & Ipeirotis, 2011). We tested different classification algorithms and compared the performance of Logistic Regression, Naïve Bayes, k-Nearest Neighbor, Decision Trees, and Random Forest for this study. We found the Random Forest algorithm to perform substantially better in classifying the contributions compared to the other approaches – both regarding the accuracy and the receiver operating characteristic. Our findings are consistent with comparative experiments conducted for similar classification tasks (Ghose & Ipeirotis, 2011). Thus, we focus on the results of the Random Forest algorithm.

The Random Forest algorithm (Breiman, 2001) builds a large number of decision trees with different combinations of the given variables. These decision trees are internally trained and evaluated using random subsets of the same data. The Random Forest model then averages the decision trees. In this vein, it reduces the variance that comes with individual decision trees, provides information about the importance of the variables for the classification, and overcomes the risk of overfitting (Hastie, Tibshirani, & Friedman, 2009).

We use 100 decision trees for our Random Forest model and set the cutoff for the model's probability estimates at the standard value of 0.5. To build and evaluate the classifier, we followed the widely used k-fold cross-validation approach with 5 folds. That is, we randomly split our data set in a stratified manner into 5 subsets. 4 subsets are used to train the classifier with the given labels. The remaining subset does not include the quality labels and is used to evaluate the performance of the classifier by comparing the labels predicted by the Random Forest algorithm to the actual labels provided by the experts. We measure the accuracy, the sensitivity, the specificity, and the receiver operating characteristic (Fawcett, 2006). This procedure is repeated until each split of the data set has been used to train and evaluate the classifier.

The results of the cross-validation reveal an accuracy of 80.03% on average for our Random Forest model. Thus, by only using the four variables based on our proposed textual features, the algorithm is able to automatically predict the quality of the crowdsourced contributions and correctly classify them in over 80% of the cases. The classifier shows a very high specificity of 87.73%, indicating that it performs exceptionally well at recognizing and filtering low quality contributions. As suggested by the slightly lower sensitivity measure (60.27%), it is more difficult for the algorithm to achieve a high true positive rate. The sensitivity of the classifier can be increased by adjusting the cutoff for the probability estimates. Lowering the cutoff by 20% increases the classifier's sensitivity to 75.30%. Naturally, however, this comes at the expense of reducing its specificity to 76.56%.

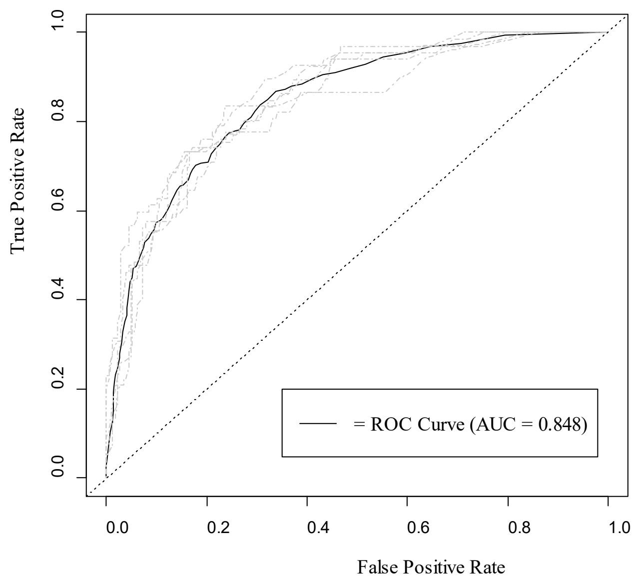


Figure 4: Receiver Operating Characteristic

Source: Own Illustration

The curve of the classifier's receiver operating characteristic (ROC) is depicted in Figure 4. It plots the true positive rate against the false positive rate (Fawcett, 2006). The diagonally plotted line represents the strategy of randomly guessing the quality of the contributions. A classifiers that reaches the upper triangular region of this line exploits information in the data and performs better than the random classification strategy (Fawcett, 2006). The area under curve (AUC) is equivalent to "the probability that the classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance" (Fawcett, 2006, p. 868), making it also equivalent to the Wilcoxon test of ranks. Here, the AUC reveals a high value of 0.848. Our classification algorithm performs very well.

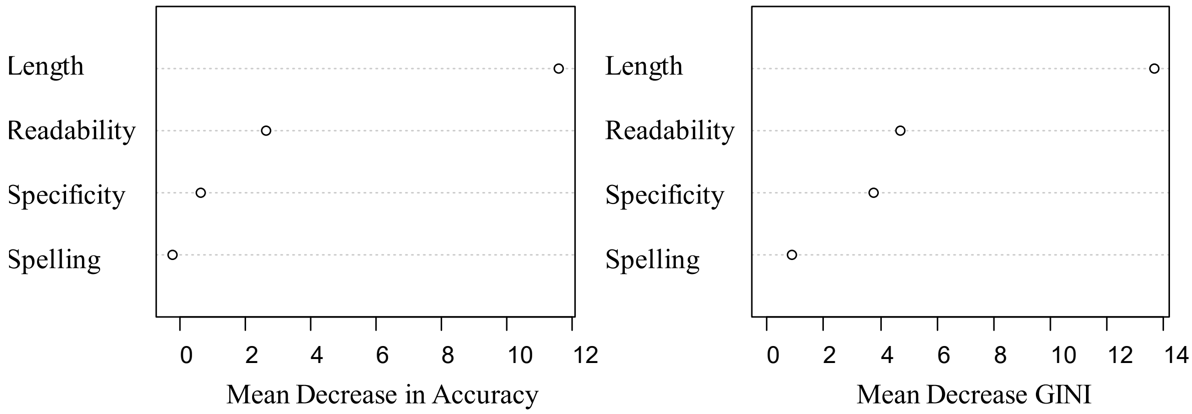


Figure 5: Variable Importance Plots

Source: Own Illustration

Finally, Figure 5 displays the importance of the four proposed variables, measured by the mean decrease in accuracy and the mean decrease in node impurity (i.e., Gini index) for each variable (Liaw, 2015). All variables were used by the Random Forest algorithm and have predictive power. The length of the contribution is by far the most important variable for the classification. When aiming for a sparse prediction model, the variable "Spelling" may be omitted without risking much worse results.

4.4 Discussion of Predictive Modeling Results

The models and results presented in the previous section yield two important findings. First, we find support for our hypotheses and show that the length of a contribution, the specificity of the terms, the readability of the text, and the number of spelling errors are all associated with contribution quality in crowdsourcing at statistically significant levels. Therefore, in a second step, we used the textual characteristics in combination with an expert-based baseline measure for contribution quality to train and evaluate an algorithm capable of predicting the contribution quality and classifying the data. Even for a small data set of 309 contributions, our Random Forest classifier achieves an accuracy of 80.03%. The algorithm has shown to perform especially well at recognizing and filtering low quality contributions. It outperforms random classification substantially and also achieves a much higher accuracy compared to a naïve classifier that would always predict the category with the majority of the ratings (i.e., 73.14%). Thus, our Random Forest algorithm proves to be very reliable. These findings have valuable implications for both researchers and practitioners.

4.4.1 Theoretical Implications

To the best of our knowledge, we are the first to show that it is possible to reliably explain and predict the quality of contributions in crowdsourcing based on textual features of the data alone. We provide empirical evidence for the relationship between both contextual and representational characteristics of contributions and their quality in crowdsourcing. This indicates that well elaborated and precise solutions, ideas, or suggestions are vital for companies trying to leverage the information submitted by a crowd. Furthermore, our results suggest that companies require textual contributions to be presented in a clear and easily interpretable manner to fully benefit from them.

Moreover, we contribute a set of models and variables to operationalize these contextual and representational characteristics. The models and variables proposed in our study have been shown to work well with algorithms capable of automatically assessing and classifying textual contributions. In this vein, we provide the foundation for partially automating the evaluation of textual data in crowdsourcing, which has frequently been requested by related literature. Kittur et al. (2013) emphasized that, while "quality control is improving for tasks with a closed set of possible answers, we still have few techniques for open-ended work and highly skilled tasks" (p. 7-8). The authors specifically asked for studies to analyze potential metrics and propose feasible approaches to predict output quality. In crowdsourced software testing in particular, related work has expressed the need for efficient mechanisms to assess the quality of crowdsourced contributions and automate the evaluation of the data (Zogaj et al., 2014). With our study, we close this gap and extend existing research that already uses machine learning and text mining algorithms to cluster the variety of topics covered in crowdsourcing projects (Barbier et al., 2012; Feng et al., 2015; Rogstadius et al., 2013; Walter & Back, 2013) by providing both the appropriate variables and models for an automated evaluation of high quality and low quality contributions in the potentially large sets of textual data. Regarding the importance of different variables, we found the length of a contribution to be the most effective indicator for explaining and predicting its quality. Both the readability of the contributions and the specificity of the terms are positively associated with the quality of the contributions at highly significant levels but reveal only moderate predictive power for classification algorithms. Interestingly, spelling errors have shown to be the least important feature for the classification and may even be omitted for sparse models. Therefore, our findings may help researchers in selecting variables for predictive modeling in crowdsourcing.

4.4.2 Practical Implications

Our proposed Random Forest classifier allows companies to substantially reduce the amount of information that needs to be reviewed manually. It shows that the classification algorithm is capable of automatically identifying high quality contributions in large data sets and removing those that do not fulfill the quality requirements defined by the companies or platforms. In this study, we set the threshold to only include the top 30% of the contributions. Hence, the algorithm can make the evaluation of the results submitted to crowdsourcing projects much more efficient and offers both time and cost savings. It is possible to incorporate the algorithm directly as a filter mechanism on the platforms or in tools for companies retrieving data from these platforms.

We also show that the sensitivity and specificity of the Random Forest algorithm can be adjusted to fit the preferences of practitioners. As both measures are inherently linked to each other, the decision to increase one measure will always come with the trade-off of decreasing the other. If the costs of wrongfully rejecting a high quality contribution is higher than the cost of wrongfully including a low quality contribution in the evaluation process, this is a trade-off that should potentially be considered.

Finally, our automated machine learning and text mining approach also contributes to practitioners in the domain of software testing. Related work already proposes algorithms that can be used to evaluate technical bug reports more efficiently. For example, it is possible to automatically assess the severity of the bug reports (Lamkanfi, Demeyer, Soetens, & Verdonckz, 2011) and detect duplicates in the data sets (Runeson, Alexandersson, & Nyholm, 2007). As our data stem from crowdsourced software testing, we extend these findings and provide developers with an approach to facilitate the evaluation of test reports obtained in user acceptance testing, user experience testing, or usability testing. These contributions are typically submitted in a free text format and entail a high workload for the developers (Zogaj et al., 2014). Our proposed classifier may help developers in evaluating these types of test reports more efficiently.

4.4.3 Limitations and Outlook

As with any research, our work does have its limitations. First, the manually assigned quality labels used for our data set are inherently dependent upon the rating scales and the subjective judgements of the experts. We attempted to address this issue by using scales that have been developed specifically for crowdsourced contributions as analyzed in this study (Blohm et al., 2016). Furthermore, we let two experts independently review the contributions. The Cohen's Kappas indicate an intersubjective agreement between the experts. Second, the data set stems from a crowdsourcing project in the field of software testing. We aimed to provide as much generalizability as possible by choosing a user acceptance test setting that yields contributions similar to other crowdsourcing contexts that are based on textual data, such as ideas, feedback, or reviews.

The findings presented in this study may encourage future efforts to analyze the performance of the proposed features or models in different crowdsourcing settings and expand on our initial results. There is still great potential in making the algorithms cost-sensitive and studying the optimal trade-off between sensitivity and specificity in crowdsourcing. Furthermore, as we focused on the textual characteristics of a contribution, future work may also examine the role of non-textual characteristics and analyze features such as the experience or the expertise of the individuals who submitted the contributions. Finally, text mining and machine learning methods benefit from large data sets. Hence, we need scalable concepts for labeling crowdsourced contributions and training algorithms with more data. Addressing these issues would pave the way for leveraging the full potential of machine learning in crowdsourcing.

4.5 Conclusion of Chapter

The process of manually reviewing and filtering large volumes of textual contributions has been a longstanding challenge in crowdsourcing. Given the unstructured format of textual data and the diversity of inputs submitted by a crowd, identifying valuable inputs and separating them from low quality contributions that cannot be used by the companies is very time-consuming and cost-intensive. In this study, we propose an approach based on the principles of text mining and machine learning to partially automatize this process. Our results indicate that it is possible to explain the quality of crowdsourced contributions purely based on textual features, such as the length of a contribution, the specificity of the words, the readability of the text, and the number of spelling errors. We use these textual features in combination with an expert-based baseline measure to train and evaluate a classification algorithm that is capable of reliably predicting the quality of the contributions and automatically filtering them for companies.

5 IDENTIFYING USEFUL IDEAS IN CROWDSOURCING

This chapter addresses the second research question of the dissertation. The focus lies on improvements to the effectiveness of decision making. Existing research shows that, when faced with a large number of contributions in crowdsourcing, decision makers often attend to only a subset of contributions due to their limited ability to process all available information (Piezunka & Dahlander, 2015). This ultimately hampers their ability to make effective decisions as it becomes increasingly difficult to identify useful contributions in the vast pools of data generated by crowds. Thus, this chapter presents the results of a study[4] that investigated the potential of network analysis and text mining to support decision makers in tracking the origin of contributions, analyzing their content, and ultimately spotting the most useful ones. The chapter is organized as follows: First, section 5.1 explains the motivation and objectives of the study in more detail. Second, section 5.2 outlines the underlying hypotheses regarding the emergence and recombination of knowledge on crowdsourcing platform and explains the methodology to test the hypotheses. Third, sections 5.3 and 5.4 describe the results and discuss their implications for research and practice. Section 5.5 concludes the chapter.

5.1 Studying the Emergence of Useful Ideas in Crowdsourcing

Over the past decades, crowdsourcing has attracted much attention for its competitive advantages over traditional work structures in mobilizing distributed workforce and leveraging innovation (Thuan, Antunes, & Johnstone, 2016). In crowdsourcing, an organization uses an open call to outsource tasks that have previously been performed by dedicated employees or contractors to an independent network of people. Compared to traditional sourcing mechanisms that rely on only few designated agents, crowdsourcing deliberately seeks to harness the collective knowledge or creativity of the masses (Schenk & Guittard, 2011). It allows organizations to move away from predefined routines and facilitates the search for distant knowledge outside their existing boundaries (Afuah & Tucci, 2012). In consequence, crowdsourcing is currently being applied in a variety of different domains, including innovation management (e.g., Leimeister et al., 2009; Poetz & Schreier, 2012), software development (e.g., Leicht et al., 2017; Stol, LaToza, & Bird, 2017), and the humanitarian aid sector (e.g., Barbier et al., 2012; Rogstadius et al., 2013).

[4] A previous version of this chapter (sections 5.1 to 5.5) has been published in Rhyn et al. (2017). The content has been reformatted, modified, and updated for this dissertation.

While crowdsourcing offers the potential to search for distant knowledge in a very efficient and effective manner (Afuah & Tucci, 2012), the quantity and complexity of information impinging on organizations is exceptionally high with this approach. IBM, for example, faced more than 46’000 ideas submitted by 150’000 contributors during its Innovation Jam (Bjelland & Wood, 2008). Similarly, Dell’s innovation platform IdeaStorm has yielded more than 26’000 ideas with over 100’000 comments since its inception (Dell, 2017). Given the limited ability of organizations to process information, they frequently fail to harness the full potential of crowdsourcing when searching for new and useful knowledge in such large pools of contributions. More specifically, Blohm et al. (2013) emphasize that crowdsourcing typically yields a large number of contributions that only have limited value for organizations and that the search for useful suggestions often represents a great challenge on crowdsourcing platforms. Similarly, Piezunka and Dahlander (2015) analyzed how 922 organizations responded to contributions submitted by crowds and found that organizations often miss valuable information because they are exposed to an overload of worthless information. They argue that organizations in crowdsourcing may “succeed in generating a particularly large amount of new knowledge, but that they fail to pay attention to the knowledge that has the most potential for innovation” (Piezunka & Dahlander, 2015, p. 875).

In existing literature, little is known about the emergence and evolution of new knowledge on crowdsourcing platforms and how organizations may identify contributions that capture such knowledge. Instead, much research has focused on well-approved contributions generated by experienced lead users in a crowd. Li et al. (2016) found that popular ideas submitted by contributors with prior experience on crowdsourcing platforms are more likely to be implemented than less popular ideas submitted by unknown contributors. Similarly, Schemmann et al. (2016) show that the chance of an idea being implemented by an organization increases when the contributor of the idea has previously examined other crowdsourced ideas and when the idea is popular within the crowd. In earlier studies, Huang et al. (2014) observed that contributors on crowdsourcing platforms learn how to come up with promising ideas over time through increased participation and peer voting. Bayus (2013) provides evidence that serial contributors are more likely to generate an idea that will be implemented than contributors with few ideas. However, popular ideas or solutions generated by experienced members of a crowd may not be the most useful contributions for organizations seeking to span their boundaries and gain new knowledge. Lüthje et al. (2005), for example, show that user-innovators in crowds almost always use local information to determine the need for new solutions and develop them. Thus, in the words of Piezunka and Dahlander (2015), organizations that engage in crowdsourcing face the risk of being lured and lulled by their crowds – "lured into wasting attention on the process of discerning good ideas from bad, and lulled into believing that the ideas expressed most often or most loudly are also the best" (p. 876).

We argue that, in order to find useful contributions in their search for new knowledge on crowdsourcing platforms, organizations must analyze the network structure of their crowds and monitor the topics that are being discussed among its members. We analyze cross-sectional data from a large crowdsourcing platform in Europe and combine statistical approaches from the fields of information retrieval, text mining, and network analysis to answer the following research question: How do useful contributions emerge and evolve on crowdsourcing platforms? We find evidence that such contributions are more likely to originate from members with only few effective network ties in the crowd. These contributions introduce new information and distant perspectives to the crowd, which are then further enriched and combined with local knowledge provided by experienced members on the platform. We argue that organizations searching for new and useful knowledge through crowdsourcing should pay attention to these types of contributions.

With these findings, our contribution is twofold. For research, we provide an extended understanding of how crowdsourcing may be employed for distant search in organizations. We examine the effects of network relationships and knowledge collaboration on the usefulness of crowdsourced contributions and find that simply engaging in crowdsourcing in an attempt to span organizational boundaries may not suffice to find distant knowledge. We show that, when searching for distant knowledge, it is important to differentiate the submitted contributions with regard to their origin in the crowd and their similarity to existing knowledge. We also show that useful contributions not only emerge from isolated contributions alone but from a combination of different topics brought together by members of a crowd. From a practical perspective, our results provide guidance for crowdsourcing intermediaries or organizations that host their own crowdsourcing platforms on how to identify useful contributions in the vast pool of data generated by their crowds. Based on our findings, we urge organization not to rely exclusively on common rating scales or voting systems for assessing crowdsourced contributions. Instead, we see great potential in network analysis and text mining to support organizations in tracking the origin of contributions in crowdsourcing, analyzing their content, and ultimately identifying the most useful ones.

5.2 Statistical Analysis of Useful Ideas in Crowdsourcing

In order to address our research question and empirically test the outlined hypotheses, we combine statistical approaches from the fields of information retrieval, text mining, and network analysis to examine cross-sectional data retrieved from a crowdsourcing platform of a large Swiss organization. All calculations were performed with the R Language for Statistical Computing. For processing the textual data and analyzing their content, we employed algorithms provided by the tm, openNLP, and topicmodels packages. The network structure of the crowd was mapped and analyzed using the igraph package.

5.2.1 Development of Hypotheses

Organizations can use or host crowdsourcing platforms in an effort to span their organizational boundaries and engage in distant search. The platforms grant them access (i.e., provide an interface) to a large and diverse network of people who are willing to contribute their ideas or solutions to a particular problem. However, when searching for actually useful contributions and new knowledge on these crowdsourcing platforms, it is important to understand how networks of people form around organizations and how they create and share knowledge. Thus, for the development of our hypotheses, we draw upon theoretical findings from two different streams of research. First, we ground our study on prior work in the fields of network theory and problem solving (e.g., Perry-Smith & Shalley, 2003). That is, we focus on crowdsourcing settings that allow knowledge to be shared and jointly developed amongst members of a crowd. A number of studies have already employed a network perspective to analyze such connected crowds that form around focal organizations or topics (e.g., Lu, Singh, & Sun, 2017; Simula & Ahola, 2014; Stephen, Zubcsek, & Goldenberg, 2016). This stream of research provides valuable insights on how individuals with different positions in the network (i.e., the crowd) introduce new ideas or problem-solving approaches that may represent distant knowledge for organizations. A second, relevant stream of research is concerned with knowledge collaboration and offers insights on how crowds discuss and enrich these initial ideas or solutions to make them useful for organizations (e.g., Faraj, Jarvenpaa, & Majchrzak, 2011).

First, we refer to research in the field of problem-solving and network theory. Existing literature in this field generally suggests that individuals have different perspectives on problems and employ different heuristics to derive solutions based on their prior experiences and their domain of expertise (cf. Jeppesen & Lakhani, 2010). This is based on findings that human problem solving involves the construction of an internal representation of the problem and the application of an appropriate heuristic to search for a potential solution (Dunbar, 1998; Jeppesen & Lakhani, 2010). It has been found that individuals overwhelmingly use familiar knowledge and prior experience in developing solutions to problems they encounter (Jeppesen & Lakhani, 2010; Lovett & Anderson, 1996; Lüthje et al., 2005). Although expertise in a domain and familiarity with existing knowledge may be useful, the most innovative or useful solutions may not necessarily be brought up by individuals affiliated with the actual domain of the problem. As shown by Jeppesen and Lakhani (2010), individuals that are technically and structurally distant from the problem domain "can offer perspectives and heuristics that are novel and thus useful for generating solutions to these problems" (p. 1019). They are often naïve with regard to the prevailing assumptions or theories in a particular domain (Gieryn & Hirsh, 1983) and have access to differing knowledge and perspectives compared to individuals that are local to the domain. Research also shows that psychological distance greatly benefits creativity (Förster, Friedman, & Liberman, 2004; Trope & Liberman, 2010). Being remote to a problem and thinking in abstract terms may lead to more diverse and original solutions whereas thinking in concrete, technical terms often impedes innovation (Förster et al., 2004).

These arguments are also supported from a network perspective. Perry-Smith and Shalley (2003) illustrate that mental representations tend to converge in local networks (such as crowds) due to common experiences and an increased sharing of redundant information. As individuals immerse in a particular network, it becomes more difficult for them to see beyond their direct ties, which provide mostly conformant information (Perry-Smith & Shalley, 2003). Similarly, as the ties within a network or domain become stronger, conformity will hamper creativity. In consequence, it is argued that perspectives drawn from distant positions in a network will likely be more novel relative to existing standards within the domain (Perry-Smith & Shalley, 2003).

Following this stream of research, we hypothesize that individuals with few network ties who are not already immersed in the crowd are more likely to provide new paradigms or problem-solving approaches to the platform than experienced individuals with many network ties in a crowd. Ultimately, these contributions should be more useful for organizations that search for new and distant knowledge through crowdsourcing. Furthermore, in line with the previous argument, we expect contributions that offer novel information to be more useful for organizations than contributions that refer to already available information.

H1. Contributions that are created by members of a crowd with few network ties are more likely to be useful for organizations than contributions created by members of a crowd with many network ties.

H2. Contributions that contain novel information are more likely to be useful for organizations than contributions that refer to already available information.

A second stream of research that provides valuable insights into the development of knowledge in crowdsourcing revolves around knowledge collaboration (Faraj et al., 2011). Existing literature in this field suggests that knowledge usually emerges not from a single contribution alone but from a recombination, modification, and integration of knowledge provided by different individuals in a network (Ye et al., 2016). Given that individuals employ different perspectives and problem-solving heuristics, information to develop innovations or solve problems have generally been found to be widely distributed among many people rather than concentrated among only few prolific individuals (Von Hippel, 2005). When perspectives from distinct fields are brought together and combined, the proposed solutions or ideas are argued to have a high potential to be novel and deviate from established mindsets (Perry-Smith & Shalley, 2003). Mumford and Gustafson (1988) in particular suggest that high levels of creativity usually emerge from very different schemata or cognitive structures being combined. Even in design science, it is argued that innovation and new knowledge are typically developed from an iterative expansion or revision of existing concepts and knowledge during the design process (e.g., Braha & Reich, 2003; Hatchuel & Weil, 2009).

A large number of studies indicate that collectively recombining knowledge is essential for developing innovative ideas or problem-solving approaches (e.g., Faraj et al., 2011; Ye et al., 2016). Recombination enables crowds to further enrich and develop ideas or solutions and aggregate the content for a more in-depth and comprehensive understanding. Furthermore, crowdsourcing greatly benefits from different topics or perspectives coming together on the platforms. Exposure to different alternatives or new perspectives has been found to trigger a process of using wider categorizations and generating more divergent solutions (Kanter, 1988; Perry-Smith & Shalley, 2003). Especially for tasks that revolve around the development of alternative ideas, access to diverse knowledge and perspectives lead to greater quantities of non-redundant solutions (Chatman, Polzer, Barsade, & Neale, 2007; De Dreu & West, 2001; Riedl & Woolley, 2017). In crowdsourcing specifically, Riedl and Woolley (2017) observed that crowds coming up with useful solutions draw on diverse sets of topics and exhibit less redundancy regarding the exchanged information. The diversity of information in their discussions has been found to be much higher. Results presented in a related study conducted by Bayus

(2013) points in a similar direction, showing that an individual's likelihood of generating promising ideas in crowdsourcing is positively affected by the diversity of his or her commenting activity on other ideas. In general, these findings suggest that useful knowledge in crowdsourcing may especially emerge from collective contributions provided by individual members of a crowd that collaborated and became creative (Geiger & Schader, 2014; Ye et al., 2016).

These insights resonate strongly with the previously discussed (re)combination of knowledge in search theory (Fleming, 2001). As assumed in H1 and H2, members with few network ties in the crowd are likely to introduce novel information, perspectives, or problem-solving approaches. Members who have established many network ties and who are immersed in the network may combine or recombine this new information with local knowledge that has already matured on the platform. They add more in-depth information to the initial contribution and combine it with different topics or insights that have already been accumulated. Thus, we assume that a contribution becomes more useful for organizations that search for new knowledge when the crowd combines it with novel information and adds depth to it. We also argue that contributions that offer a broader scope of perspectives and draw upon a diverse set of topics on the crowdsourcing platform are more likely to be useful for organizations than contributions that draw upon a less diverse set of perspectives.

H3. Contributions whose discussion adds novel information are more likely to be useful for organizations than contributions whose discussion refers to available information.

H4. Contributions that combine information from a broad set of topics on the platform are more likely to be useful for organizations than contributions that refer to a narrow set of topics.

Figure 6 depicts the underlying research model of this study. For organizations that search for new knowledge on crowdsourcing platforms, we expect the number of network ties a contributor has established in the crowd (H1) to be negatively related to the usefulness of a contribution and the novelty of the contribution (H2) to be positively related to the usefulness of a contribution. Furthermore, we assume that the usefulness of a contribution is positively affected when novel information is added by other members of the crowd (H3) and when a broad set of topics are combined (H4). We also control for alternative effects that are explained in more detail in the subsequent section below.

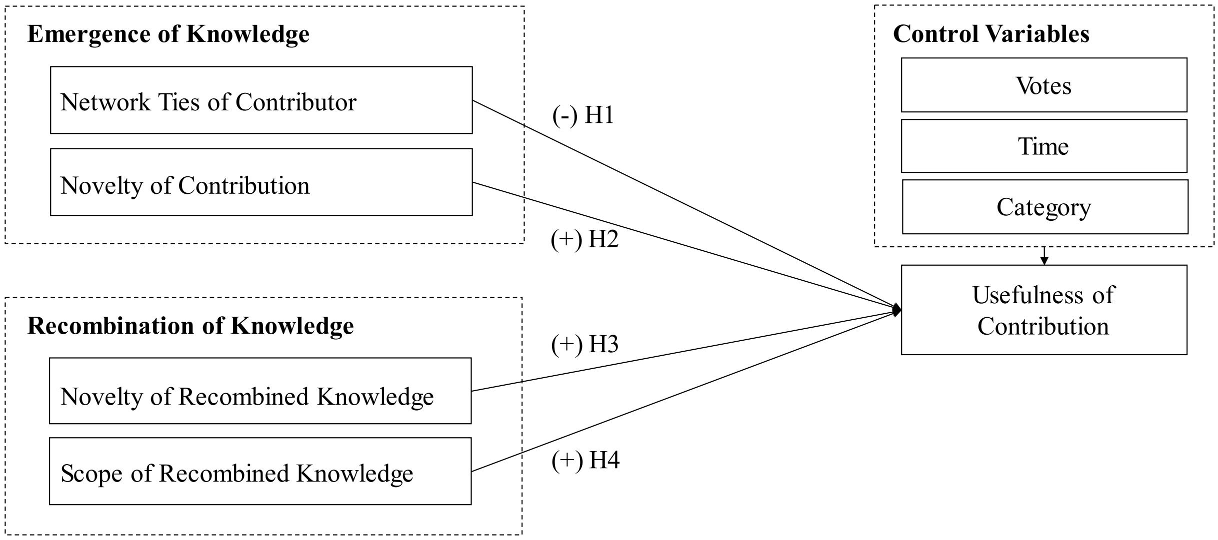


Figure 6: Research Model I

Source: Own Illustration

5.2.2 Description of Dataset

For our study, we use a unique data set retrieved from a German-speaking crowdsourcing platform. The platform is operated by a Swiss transportation and logistics organization. Most notably, it revolves around web services and smartphone applications developed by the organization that include, for example, a ticket purchasing system, times schedules, or itineraries. On the platform, members of the crowd may either contribute their own ideas or comment and vote on the contributions of others. In addition, the organization frequently issues open calls on the platform asking the crowd for their feedback or ideas on new product releases. In this way, the platform aims to foster a collaborative discourse amongst its users and ensure interaction in the crowd. Thus, the platform hosts a rather stable, slowly growing crowd and does not rely on temporary teams or contests with time constraints. The ideas on the platform are continuously reviewed and evaluated by a team of 29 administrators. The administrators work for the organization and are responsible for identifying the most useful contributions for implementation.

As of August 2016, the platform has a total of 11'408 registered members in the crowd. While it is hosted by the organization and predominantly appeals to its customers, there are no prerequisites for joining the crowd. Hence, it represents an open crowd (cf. Corney, Torres-Sánchez, Jagadeesan, & Regli, 2010; Zwass, 2010) that includes potentially everyone who's interested and willing to contribute ideas or comments to the organization. 1'700 (or 14.9%) of all registered members are active and made at least one contribution (i.e., idea, solution, or comment). On average, active members of the crowd made 2.86 (SD: 13.5) contributions on the platform. They receive no monetary compensation for their efforts. However, the organization occasionally awards the most active and valuable contributors in the crowd by inviting them to sponsored events or dinners.

With regard to the contributions, the original dataset retrieved from the platform contained a total of 2’304 ideas and 2’564 comments that discuss them. In preparation for the analysis, we cleaned the dataset. First, we removed all ideas and comments that were not written in German (i.e., only 2.3% of all contributions). Since we are analyzing the content of the contributions with text mining algorithms, this step was necessary in order to ensure that the language of our text corpus is uniform and does not confound the results. Second, we excluded all responses by the administrators and moderators for our analysis. This was done because we are only interested in contributions generated by actual members of the crowd and not the organization. The administrators’ responses are mostly “standard” responses, for example, to thank the crowd for their contributions. Third, we followed the commonly used bag-of-words approach and applied standard preprocessing steps in order to make the unstructured, user-generated text in our dataset compatible for text mining algorithms (cf. Feldman & Sanger, 2007). That is, we tokenized the contributions and broke them up into individual terms. We applied standard transformations to the terms, including normalization (i.e., transforming all characters to lower-case) and stop word filtering (i.e., removing terms such as prepositions or articles that bear no value for our analysis). We also excluded extremely scarce terms (e.g., terms that were misspelled and, thus, treated as a separate, “new”, terms). Our final dataset comprises a total of 1’927 ideas or suggestions for improvements with 1’859 comments. Of these 1’927 ideas or suggestions for improvements, 258 contributions have been implemented by the organization. The dataset has been retrieved in August 2016 and contains all information since the initial launch of the crowdsourcing platform in October 2015.

The dataset is suitable for our analysis for several reasons. First, the crowdsourcing platform was specifically created by the transportation and logistics provider for the purpose of spanning its organizational boundaries and eliciting new ideas or feedback from an independent crowd of users on how to improve its products and services. Thus, the explicit purpose of the platform is to access distant knowledge. Second, as representatives of the organization use the platform themselves to evaluate and select useful contributions, we have reliable expert labels and statuses on which our analysis can be grounded (see subsequent section below). Third, we note that the design of the platform and the characteristics of our dataset are comparable to those reported in related studies (e.g.,

M. Li et al., 2016; Schemmann et al., 2016). This should benefit the generalizability of our results.

5.2.3 Variables and Measures

Usefulness of the Contribution. We use the decision to implement a contribution as our dependent variable. Ultimately, the implementation indicates whether a contribution has been deemed useful by representatives of the organization for solving a problem or serving as innovation. In this way, we follow a large number of related studies that have already used this rationale to address similar research questions in crowdsourcing (e.g., M. Li et al., 2016; Piezunka & Dahlander, 2015; Schemmann et al., 2016) and the notion Levitt (1963) who states that “ideas are useless unless used” (p. 79). The decision to implement an idea is made by representatives of the organization who serve as administrators on the platforms. These representatives are responsible for identifying the most useful contributions submitted by the crowd. They manually review and evaluate the contributions on the platform and assign a status to them depending on their decision. Thus, we have reliable labels on whether a contribution has been deemed useful or not. The status of a contribution is a binary variable (0 = not implemented, 1 = implemented) and has been retrieved from the platform.

Network Ties of Contributor. To measure the network ties of a contributor, we constructed a social network of the crowd on the platform. Each node in the network represents a member of the crowd. An edge or tie (i.e., a relationship) between two members of the crowd was established when two members exchanged information (i.e., when they commented or voted on their ideas or solutions). For simplicity, we constructed a non-directional network. In this way, we followed Hautz et al. (2010) and adopted a weak notion of network relationships. We calculate the effective network size (Borgatti, 1997; Burt, 1995) for each member of the crowd. The effective network size measures the number of ties that a member of the crowd has established in the network while discounting redundant ties. Accounting for redundant ties is important in our case as it can affect the type of knowledge that is exchanged between users (i.e., redundant information or diverse information from different network relationships). As each contribution on the platform was created by a member of the crowd, we are able to examine the origin of the contribution in the network. Contributions that are created by members of the crowd with many effective network ties are characterized by a large effective network size while contributions that are created by members of the crowd with only few effective network ties are characterized by a small effective network size.

Novelty of Contribution. In order to measure the novelty of information in a contribution, we process the content of the contribution with text mining algorithms and calculate an aggregated TF-IDF-index. TF-IDF refers to a term weighting scheme in information retrieval (Salton and Buckley 1988) that accounts for the importance of a particular term (i.e., word) in a document (i.e., crowdsourced contribution). It measures the frequency of a term in a document normalized by the document length (TF) and multiplies this value with the inverse document frequency of the term (IDF). Generally speaking, a term that is frequently used in a contribution but rarely used in other contributions on the crowdsourcing platform will receive a high TF-IDF-value. In this way, it is possible to measure the novelty of the words in a contribution. We aggregated the TF-IDF-values for all terms in a contribution using the sum. Thus, contributions with a high TF-IDF-index include more novel information than contributions with a low TF-IDF-index. Related studies have already used aggregated TF-IDF-indices to analyze textual contributions in crowdsourcing (Rhyn & Blohm, 2017a; Zhang, Zeng, Wang, Breiger, & Hendler, 2016).

Novelty of Recombined Knowledge. In order to measure the novelty of the recombined knowledge, we analyzed the content of the discussions that developed around a contribution. For each contribution, the platform offered the possibility for other members of the crowd to add comments and provide their own perspectives or experiences. Thus, not only the initial contribution provides information for the organization but also the discussion that potentially combines the initial contribution with other perspectives and adds novel insights. Consistent with our previous measure, we calculated the average TF-IDF-indices of all comments for a contribution in order to measure the novelty of this recombined knowledge. Thus, a high TF-IDF-index suggests that, on average, the comments in the discussion added more novel information to a contribution than comments in a discussion with a low TF-IDF-index.

Scope of Recombined Knowledge. In order to measure the scope of the recombined knowledge, we analyze the topics that are being addressed in the contributions and their discussions. Topics can be interpreted as a distribution of words. A “design” topic, for example, will likely include words referring to colors or shapes and less likely include words referring to cars or trains. A “transportation” topic, on the other hand, will likely include words referring to cars or trains and less likely words referring to colors or shapes. We measure the distribution of the topics in a crowdsourced contribution and compare it to the average topic distribution on the crowdsourcing platform. Thus, contributions whose topic distribution is similar to the average topic distribution on the platform will combine a broad set of topics (i.e., large scope). Contributions whose topic distribution is less similar to the average topic distribution on the platform focus on only one or few specific domains (i.e., narrow scope). In order to calculate this measure, we use topic modeling based on the Latent Dirichlet Allocation (LDA) with a Gibbs sampler (Blei, Ng, & Jordan, 2003). LDA refers to a generative probabilistic model that can be used to automatically detect topics that are underlying a collection of text documents (i.e., contributions on the platform). The process for detecting and analyzing the topics in text documents includes two essential steps. First, we used all contributions and comments created by the crowd to uncover the topics that are present on the crowdsourcing platform. Based on the approaches proposed by Griffiths and Steyvers (2004) and Arun et al. (2010), we found 36 topics on our platform, which are outlined in more detail in the discussion section below. In a second step, it is possible to assign each individual contribution with probabilities for addressing each topic. The distribution of topics can be represented by a vector. We used the cosine similarity to calculate the similarity between the topic vector of a contribution with its discussion and the mean topic vector on the platform. The cosine similarity has been found to be a valid measure for the similarity of posterior distributions as retrieved in topic modeling (Niekler & Jähnichen, 2012). In this vein, the measure is also akin to the concept of information diversity as used in a related study by Riedl and Woolley (2017).

Control Variables. We use several additional variables to control for alternative effects that could influence the likelihood of a contribution being implemented by the organization. First, as shown by related literature (e.g., M. Li et al., 2016; Schemmann et al., 2016), popular ideas have a higher chance of being implemented by organizations than less popular ideas. Thus, on our platform, the voting behavior of the crowd might have influenced the decision of the administrators to implement a contribution. We account for this effect by including the number of votes per contribution as a control variable. Second, since we are analyzing cross-sectional data, older contributions have had more time to being discussed and noticed by the crowd or the administrators than contributions that have been submitted just recently. We control for this effect by measuring the time (in number of days) a contribution has been on the platform. Third, it is possible that a certain type of contribution is deemed more important and thus prioritized by the organization for implementation. On our platform specifically, the crowd was able to submit contributions that address two categories: ideas and problems (e.g., with using the organization's smartphone application or ticketing system). We control for the category in which the contribution was submitted by using a dummy variable (0 = idea, 1 = problem).

The descriptive statistics and the correlation matrix for the variables in our study are listed in Table 7 and Table 8. Most importantly, the correlations and the variance inflation factors for our independent variables, which range from 1.010 (for the novelty of the contribution) to 1.220 (for the votes), raise no concerns for multicollinearity.

	Mean	Std. Dev.	Min.	Max.
Dependent Variable				
Usefulness of Contribution	0.13	0.34	0	1
Independent Variables				
Network Ties of Contributor	5.13	23.71	0	178.75
Novelty of Contribution	2.65	0.88	0.27	6.87
Novelty of Recombined Knowledge	0.27	0.52	0	4.56
Scope of Recombined Knowledge	0.92	0.08	0.28	0.99
Control Variable				
Votes	2.77	9.79	0	218
Time	91.38	84.70	0	274
Category	0.36	0.48	0	1

Table 7: Descriptive Statistics I

Source: Own Illustration

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1) UC	1							
2) NTC	-0.04	1						
3) NC	0.07***	-0.11***	1					
4) NRK	0.07***	0.00	0.00	1				
5) SRK	-0.01	-0.08***	0.04*	-0.25***	1			
6) VT	0.15***	-0.01	0.04	0.14***	-0.19***	1		
7) TI	-0.29***	-0.03	-0.01	-0.04*	0.06**	-0.21***	1	
8) CAT	0.08***	0.01	0.02	0.12***	-0.08***	-0.16***	0.04*	1

Note: ***p<0.01; **p<0.05; *p<0.10; UC = Usefulness of Contribution, NTC = Network Ties of Contributor, NC = Novelty of Contribution, NRK = Novelty of Recombined Knowledge, SRK = Scope of Recombined Knowledge; VT = Votes; TI = Time, CAT = Category

Table 8: Correlation Matrix

Source: Own Illustration

5.3 Models and Results

Given that the dependent variable is binary, we use binary logistic regression to analyze the dataset and test our hypotheses. Our full model reads as follows:

$$P_{implemented} = \frac{1}{1 + e^{-(\beta_0+\beta_1 \cdot x_1+\beta_2 \cdot x_2+\beta_3 \cdot x_3+\beta_4 \cdot x_4+\beta_5 \cdot x_5+\beta_6 \cdot x_6+\beta_7 \cdot x_7+\varepsilon)}}$$

where β_0 represents the constant term, x_1 represents the effective network ties of the contributor, x_2 represents the novelty of the contribution, x_3 represents the novelty of the recombined knowledge, x_4 represents the scope of the recombined knowledge, x_5 represents the votes, x_6 represents the time, and x_7 represents the category. The β-coefficients can be interpreted as the change in log odds for a one unit change in the variables. The error term is represented by ε. The results of the logistic regression are listed in Table 9. We report the coefficients, the standard errors, and the Wald statistics to assess the significance of the coefficients. The Maximum-Likelihood-Estimation (MLE) method was used to estimate the coefficients. Since some contributions were created by the same person in the crowd, we ran our analysis with clustered robust standard errors to control for potential dependencies. This approach has been widely used by related studies to control for similar effects (e.g., M. Li et al., 2016; Schemmann et al., 2016).

	Est.	S.E.	Wald Z
Intercept	-3.526	1.030	-3.42***
H1: Network Ties of Contributor	-0.009	0.003	-3.45***
H2: Novelty of Contribution	0.205	0.082	2.49**
H3: Novelty of Recombined Knowledge	0.310	0.144	2.15**
H4: Scope of Recombined Knowledge	1.737	1.042	1.67*
Control: Votes	0.019	0.012	1.66*
Control: Time	-0.016	0.002	-9.76***
Control: Category	0.735	0.162	4.53***
Pseudo R-Sq.[1]	0.228		
Chi-Sq.	255.91*** (df = 7)		

Note: ***p<0.01; **p<0.05; *p<0.10; [1] Nagelkerke (1991)

Table 9: Results of the Logistic Regression

Source: Own Illustration

As shown in Table 9, the test of our logistic regression model against the constant-only model is highly statistically significant ($Chi^2(7) = 255.91$, $p < 0.00$). The Nagelkerke's

Pseudo R^2 amounts to 0.228 (Nagelkerke, 1991). All hypotheses with the exception of H4 are supported at highly significant levels of $p < 0.05$. For H4, we only find support at a confidence level of $p < 0.10$ but not at a confidence level of $p < 0.05$. The coefficients listed in Table 9 can be exponentiated and interpreted as odds ratios. The following paragraphs discuss these odds ratios and summarize our findings with regard to our hypotheses.

In H1, we hypothesized that contributions that are created by members of a crowd with few network ties are more likely to be useful for organizations than contributions created by members of a crowd with many network ties. Indeed, we find that the odds of a contribution being implemented decreases by 0.90% ($\beta = -0.009$; $p = 0.001$) for every one unit increase in a user's effective network size on the platform. In essence, this suggests that, as contributors become more immersed in the crowd and exchange a lot of information with other users, they become familiar with dominant perspectives on the platform and are less likely to come up with new contributions that are useful for organizations. The effect is statistically significant ($p < 0.05$).

In H2, we hypothesized that contributions that contain novel information are more likely to be useful for organizations than contributions that refer to already available information. Our results reveal that, for every one unit increase in the TF-IDF-index, the odds of a contribution being implemented increase by 22.75% ($\beta = 0.205$; $p = 0.013$). This shows that contributions which include a lot of terms that are rarely used in other contributions (i.e., contributions that offer novel information) are more likely to be useful for organizations than contributions that include rather common terms (i.e., contributions that refer to already existing information). The effect is statistically significant as well ($p < 0.05$).

In H3, we hypothesized contributions whose discussion adds novel information are more likely to be useful for organizations than contributions whose discussion refers to available information. We find that, for every one unit increase in the average TF-IDF-index for a discussion, the odds of a contribution being implemented increase by 36.34% ($\beta = 0.310$; $p = 0.031$). This goes to show that not only the initial contributions are important, but also the discussions that develop around them. A contribution becomes more useful for organizations when the crowd combines it with novel insights and further enriches the content. The effect is statistically significant ($p < 0.05$).

In H4, we hypothesized that contributions that combine information from a broad set of topics on the platform are more likely to be useful for organizations than contributions that refer to a narrow set of topics. On a confidence level of $p < 0.10$, it shows that if the distribution of the topics in a contribution perfectly matches the broad distribution of topics on the whole platform (i.e., reach a perfect value of 1), the contribution would be 5.68 times more likely to be implemented than contributions whose topic distribution refers to a more narrow set of topics ($\beta = 1.737$; $p = 0.096$).

The results presented in Table 9 also illustrate the importance of the control variables. Both the control variable for the time effect and the control variable for the category are highly statistically significant. The control variable for the votes is statistically significant at a confidence level of 0.10. Thus, contributions with a higher number of votes are more likely to be implemented than contribution with a lower number of votes. Consistent with prior work (e.g., M. Li et al., 2016; Schemmann et al., 2016), this indicates that the voting behavior of the crowd influences the decision of organizations to implement ideas. Second, recently submitted contributions are less likely to be implemented by the organization than older contributions. This can be explained by the fact that older contributions have had more time to be discussed, enriched, and selected by the administrators than recently submitted contributions. Furthermore, it takes time to implement contributions. These findings are also consistent with prior studies (e.g., Schemmann et al., 2016). Table 9 also shows that it is important to control for the category of the contribution. Contributions that address problems (e.g., ticket purchasing system not working, crashes of the organization's mobile application) are more likely to be implemented by the organization than ideas. Finally, we also conducted an additional robustness check and reran our model using centrality measures (i.e., the degree centrality) instead of the effective network size as an indication of an individual's connections in the crowd. The results remain consistent and show that contributions created by members of a crowd with a low degree centrality are more likely to be useful for organizations than contributions created by members of a crowd with a high degree centrality.

5.4 Discussion of the Emergence of Useful Ideas in Crowdsourcing

The results of our study yield several important findings. First, they indicate that useful contributions typically originate from individuals with only few network ties in the crowd and that these contributions generally introduce new information to the platform. Both effects are statistically significant and correspond to theoretical insights from the fields of problem solving and network theory. These effects suggest that contributions created by members of a crowd who are not already immersed in the network offer novel insights and perspectives. Thus, they are especially useful for organizations in their search for new knowledge on crowdsourcing platforms. Similar findings have already been documented by Jeppesen and Lakhani (2010) who note that that technical and social marginality are positively related to problem-solving success in crowdsourcing. However, the results of our study also emphasize the effects of collaboration and knowledge sharing on crowdsourcing platforms that foster interaction between members of crowd. As members of a crowd become more familiar with both their peers and the organization, an increased exchange of mostly local information seems to lead to a convergence of perspectives or cognitive schemata that ultimately stifles their creativity and problem-solving capabilities (Perry-Smith & Shalley, 2003). It is also likely that the problem representations of members in the crowd become more confined and concrete as their knowledge and expertise in the crowd grows. Such effects have typically been found to impede diverse solutions and innovation (Förster et al., 2004; Marsh, Ward, & Landau, 1999). Another potential explanation is given by related research suggesting that users typically have only a few truly innovative ideas or solutions to offer before they begin to generate less innovative contributions and provide redundant information (Bayus, 2013; Von Hippel, 2005). Our findings hint that such effects can also be observed on crowdsourcing platforms, which are often specifically used or created by organizations for the purpose of searching for distant knowledge (Afuah & Tucci, 2012). Organizations should be aware that, especially when hosting their own crowdsourcing platforms with dedicated crowds, increased participation, feedback, and peer-voting systems may quickly lead to bounded rationalities and local knowledge bases. Thus, it might not always be recommended to rely exclusively on rating scales or preference markets in crowdsourcing when searching for distant knowledge (cf. Blohm et al., 2016), as they might be biased by these bounded rationalities on the platforms. As shown by our results, both the structural origin of a contribution in a crowd as well as the textual characteristics of its content can be used as additional measures to identify potentially useful contributions in crowdsourcing. We argue that organizations searching for useful contributions should pay special attention to novel information provided by new members in a crowd who have not yet made strong connections to representatives of the organization or peers within the crowd.

Second, our study also reveals important insights on how knowledge is absorbed, recombined, and enriched by crowds on crowdsourcing platforms. In traditional group settings, some studies suggest that information diversity and the integration of different perspectives may derail discussions or lead to coordination problems. Cronin and Weingart (2013), for example, argue that diversity bears the risk of making it more difficult for groups to develop a shared understanding and communicate efficiently and effectively. Similarly, van Knippenberg et al. (2004) show how group diversity can lead to social categorizations. In the case of crowdsourcing, however, we found organizations to greatly benefit from different topics or perspectives coming together on the platforms. Consistent with related research (e.g., Faraj et al., 2011; Ye et al., 2016), our findings suggest that (re)combining different perspectives enables crowds to enrich and further develop ideas or solutions for organizations. On our platform specifically, we detected 36 topics that are underlying the contributions and comments created by the crowd. As the platform is hosted by an organization in the transportation and logistics sector, the topics relate to rather particular aspects in this domain, such as “tickets”, “connections”, “calendar”, or “time tables”. Across the platform, we note a relatively even distribution of the topics with around 3% per topic. Our results suggest that contributions whose discussion resembles this distribution and combines a diverse set of topics are more likely to be implemented than contributions whose discussion focuses on only one or few specific topics. For example, in one of the most commented, implemented idea, a member of the crowd suggested that the smartphone application, which the transportation and logistics provider offers to its customers, should automatically track the time and location when travelling by train. The user argued that this would make it easier to check for connections and platforms of departure and arrival. The idea was then taken up by other members of the crowd who linked the suggestion to different topics and became creative, for example, by proposing how the design should look like or how push notifications may be used for even more convenience. Furthermore, it was also discussed whether the idea is technically feasible and how it could be implemented. In addition, we do find statistically significant evidence that contributions whose discussion fosters new information and make the crowd become creative are more likely to be useful for organizations than contributions whose discussion is only based on existing information.

Taken together, this suggests that organizations searching for new knowledge on crowdsourcing platforms should not only pay attention to the contributions themselves, but also how these contributions are being discussed and developed by the crowd. Contributions provided by individuals with few network ties are likely to introduce new information and perspectives which can then be combined or recombined in a creative process with local knowledge by already experienced and immersed members in the crowd. Our findings suggest that when information or insights from different perspectives and topics are brought together and combined on crowdsourcing platforms, the related contributions and the information exchange around these contributions are especially useful for organizations. These findings also resonate with insights from search theory emphasizing the role of knowledge combination and recombination for organizations (e.g., Fleming, 2001; Fleming & Sorenson, 2004). Thus, our results entail a number of important theoretical and practical implications.

5.4.1 Theoretical Implications

From a theoretical perspective, we are able to contribute novel insights for both search theory and research on crowdsourcing. In recent years, a large body of literature on search theory has emphasized the importance of spanning organizational and technological boundaries in order to find distant knowledge (Katila & Ahuja, 2002; Katila et al., 2012; Rosenkopf & Nerkar, 2001). The interaction with external sources has been deemed especially promising for coming up with innovative ideas or solving existing problems in organizations (e.g., Chesbrough, 2003). As a result, novel, IT-facilitated approaches for distant search, such as crowdsourcing, have emerged (Afuah & Tucci, 2012). Crowdsourcing platforms offer organizations an interface to access knowledge distributed amongst large and diverse networks of people. They serve as focal points in this distributed network at which knowledge emerges and evolves.

Based on our findings, however, we argue that simply engaging in crowdsourcing or building crowdsourcing platforms in an attempt to span organizational boundaries may not suffice to find new and distant knowledge. The result presented in this study suggest that, even on crowdsourcing platforms, increased participation, feedback, and information sharing may lead to dominant schemata and mindsets which, ultimately, create local knowledge bases. Similar effects have already been described by Laursen (2012) and Christensen (1997) who argue that searching across organizational boundaries not always implies distant search, since existing customers require organizations to follow established trajectories – even when novel opportunities emerge. We find that it is important to differentiate the contributions with regard to their origin and their content when searching for distant knowledge on crowdsourcing platforms.

Furthermore, the results of our study underline the importance of knowledge recombination for distant search (cf. Fleming, 2001; Fleming & Sorenson, 2004; Hargadon & Sutton, 1997) and extend its understanding in crowdsourcing. Our findings suggest that useful ideas and solutions not only emerge from isolated contributions alone but from a combination of different topics brought together by members of a crowd. Distant contributions from new users are likely to trigger such discussions and make the crowd mobilize its knowledge and become creative. As much as novel inputs and perspectives are important in crowdsourcing for the elicitation of distant knowledge, as much is the additional development and discussion of these inputs by experienced users in the crowd important. These individuals are familiar with the already accumulated knowledge base on the platform and may combine novel ideas or solutions with existing elements. Prior research has already discussed that combinations of existing knowledge elements (depth) with new ideas (scope) are likely to create unique solutions that can be commercialized (Katila & Ahuja, 2002; Winter, 1984). It shows that crowdsourcing can be employed as a mechanism to orchestrate these processes with knowledge distributed amongst large networks of people.

Finally, from a network perspective, our study also touches upon the classic problem of weak and strong ties in the context of innovation and knowledge sharing (Granovetter, 1973, 1983). Our results suggest that novel information is likely to be solicited from members of the crowd who are not yet highly embedded and connected in the network but may bring diverse perspectives from outside the community to the platform. This is consistent with findings presented in earlier research conducted by Hargadon and Sutton (1997) and Lingo and Mahony (2010). Individuals who are not (yet) connected to redundant sources of information may serve as knowledge brokers and draw analogies or introduce new knowledge to a particular field or network. Hence, they have unique informational benefits compared to those who are structurally central and immersed in the network (Lingo & O'Mahony, 2010). However, when complex or diverse knowledge needs to be shared and further developed, strong ties become more advantageous. Hansen (1999), for example, shows that weak interunit ties help project teams to search for useful knowledge in other subunits but impede the transfer of complex knowledge, which requires rather strong ties. In crowdsourcing, we found similar effects. New and useful knowledge initially emerges from ideas or solutions provided by members of the crowd who have not yet established a large number of connections in the network and are able to provide novel information to the crowd. Immersed users, on the hand, may add to these contributions by integrating and transferring complex and local knowledge through discussions.

5.4.2 Practical Implications

There are a number of practical implications that can be drawn from our results. First, our result show that useful ideas and solutions are often created by new members in a crowd with only few effective network ties. While it is common for organizations to pay special attention to lead users and follow suggestions that are popular in their established networks, we propose that organizations seeking to gain access to distant knowledge through crowdsourcing should rather focus on contributions that are generated by less immersed members of the crowd. Especially on well-established platforms with stable crowds, using rating scales and voting systems alone may not be the most appropriate mechanisms to identify such contributions. Instead, integrating crowdsourcing platforms with novel business analytics that offer capabilities for social network analysis, text mining, and topic modeling could bridge this gap. Based on our findings, we see great potential in these technologies to support organizations in tracking the origin of contributions in crowdsourcing, analyzing their textual characteristics, and ultimately identifying the most innovative ones. We propose that organizations should delve deeper into such possibilities and make use of novel business analytics or decision support systems when engaging in crowdsourcing.

Second, our results also offer valuable implications for managing crowds and collaboration on crowdsourcing platforms. The results of this study not only emphasize the importance of an initial idea but also the importance of the discussion that unfolds around these contributions. We show that contributions are more likely to be useful for organizations when they spur an exchange of diverse information and when they combine a broad set of topics. This is especially relevant for crowdsourcing intermediaries or organizations that host their own platform. Based on our findings, we urge organizations to encourage the exchange of information between different members of the crowds and integrate incentives for collaboration. We find crowdsourced contributions to be especially useful for organizations when they combine information from a broad range of topics and when the members of a crowd become creative. As unfamiliar and novel information may trigger this creative process, we also suggest that organizations should actively advertise ideas or solutions generated by newer users on the platform, for example, by using recommender systems that promote particularly innovative ideas for further discussion. In line with our first practical implication, not only the most popular and frequently voted contributions should be endorsed by organizations, but also contributions that deviate from well-established patterns.

Finally, our study may serve as the starting point for developers of such business analytics or recommender systems to design and customize related models on crowdsourcing platforms. We provide a set of variables that have been found to be statistically significant predictors for innovative contributions in our study. These variables are based on relatively simple measures, such as the users' effective network size in social network analysis or the TF-IDF-index in information retrieval. They may be used as a foundation or addition for predictive modeling. We thus encourage practitioners to build upon our findings and develop more sophisticated algorithms or models to facilitate the evaluation of large amounts of contributions on crowdsourcing platforms with related business analytics or decision support systems.

5.4.3 Limitations and Outlook

As with all research, the results and implications presented in this study should be regarded in light of its limitations. First, we analyze crowds from a network perspective in this study. The effects described and examined in our study are inherently based on interactions that unfold on crowdsourcing platforms. Hence, an important boundary condition of this study is that there are connections between members of a crowd and that crowds can be treated as networks of people. As for platforms or crowdsourcing settings with little interaction between contributors, our findings might be less applicable. For future research, it would thus be interesting to further study how different forms of collaboration or interaction on crowdsourcing platforms affect the emergence and recombination of distant knowledge. Our findings should be viewed as initial insights from a collaborative crowdsourcing setting.

Second, this study examines crowdsourcing in an organizational setting. Thus, we focus on a context that leverages crowdsourcing as an approach for distant search in order to harness knowledge outside existing, organizational boundaries (Afuah & Tucci, 2012). Furthermore, it must be noted that our results are based on data retrieved from a crowdsourcing platform hosted by a transportation and logistics provider. While the characteristics of the dataset and the platform are similar to those reported in related studies (e.g., M. Li et al., 2016; Schemmann et al., 2016), there is still the possibility that our findings may not apply to every other industry or application of crowdsourcing to the same extent (e.g., for local governments; see Masdeval & Veloso, 2015). In this sense, additional data from more diverse crowdsourcing platforms would greatly benefit the generalizability of our findings. Especially with regard to topic distributions, we expect that platforms hosted by intermediaries who specifically target a broad and diverse crowd may benefit much more from a combination or recombination of knowledge brought together from different domains. In our study, this effect could not be empirically supported at a significance level of 0.05 but only the 0.10 level.

Third, our results are based on a cross-sectional study of the data. Thus, we measured the characteristics of the contributions at a particular point in time on the platform. While cross-sectional analyses are also commonly applied in related studies (e.g., M. Li et al., 2016; Schemmann et al., 2016), they provide only a static perspective on the dataset and the underlying effects. Another interesting perspective on where and how distant knowledge emerges in crowdsourcing or other distributed networks may be achieved by conducting longitudinal studies or using survival analysis. This dynamic perspective represents a promising avenue for future research to analyze in more detail how knowledge or topics emerge and evolve on crowdsourcing platforms over time.

Fourth, future research may examine in more detail how crowdsourcing platforms and evaluation processes for crowdsourced contributions should be designed. Based on this study, we argue that rating scales and voting systems may not be the most adequate mechanisms for identifying innovative contributions as their results are prone to being biased by already familiar perspectives and popular opinions amongst experienced members of the crowds. Thus, it would be interesting for future research to investigate how IT-supported processes with systems capable of tracking the origin of crowdsourced contributions and analyzing their textual characteristics may support organizations in their evaluation of large amounts of ideas and solutions on their platforms. From a theoretical perspective, we also urge future research to delve deeper into the possibilities of assessing and measuring innovative contributions in crowdsourcing. For example, studies may use textual features, similarity measures, or redundancy indices as proxies for the innovativeness of contributions on crowdsourcing platforms.

5.5 Conclusion of Chapter

Crowdsourcing represents a powerful approach for organizations to engage in distant search and mobilize knowledge distributed amongst a diverse network of people. While organizations generally succeed in generating large amounts of new knowledge on crowdsourcing platforms, it represents a latent challenge to find useful contributions that have the potential to actually solve problems or serve as innovation. In existing literature, little attention has been paid to the emergence and evolution of new knowledge on crowdsourcing platforms and how organizations may identify contributions that capture such knowledge. In this study, we address this gap by analyzing cross-sectional data from a large crowdsourcing platform in Europe and combining statistical approaches from the fields of network analysis and information retrieval to empirically test a set of hypotheses. We find that new and useful contributions are typically created by individuals who have not (yet) established a large number of network ties in the crowd. They have unique informational benefits compared to those who are already immersed in the network. However, their contributions become especially useful when they are further enriched and combined with local knowledge provided by experienced members on the platforms. For researchers in the fields of distant search, knowledge collaboration, and crowdsourcing, we provide a more thorough understanding on how network relationships and information sharing affect the emergence and evolution of knowledge on crowdsourcing platforms. From a practical perspective, we offer guidance for crowdsourcing intermediaries or organizations that host their own crowdsourcing platforms on how to identify potentially useful contributions in the vast pool of data generated by their crowds. We see great opportunity for business analytics to support organizations in tracking the origin of contributions in crowdsourcing, analyzing their textual characteristics, and ultimately identifying the most useful ones. In light of these insights, organizations may leverage distant search on crowdsourcing platforms to its fullest extent.

6 THE EFFECTS OF COLLABORATION IN CROWDSOURCING

This chapter[5] addresses the second research question of the dissertation. The focus lies on improvements to the effectiveness of decision making. It extends the findings presented in chapter 5 and examines how collaboration affects the creative performance of individuals in crowds over time. In this way, the findings may help to better understand how collaboration in crowds affects the emergence of valuable contributions. The chapter is organized as follows: First, section 6.1 explains the motivation and objectives of the study in more detail. Second, section 6.2 outlines the underlying hypotheses regarding the effects of collaboration on an individual's creative performance in an online crowd. Third, section 6.3 explains the methodology to test these hypotheses and reveals the results. Fourth, section 6.4 discusses the implications of the results for research and practice. Finally, section 6.5 concludes the chapter with a short summary.

6.1 Understanding the Effects of Collaboration in Crowdsourcing

Spurred by the growing relevance of crowdsourcing in recent years, there have been increasing efforts in research to understand how collaboration affects the creativity of individuals in online crowds (e.g., Bayus, 2013; Faraj et al., 2011; Huang et al., 2014; Majchrzak & Malhotra, 2016; Ransbotham, Kane, & Lurie, 2012). However, existing findings in this field are mixed. A large group of studies provide evidence that collaboration in crowds serves as a crucial prerequisite for eliciting creative solutions to innovation problems (e.g., Faraj et al., 2011; Majchrzak & Malhotra, 2016; Von Hippel, 2005). The general consensus found in these studies is that creative outcomes are likely to occur when perspectives from distinct fields are brought together, discussed, and jointly developed in a crowd (Faraj et al., 2011). A second group of studies take a difference stance and suggest that collaboration is prone to hamper creativity (e.g., Perry-Smith & Shalley, 2003; Stephen et al., 2016). According to related research, collaboration may lead to a convergence of mental representations (Perry-Smith & Shalley, 2003), an assimilation to popular ideas (Stephen et al., 2016), and a fixation on dominant themes (Goldenberg, Lehmann, & Mazursky, 2001). A third group of studies present evidence that collaboration is beneficial only up to a certain threshold, after which the positive effects begin to reverse (Uzzi & Spiro, 2005).

Given these mixed results, it is difficult to draw a comprehensive picture of how collaboration affects an individual's creative performance and his or her ability to solve innovation problems in crowds. In this study, we address this gap and answer the following

[5] This chapter builds upon prior research published in Rhyn et al. (2017). It extends the findings with a new study that includes new hypotheses, a new research model, new data, and new contributions for the dissertation.

research question: How does collaboration affect an individual's creative performance in an online crowd? We argue that, as individuals collaborate, they are exposed to perspectives of their peers and become assimilated to dominant paradigms in a crowd. Our core hypothesis is that this assimilation to dominant paradigms changes the manner in which individuals become creative and are able to solve innovation problems with their ideas. We refer to Nagasundaram and Bostrom's (1994) notion of *paradigm-relatedness* and hypothesize that, over time, collaboration decreases individuals' ability to create value through paradigm-modifying (PM) ideas but increases their ability to create value through paradigm-preserving (PP) ideas. The former represent radical shifts in a paradigm while the latter represent gradual improvements in a paradigm. In other words, we believe that, over time, collaboration makes individuals gradually change from *innovators,* who perform well at challenging paradigms and introducing novel solutions through their ideas, to *adaptors,* who perform well at refining paradigms and building upon existing solutions with their ideas (Nagasundaram & Bostrom, 1994).

We test our theory with a unique longitudinal dataset that captures over 8 years of activity by a crowd of 7'832 individuals who developed 222'259 ideas and comments to solve 476 innovation problems issued by different organizations. We combine statistical approaches from the fields of network analysis and text mining to examine how individuals in the crowd collaborated and how the content of their ideas relates to paradigms in the crowd. Based on a multilevel mediation analysis, we study the effects of collaboration on an individual's ability to create paradigm-modifying and paradigm-preserving ideas and, thus, on their creative performance in solving innovation problems. Our results show that, over time, collaboration (1) decreases an individual's ability to make paradigm-modifying ideas but (2) increases his or her ability to make paradigm-preserving ideas. That is, collaboration makes individuals change from *innovators* to *adaptors* over time. Importantly, however, both innovators and adaptors are able make valuable contributions to innovation problems in their respective way.

These results have a number of important theoretical implications for research on creativity, knowledge collaboration, and crowdsourcing. First, we contribute to creativity research by explaining how collaboration changes the manner in which individuals are able to operate within – or challenge – paradigms and thus become creative when embedded in an online crowd. Second, we contribute to research on knowledge collaboration by unraveling the effects of collaboration on problem solving in creative tasks. Our findings show two opposing effects that unfold over time and may help to bridge the gap between hitherto mixed findings presented in related research. Third, we contribute to research on crowdsourcing by identifying important limitations of this approach for developing innovations. We show why crowds are prone to develop dominant paradigms over time and thus become less capable of coming up with truly innovative, paradigm-modifying ideas in the long run.

6.2 Development of Hypotheses

For the development of our hypotheses, we draw upon prior work on knowledge collaboration and creativity research. *Knowledge collaboration* in this context is defined as "individual acts of offering knowledge to others as well as adding to, recombining, modifying, and integrating knowledge that others have contributed" (Faraj et al., 2011, p. 1224). While in a pairwise interaction this would be referred to as a conversation (Tsoukas, 2009), online crowds rely on temporal strings of posts on IT-based platforms that form "knowledge-sharing trajectories" (Majchrzak & Malhotra, 2016). Extending, recombining, modifying, and integrating knowledge in these trajectories means that individuals collaborate and become connected in a network, i.e., in a crowd (Stephen et al., 2016). The goal of these networks in crowdsourcing is to jointly solve an organizational problem formulated as a task.

Problem solving in such task settings involves the construction of an internal representation of the problem and the application of an appropriate heuristic to search for potential solutions (Dunbar, 1998; Jeppesen & Lakhani, 2010). Existing literature suggests that there is great variance between individuals with regard to how problems are represented and solved (Jeppesen & Lakhani, 2010). It has been found that individuals overwhelmingly use familiar knowledge and prior experience in a particular domain for developing solutions to problems they encounter (Lovett & Anderson, 1996; Lüthje et al., 2005). That is, their problem solving reflects particular assumptions and paradigms of a domain.

Creativity manifests itself in approaches to formulating and solving problems (Kirton, 1976) and is thus often examined in terms of "paradigm-relatedness" (Dean, Hender, Rodgers, & Santanen, 2006; Nagasundaram & Bostrom, 1994). *Paradigm* refers to "the prevailing ways of perceiving and acting in a given situation or problem", while *relatedness* refers to "the extent to which an idea operates within or challenges that paradigm" (Garfield, Taylor, Dennis, & Satzinger, 2001; Satzinger, Garfield, & Nagasundaram, 1999; Silk, Daly, Jablokow, & McKilligan, 2019, p. 31). Solving problems largely within a paradigm describes a *paradigm-preserving* way of problem solving. Solving problems with radical shifts in a paradigm describes a *paradigm-modifying* way of problem solving (Nagasundaram & Bostrom, 1994; Silk et al., 2019). Nagasundaram and Bostrom (1994) emphasize that the degree of paradigm preservation or modification is not a measure of the degree of creativity but rather a description of the manner in which creativity manifests itself; it does not explain how creative individuals are but how individuals are creative. Both paradigm-modifying and paradigm-preserving forms of creativity can be valuable for organizations and complement each other. While the former can be important for exploring and formulating new opportunities for organizations, the latter often ensure that those opportunities are implemented and monitored efficiently (Nagasundaram & Bostrom, 1994).

Collaboration has been found to greatly affect creativity in problem solving (e.g., Hargadon & Bechky, 2006; Ransbotham et al., 2012; Uzzi & Spiro, 2005). By collaborating, individuals are exposed to perspectives and alternatives provided by peers. Exposure to diverse perspectives and alternatives has been found to elicit wider categorizations and more divergent solutions (Kanter, 1988; Perry-Smith & Shalley, 2003). It offers a basis for individuals to transform information, combine it with their own experience, and create new knowledge (Carlile & Rebentisch, 2003; Ransbotham et al., 2012). Studies show that access to information and the ability to transform knowledge can be particularly beneficial for creative tasks, given that the extension, modification, and recombination of concepts from existing knowledge represent the most common processes through which novel solutions are formed (Dahl & Moreau, 2002; Goldenberg, Mazursky, & Solomon, 1999; Stephen et al., 2016). Thus, as individuals collaborate, the greater the access to information and the greater the opportunity for an exchange and (re-)combination of knowledge (Ransbotham et al., 2012).

Our core argument in this study, however, is that collaboration changes the manner in which individuals are able to operate within – or challenge – paradigms and, thus, affects their creative performance over time (see Figure 7 below). That is, we believe collaboration has opposing effects on an individual's ability to create value through paradigm-modifying ideas and paradigm-preserving ideas.

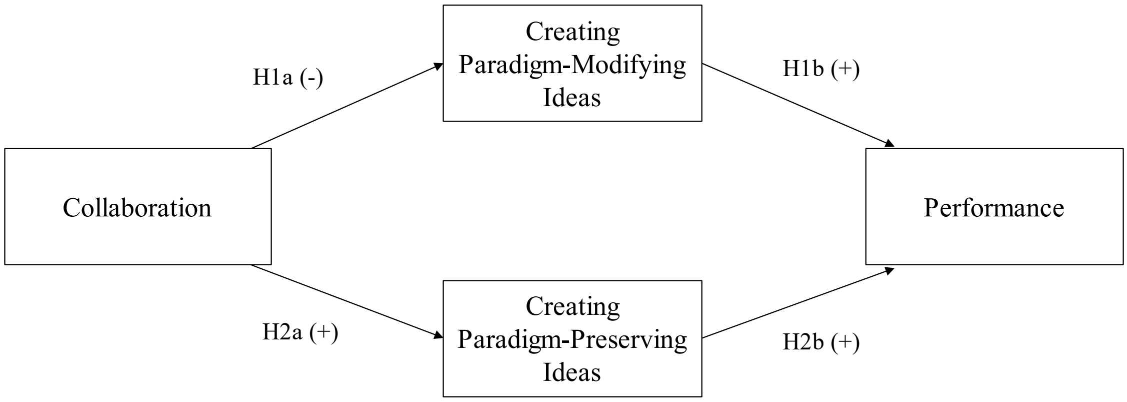


Figure 7: Research Model II

Source: Own Illustration

6.2.1 Paradigm-Modifying Ideas

Individuals that challenge prevailing ways of solving a particular problem operate in a *paradigm-modifying* way. When confronted with a problem, they attempt "to restructure the problem by approaching it from a new angle, thus breaking the customary starting point for its solution" (Kirton & De Ciantis, 1986, p. 141). In collaborative settings, they typically create value by either introducing new elements into a problem context or altering the relationships between such elements (Kirton, 1976; Nagasundaram & Bostrom, 1994). Kirton (1976) refers to these individuals as "innovators".

Studies suggest that individuals who are technically and structurally distant from a domain are more likely to act as innovators and face problems in a paradigm-modifying way than those already familiar with the domain (e.g., Gieryn & Hirsh, 1983; Jeppesen & Lakhani, 2010). They are naive with regard to the prevailing paradigms in the domain and have access to differing knowledge (Gieryn & Hirsh, 1983). Thus, they are able to offer perspectives and heuristics that are novel and innovative for solving a problem (Jeppesen & Lakhani, 2010). Similarly, studies show that being remote to a problem and thinking in abstract terms may lead to more diverse and original solutions than thinking in concrete, technical terms, which often impedes innovation (Förster et al., 2004).

With increased collaboration, however, mental representations tend to converge in crowds due to common experiences and greater information exchange (Perry-Smith & Shalley, 2003). As individuals become immersed in a particular network, they become assimilated to their direct ties that provide mostly conformant information. With ties becoming stronger in a network, this conformity is argued to hamper creativity (Perry-Smith & Shalley, 2003). Related research also argues that assimilation to direct ties and increasing conformity in networks makes it more likely for individuals to fixate on prevalent ideas, themes, or concepts at the expense of more radical, distant ones – i.e., they focus on ideas, themes, or concepts that relate to dominant paradigms in their network (Goldenberg et al., 2001; Stephen et al., 2016). Thus, perspectives drawn from distant positions in a network are typically argued to be more novel relative to existing standards within a domain (Perry-Smith & Shalley, 2003).

Following these findings, we argue that, over time, collaboration will have a negative effect on an individual's ability to come up with paradigm-modifying ideas. We believe individuals in a crowd to initially exhibit "a useful ignorance of prevailing assumptions and theories" (Gieryn & Hirsh, 1983, p. 91), that makes them likely to create value through paradigm-modifying ideas. Such paradigm-modifying ideas are useful for solving organizational problems on crowdsourcing platforms (Jeppesen & Lakhani, 2010). As individuals collaborate with peers, they become assimilated to prevailing assumptions and theories and their solutions relate more closely to dominant paradigms in the crowd. We hypothesize the following:

H1a. Collaboration has a negative effect on an individual's ability to create paradigm-modifying ideas in crowds.

H1b. An individual's ability to create paradigm-modifying ideas is positively associated with the individual's creative performance.

6.2.2 Paradigm-Preserving Ideas

While collaboration may have a negative effect on an individual's ability to create paradigm-modifying ideas, a large body of literature suggests that becoming familiar with paradigms and gaining domain-relevant knowledge may also entail positive effects for problem solving. Such positive effects may be particularly related to paradigm-preserving ideas. Individuals that follow prevailing ways of solving a particular problem operate in a *paradigm-preserving* way. When confronted with a problem, they turn "to conventional procedures and consensus of the group to which they belong, and derive their ideas towards the solution of the problem from established procedures" (Kirton & De Ciantis, 1986, p. 141). In collaborative settings, they typically create value through refinements and improvements to solutions (Nagasundaram & Bostrom, 1994; Silk et al., 2019). Kirton (1976) refers to these individuals as "adaptors".

Domain-relevant knowledge, which is required for – or at least benefits – such refinements and improvements, refers to "an individual's knowledge of facts, circumstances, and issues surrounding a given problem or area" (Perry-Smith & Shalley, 2003, p. 91). It involves technical expertise necessary to develop feasible solutions to a given problem and is argued to increase the ability to generate potential solutions and validate them with regard to their appropriateness (Amabile, 1996; Perry-Smith & Shalley, 2003). A large number of studies outline that, as individuals gain domain-relevant knowledge and expertise, their solutions are associated with higher expected quality (e.g., Larkin, McDermott, Simon, & Simon, 1980; Magee, 2005) and they show superior task performance (Alba & Hutchinson, 1987). Particularly for creating user-generated content in crowdsourcing, Ransbotham et al. (2012) suggest that gaining domain-relevant knowledge and experience through collaboration may benefit individuals in identifying and transforming valuable information into useful formats (Spence & Brucks, 1997), transferring relationships among different trajectories in ways that make content more informative (Gregan-Paxton & John, 1997), and providing more comprehensive information (Alba & Hutchinson, 1987).

In line with these findings, we argue that, over time, collaboration will have a positive effect on an individual's ability to create valuable, paradigm-preserving ideas. We believe that individuals will become familiar with the prevailing paradigms and assumptions in a crowd and thus become better at refining and recombining existing knowledge. In this way, collaboration can have a positive effect on an individual's creative performance in a crowd. We hypothesize the following:

H2a. Collaboration has a positive effect on an individual's ability to create paradigm-preserving ideas in crowds.

H2b. An individual's ability to create paradigm-preserving contributions is positively associated with the individual's creative performance.

6.3 Methodology of the Study

To empirically test our hypotheses, we conduct a multilevel mediation analysis. We combine statistical approaches from the fields of network analysis and text mining to analyze longitudinal data from a crowdsourcing intermediary in Europe.

6.3.1 Dataset

We used a unique data set retrieved from one of Europe's largest crowdsourcing platforms in the innovation domain. The platform is operated by a crowdsourcing intermediary. On the platform, organizations can organize campaigns and invite volunteers in the crowd to solve their innovation problems. These innovation problems typically require the crowd to jointly develop a new product or service. Solutions to these innovation problems submitted as "ideas", which are then collaboratively developed in strings of posts (i.e., textual contributions) that form "knowledge trajectories" for the organizations. After each campaign, the trajectories generated by the crowd are evaluated by the organizations who issued the innovation problem and organized the campaign on the platform. The organization is responsible for selecting the most useful ideas for implementation.

The data contains more than eight years of longitudinal data from the platform, capturing all activity in the crowd from November 2008 until September 2017. During this time, 476 innovation campaigns by different organizations were conducted. The crowd comprised a total of 25’024 registered individuals. Since there were no prerequisites for joining the crowd, it represents an open crowd (Zwass, 2010) that includes potentially everyone who is willing to make contributions to a knowledge trajectory for a particular innovation problem. Of all registered individuals, 7’832 were active and made at least one contribution to a trajectory. A total of 222’259 textual contributions were made to knowledge trajectories and 6’933 trajectories were selected for implementation by the organizations.

We used social network analysis (SNA) to study how individuals in the crowd collaborated and with whom they worked on solutions. For the SNA, we constructed a social network of the crowd and updated the network each week from November 2008 until September 2017. In this way, we are able to capture the collaborative ties of each individual in the crowd and analyze changes over time. Each node in the social network represents an individual in the crowd. A tie between two individuals in the crowd is created when they collaborate, i.e., when they make comments on each other’s ideas and jointly develop a solution in a knowledge trajectory.

In order to analyze the type of information that the individuals generated and discussed, we use text mining to statistically examine the 222’259 textual contributions forming the knowledge trajectories. For the analysis, we followed the commonly used Bag-of-Words approach and applied standard preprocessing steps to make the unstructured, user-generated content compatible for text mining algorithms (cf. Feldman & Sanger, 2007). We only applied minimal transformations to the texts and limited the preprocessing to tokenization (i.e., breaking up the texts into individual terms) and normalization (i.e., transforming all characters to lower-case) in order to keep as much data as possible in its original form.

The dataset is suitable for our analysis for several reasons. First, it stems from an intermediary platform and therefore includes data from a broad set of 476 different innovation campaigns and organizations, which mitigates concerns of a potential sample bias. Second, as representatives of the organization use the platform themselves to evaluate the trajectories, we have objectively measurable labels on how well individuals have performed on the innovation problems. Third, we note that the characteristics of our dataset are comparable to those reported in related studies (e.g., Bayus, 2013; M. Li et al., 2016; Piezunka & Dahlander, 2015; Schemmann et al., 2016). This should benefit the generalizability of our results.

6.3.2 Variables and Measures

Based on the available data from the crowdsourcing platform, we constructed an unbalanced panel data set. Each time an individual *i* posted a new idea in a given week *t*, we measured (1) the extent to which the individual has previously collaborated, (2) the degree of paradigm-modification of the individual's idea, (3) the degree of paradigm-preservation of the individual's idea, (4) and the performance of the individual in solving an innovation problem. The independent variable is collaboration. The dependent variable is the individual's performance in solving innovation problems. The mediators are the ability to create paradigm-modifying ideas and the ability to create paradigm-preserving ideas. Data were collected from November 2008 to September 2017.

Performance. Since we aim to examine how collaboration affects an individual's performance in solving a given innovation problem, we measure whether the idea was deemed valuable by the organization for solving the problem. Here, the implementation of the idea is a common and reliable indicator for its value that has been used extensively in related research (e.g., Jeppesen & Lakhani, 2010; M. Li et al., 2016; Piezunka & Dahlander, 2015; Schemmann et al., 2016). The implementation of an idea shows that an individual performed well in solving an organizational problem. In our case, the decision to implement an idea is made by representatives of the organizations (i.e., the problem owners) who are conducting the innovation campaigns on the platform. They manually review and evaluate the proposed ideas and assign a status to it depending on their decision (0 = idea was not implemented; 1 = idea was implemented). Thus, for each observation in the data set (individual *i* created a new idea on how to solve a given problem in week *t*), we have a reliable measure of performance.

Collaboration. We use the effective ego network size (Borgatti, 1997; Burt, 1995) of individual *i* in week *t-1* to measure the extent to which the individual has collaborated prior to creating a new idea. The effective network size measures the number of collaborative ties in an individual's network (i.e., an *ego network*) and discounts for redundancy (Burt, 1995). Redundancy measures the extent to which an individual's contacts are connected to each other as well and makes it possible to account for the structure of the ego network, i.e., for local clusters (Borgatti, 1997). Individuals who have extensively collaborated with many different peers have large effective network sizes while those who have not extensively collaborated (or only in local clusters) have smaller effective network sizes. For non-valued, undirected graphs, an individual's effective network size can be calculated as follows (see Borgatti, 1997):

$$effective\ network\ size = v - \frac{2e}{v}$$

where e is the number of ties in the ego network and v is the number of nodes in the ego network.

Paradigm-Modifying Ideas. In order to measure the extent to which an individual creates a paradigm-modifying idea, we analyze the topics addressed in the idea made by an individual i in week t. A topic can be interpreted as a distribution of words that are statistically related to each other. A topic distribution (i.e., a vector of probabilities) indicates the degree to which an idea relates to a given set of topics. The extent to which an individual introduces novel topics (i.e., creates a paradigm-modifying idea) can then be calculated as the distance of the topic distribution in the individual's contributions to the topic distribution of all other contributions on the platform. This is a widely used approach for measuring "novelty" in information retrieval (e.g., Vosoughi, Roy, & Aral, 2018). In order to calculate this measure, we employ topic modeling based on the Latent Dirichlet Allocation (LDA) with a Gibbs sampler (Blei et al., 2003). LDA refers to a generative probabilistic model that can be used to automatically detect topics that are underlying a collection of text documents. The process for detecting and analyzing the topics in text documents includes two essential steps. First, we used all contributions created by the crowd to uncover the topics that are present on the platform. We modeled 50 topics using the entire corpus of all 222'259 contributions. Second, we assigned each individual idea with probabilities for addressing each topic. We used the Information Uniqueness (IU; Vosoughi et al., 2018), which is based on the cosine similarity function, to measure the distance between two probability distributions (i.e., the topic vector of the contribution and the mean topic vector of all other contributions on the platform). The IU can be calculated as follows (Vosoughi et al., 2018):

$$IU(\Gamma_c, \Gamma_p) = 1 - cos(\Gamma_c, \Gamma_p)$$

where Γ_c represents the topic distribution of an idea while Γ_p represents the general topic distribution on the platform. *cos* refers to the cosine similarity function. A higher value for the IU indicates greater distance between the topic distributions in Γ_c and Γ_p.

Paradigm-Preserving Ideas. In order to measure the extent to which an individual creates a paradigm-preserving idea, we calculate an aggregated TF-IDF-index for the idea made by an individual *i* in week *t*. TF-IDF refers to a term weighting scheme in information retrieval (Salton & Buckley, 1988) that indicates how relevant a word is to a document (i.e., a contribution) in a corpus (i.e., knowledge trajectories for an innovation problem). It measures the frequency of a word in a contribution normalized by the document length (TF) and multiplies this value with the inverse document frequency of the word (IDF). A word that is frequently used in a contribution but rarely used in other contributions submitted to an innovation problem will receive a high TF-IDF-value. In this way, it is possible to measure the relevancy of the words in an idea and examine whether individuals are able to add relevant content to a topic in an idea. The TF-IDF-index for a term can be calculated as follows:

$$tfidf(w, c, C) = tf(w, c) * idf(w, C) \quad where$$

$$tf(w, c) = \frac{f_{w,c}}{\sum_{w' \in c} f_{w',c}} \quad and$$

$$idf(w, C) = \log \frac{N}{|\{c \in C : w \in c\}|}$$

Here, *tf(w,c)* is the frequency *f* of a word *w* in an idea *c* divided by the total number of words *w'* in idea *c*, while *idf(w,C)* is the logarithmically scaled inverse fraction of the ideas on the platform *C* that contain word *w*. We use the sum to calculate an aggregated TF-IDF index for each idea *c*:

$$tfidf\ index = \sum_{w \in c} tfidf(w, c, C)$$

Ideas with a high TF-IDF-index contain more relevant terms for a topic than ideas with a low TF-IDF-index. In this way, we follow related studies that have already used aggregated TF-IDF-indices to analyze textual contributions in crowdsourcing (e.g., Zhang et al., 2016).

Control Variables. We use several additional variables to control for alternative effects that could influence the likelihood of a solution being implemented by the organization and thus the performance of individuals in the crowd. First, a large number of related studies suggest that the activity of a crowd on a platform changes over time (see Bayus, 2013; Huang et al., 2014). Thus, increased *maturity of the platform* might influence how organizations select ideas for implementation. We control for this effect by including the number of weeks passed since the platform has been launched. Second, as shown by related literature (e.g., Schemmann et al., 2016), ideas that are popular in a crowd have a higher chance of being implemented by organizations than ideas that are less popular. We account for this effect by including the number of votes per solution (i.e., the *popularity* in the crowd) as a control variable. Third, organizations might be more likely to award individuals that convey positive ideas compared to individuals that are more critical. We control for this effect by accounting for the *sentiment* of the idea.

The descriptive statistics and the correlation matrix for the variables in our study are listed in Table 10. The measures raise no concerns regarding multicollinearity.

	Mean	SD	Min	Max	(1)	(2)	(3)	(4)	(5)	(6)
PER	0.05	0.22	0	1.00						
COL	90.53	108.89	0	503.93	-0.01**					
PMI	0.11	0.09	0	0.76	0.04***	-0.02***				
PPI	2.06	0.61	0	6.60	0.03***	0.01***	0.03***			
POP	2.51	2.84	0	41.00	0.13***	0.09***	0.02***	0.02***		
MAT	201.87	87.81	0	431.00	-0.02***	0.38***	-0.10***	0.05***	-0.04***	
SEN	0.15	0.36	-1	1.00	0.01***	0.03***	-0.02***	0.04***	0.00	0.01***

Note: PER = Performance, COL = Collaboration, PMI = Paradigm-Modifying Idea, PPI = Paradigm-Preserving Idea, POP = Popularity of Idea, MAT = Maturity of Platform, SEN = Sentiment of Idea

Table 10: Descriptive Statistics II

Source: Own Illustration

6.3.3 Estimation Approach

We use multilevel structural equation modeling (MSEM) as discussed by Preacher, Zyphur, and Zhang (2010) to test our hypotheses and conduct the mediation analysis. Multilevel modeling is the standard approach for mediation analysis with repeated measurements (cf. Bauer, Preacher, & Gil, 2006; Kenny, Korchmaros, & Bolger, 2003). In extant research, MSEM has shown to work well for mediation analysis with multiple mediators and panel data (e.g., Oishi, Kesebir, & Diener, 2011).

The MSEM consist of three regressions that examine (1) the effects of collaboration on the ability to create paradigm-modifying ideas (first mediator), (2) the effects of collaboration on the ability to create paradigm-preserving ideas (second mediator), and (3) the effects of the mediators on performance (complete model).

6.3.4 Models and Results

The results of the mediation analysis based on the MSEM are presented in Table 11 and Table 12. We report standardized coefficients for all variables.

	Estimate	Std. Err.	z-value	P(>|z|)
Performance ~				
Collaboration	-0.016	0.006	-2.654	0.008 ***
H1b: PMI	0.034	0.003	11.227	0.000 ***
H2b: PPI	0.022	0.003	7.651	0.000 ***
Control: Popularity	0.138	0.003	45.133	0.000 ***
Control: Maturity	0.003	0.004	0.644	0.520
Control: Sentiment	0.010	0.003	3.49	0.000 ***
PMI ~				
H1a: Collaboration	-0.017	0.003	-6.129	0.000 ***
PPI ~				
H2a: Collaboration	0.013	0.003	4.762	0.000 ***
R^2 (full model)	0.021			

Note: ***p<0.01; **p<0.05; *p<0.10

Table 11: Models and Results of the MSEM

Source: Own Illustration

	Estimate	Std. Err.	z-value	P(>|z|)
PMI	-0.001	0.000	-5.38	0.000 ***
PPI	0.000	0.000	4.043	0.000 ***
Total	-0.016	0.006	-2.703	0.007 ***

Note: ***p<0.01; **p<0.05; *p<0.10

Table 12: Mediation Effects

Source: Own Illustration

The results of the MSEM suggest that the hypothesized effects in the mediation model are statistically significant. It can be noted that collaboration overall negatively affects performance (β = -0.016; p < 0.001). The results also show that collaboration has a significant effect on both the ability to create paradigm-preserving ideas (β = 0.013; p < 0.001) and the ability to create paradigm-modifying ideas (β = -0.017; p < 0.05). Furthermore, both the ability to create paradigm-preserving ideas (β = 0.022; p < 0.001) and the ability to create paradigm-modifying ideas (β = 0.034; p < 0.001) are positively associated with performance at statistically significant levels. Table 12 shows that both indirect mediation effects for the ability to create paradigm-modifying ideas (β = -0.001; p < 0.001) and the ability to create paradigm-preserving ideas (β = 0.000; p < 0.001) are statistically significant. Likewise, the total effect is statistically significant ($\beta = -0.016$; $p < 0.05$)

Thus, the results of the mediation analysis support our hypotheses. As hypothesized in H1a and H2a, the effects of collaboration on the ability to create paradigm-modifying ideas and paradigm-preserving ideas are statistically significant but opposite in sign. Thus, our model provides evidence for an opposing mediation. Second, as hypothesized in H1b and H2b, our models show that both the ability to create paradigm-modifying ideas and paradigm-preserving ideas are beneficial for performance in solving innovation problems. Taken together, the results show that, with increased collaboration, individuals shift from creating value through paradigm-modifying ideas to creating value through to paradigm-preserving ideas.

6.4 Discussion of the Effects of Collaboration in Crowdsourcing

The results of our study yield several important findings. First, they show that both paradigm-modifying (H1b) and paradigm-preserving (H2b) ideas can be valuable for solving innovation problems for organizations. This extends earlier findings discussed by Nagasundaram and Bostrom (1994), who explain that "while innovators are constantly exploring and formulating new opportunities for the organization, adaptors ensure that those ideas are implemented and monitored efficiently" (p. 92). Our findings suggest that – especially in online crowds – striking a balance between innovators and adaptors and ensuring collaboration between them may greatly benefit the quality of knowledge trajectories. As shown in related studies, innovators are typically able to produce more original and radical contributions than experts in a particular field (e.g., Franke, Poetz, & Schreier, 2014; Jeppesen & Lakhani, 2010; Kristensson et al., 2004) but may struggle to assess the feasibility of their ideas (Poetz & Schreier, 2012). Adaptors are less experimental (Kirton & De Ciantis, 1986) but better able to validate the appropriateness of suggested solutions (Amabile, 1996; Perry-Smith & Shalley, 2003). Thus – just as "most organizations require a healthy blend of adaptors and innovators, each playing a special role in its growth and prosperity, and each style complementing the other" (Nagasundaram & Bostrom, 1994, p. 92) – we also see large online crowds benefiting from a blend between both innovators and adaptors.

Second, our results show how collaboration affects an individual's ability to provide paradigm-preserving and paradigm-modifying ideas over time. In existing literature, a large number of studies argues that mental representations tend to converge in crowds and thus inhibit creativity in the long run (e.g., Perry-Smith & Shalley, 2003), while others suggest that interaction and information exchange in crowds represent crucial prerequisites for creative outcomes (e.g., Faraj et al., 2011). By examining ideas with regard to their paradigm-relatedness, our findings may help to bridge the gap between these perspectives. Based on our results, we argue that collaboration affects the manner in which individuals become creative, rather than their capacity to become creative. We find that individuals in a crowd initially perform well at creating paradigm-modifying ideas and introducing novel topics. With increased collaboration, however, they become assimilated to the prevailing paradigms in the crowd and are less likely to continue introducing novel topics (H1b). We attribute this development to a bias effect. However, with increased collaboration, they become better at creating paradigm-preserving ideas and building upon existing topics in a crowd (H2b). This suggests that they are able to benefit from a learning effect based on increased domain-relevant knowledge in the crowd. In this way, our findings also extend Nagasundaram and Bostom's (1994) argument that "two individuals could be 'equally creative', but in different ways" (p. 91). Our results suggest that, over time, collaboration affects the way in which individuals solve innovation problems and become "creative".

Third, even though we argue that collaboration affects the manner in which individuals become creative, rather than their capacity to become creative, we still see risks associated with individuals gradually changing from innovators to adaptors in online crowds. Organizations often engage online crowds specifically for the purpose of searching for paradigm-modifying ideas and knowledge (Afuah & Tucci, 2012). Organizations should be aware that increased collaboration may lead to dominant paradigms and local knowledge bases over time. Our results suggest that, with increased collaboration, individuals gradually shift from being innovators to being adaptors. Thus, simply engaging online crowds in an attempt to span organizational boundaries may not suffice to find innovative, paradigm-breaking knowledge in the long run. While we see that collaboration overall has a positive effect on the ability to create valuable ideas in crowds, we argue that it is crucial for organizations to foster collaboration especially between innovators and adaptors.

6.4.1 Theoretical Implications

From a theoretical perspective, our findings have important implications for research on creativity, knowledge collaboration, and crowdsourcing.

For research on creativity, we contribute to extant literature by explaining how collaboration changes the manner in which individuals are able to operate within – or challenge – paradigms in online crowds. By examining ideas with regard to their paradigm-relatedness (Nagasundaram & Bostrom, 1994), we focus on the manner in which individuals become creative, rather than their capacity to become creative. Much related research on online crowds has focused on the latter (e.g., Bayus, 2013; Huang et al., 2014). Our results show that collaboration increases an individual's ability to operate within paradigms and refine existing topics but decreases his or her ability to challenge these paradigms and introduce novel topics. This suggests that individuals change from being "innovators" to "adaptors" in online crowds over time (Kirton, 1976; Nagasundaram & Bostrom, 1994).

Second, we contribute to research on knowledge collaboration by unraveling the effects of collaboration on solving innovation problems. Our results suggest that paradigm-modifying ideas are likely to be solicited from members of the crowd who have not yet extensively collaborated in the network and bring novel perspectives from outside the crowd to the platform. This extends findings presented in earlier research by Hargadon and Sutton (1997) and Lingo and Mahony (2010). Individuals who are not yet connected to redundant sources of information may serve as knowledge brokers and draw analogies or introduce new knowledge to a particular field or network. Hence, they offer unique informational benefits compared to those who are structurally central to and immersed in the network (Lingo & O'Mahony, 2010). However, when novel solutions need to be validated, further developed, and integrated, close collaboration and strong ties become more advantageous. Hansen (1999), for example, shows that weak interunit ties help project teams to search for useful knowledge in other subunits but impede the transfer of complex knowledge, which requires strong ties. In crowdsourcing, we found similar effects. Paradigm-modifying ideas emerge from contributions provided by individuals in the crowd who have not yet established a large number of connections in the network and are able to provide novel information to the crowd. Individuals with many network ties, on the other hand, may add to solutions by refining, integrating, and transferring knowledge.

Third, regarding research on crowdsourcing, a large body of extant literature has emphasized the importance of spanning organizational and technological boundaries in order to find new, paradigm-breaking knowledge (Katila & Ahuja, 2002; Katila et al., 2012; Rosenkopf & Nerkar, 2001). The interaction with external sources has been deemed especially promising for coming up with innovative ideas or solving existing problems in organizations (e.g., Chesbrough, 2003) and has spurred the emergence of novel, IT-facilitated approaches, such as crowdsourcing (Afuah & Tucci, 2012). Based on our findings, however, we argue that simply engaging in crowdsourcing in an attempt to span organizational boundaries may not be sufficient to find new and paradigm-breaking knowledge in the long run. The results presented in this study suggest that – even on crowdsourcing platforms – individuals gradually change from innovators to adaptors through increased collaboration and information exchange, which may lead to dominant paradigms and local knowledge bases on the platforms over time. In this way, our results unveil potential limitations of crowdsourcing and other network-based ideation approaches. In a related study, Huang et al. (2014) already examined crowdsourcing platforms under consumer learning and found that, over time, marginal idea contributors are filtered out, which is argued to be a signal of market efficiency. Consistent with these results, we also find individuals in online crowds to benefit from a "learning" effect. However, we show that this positive effect only applies to paradigm-preserving ideas. In contrast, collaboration and increased information exchange are prone to have a negative effect on the ability to make paradigm-modifying ideas.

6.4.2 Practical Implications

From a practical perspective, our results offer a number of important implications for the management of online crowds and the design of related online platforms. First, our results show that the ability to create paradigm-modifying and paradigm-preserving ideas changes over time. According to our findings, new individuals in crowds, who have not yet intensively collaborated with peers, are better at introducing novel topics to knowledge trajectories while experienced individuals, who already established many collaborative ties, are better at extending existing topics. Based on these findings, we urge organizations to encourage an exchange of information between these individuals in the crowd. While many organizations have begun to use crowdsourcing as an integral part of their innovation activity, it is challenging to maintain a steady stream of both innovative and feasible ideas with this approach (Blohm et al., 2013; Huang et al., 2014). Fostering collaboration might help to leverage the potential of both experienced and inexperienced individuals in crowds for creating both innovative and feasible solutions.

Second, since collaboration benefits the ability to refine topics at the expense of the ability to introduce novel topics in ideas, crowds may develop dominant patterns with regard to the topics they discuss over time. This is not only important with regard to how collaboration should be managed, but also with regard to how platforms and evaluation systems should be designed. Especially on well-established platforms with stable crowds, using rating scales and voting systems alone may not be the most appropriate mechanisms to identify paradigm-modifying ideas in crowds. Instead, integrating crowdsourcing platforms with novel business analytics that offer capabilities for social network analysis, text mining, and topic modeling could bridge this gap. Based on our findings, we see great potential in these technologies to support organizations in tracking the origin of contributions in crowdsourcing, analyzing their textual characteristics, and ultimately identifying the most innovative ones. We propose that organizations should delve deeper into such possibilities and make use of business analytics or decision support systems when engaging in crowdsourcing.

Third, our result show that creative ideas are often created by new members in a crowd who have only few effective network ties. While it is common for organizations to pay special attention to the most active individuals and follow suggestions that are popular in their established networks, we propose that organizations seeking to gain access to paradigm-breaking knowledge through crowdsourcing should rather focus on contributions that are generated by less immersed members of the crowd. Thus, endorsing new inputs and ensuring that contributions made by new individuals in the crowd receive attention seems critical for the sustainability and creative potential of crowds.

6.4.3 Limitations and Future Research

As with all research, the results and implications presented in this study should be regarded in light of its limitations. First, we analyze crowds from a network perspective in this study. The effects described and examined in our study are inherently based on interactions that unfold on crowdsourcing platforms. Hence, an important boundary condition of this study is that there are connections between members of a crowd and that crowds can be treated as networks of people. As for platforms or crowdsourcing settings with little interaction between contributors, our findings might be less applicable. For future research, it would thus be interesting to study how different forms of collaboration or interaction on crowdsourcing platforms affect individuals' performance. Our findings should be viewed as initial insights from a collaborative crowdsourcing setting.

Second, this study examines crowdsourcing in an organizational setting. Thus, we focus on a context that leverages crowdsourcing as an approach for sourcing new knowledge from outside existing, organizational boundaries (Afuah & Tucci, 2012). Furthermore, it must be noted that our results are based on data retrieved from a crowdsourcing platform hosted by an intermediary. While the characteristics of the dataset and the platform are similar to those reported in related studies (e.g., M. Li et al., 2016; Schemmann et al., 2016), there is still the possibility that our findings may not apply to all industries or applications of crowdsourcing to the same extent. Additional data from other crowdsourcing platforms would benefit the generalizability of our findings.

Third, future research may examine in more detail how crowdsourcing platforms and evaluation processes for crowdsourced contributions should be designed. Based on this study, we argue that rating scales and voting systems may not be the most adequate mechanisms for identifying innovative contributions, as their results are prone to being biased by already familiar perspectives and popular opinions amongst experienced members of the crowds. Thus, it would be interesting for future research to investigate how IT-supported processes with systems capable of tracking the origin of crowdsourced contributions and analyzing their textual characteristics may support organizations in their evaluation of large amounts of ideas and solutions on their platforms. From a theoretical perspective, we also urge future research to delve deeper into the possibilities of assessing and measuring creative contributions in crowdsourcing. For example, studies may use textual features, similarity measures, or redundancy indices as proxies for the innovativeness of contributions on crowdsourcing platforms.

6.5 Conclusion of Chapter

In this study, we examined how collaboration in online crowds affects the manner in which individuals become creative and solve innovation problems. Based on a mediation analysis with longitudinal data from a large crowdsourcing platform, we found that, over time, collaboration decreases an individual's ability to create paradigm-modifying ideas and introduce novel solutions but increases his or her ability to create paradigm-preserving ideas and refine existing knowledge. With these findings, we contribute to existing research in three ways. First, we explain how collaboration changes the manner in which individuals become creative and operate within – or challenge – paradigms in crowds. Our results suggest that, with increased collaboration, individuals gradually change from being *innovators,* who perform well at challenging paradigms, to *adaptors,* who perform well at refining paradigms. Second, we unravel the effects of collaboration on an individual's ability to solve innovation problems. Our findings show that collaboration has both positive and negative ramifications on an individual's ability to solve innovation problems. Third, our findings highlight limitations of crowdsourcing. They indicate that crowds are prone to develop dominant paradigms over time and simply broadcasting innovation problems to the same crowds may not be sufficient to find truly innovative ideas in the long run.

7 DECISION SUPPORT SYSTEMS DESIGN IN CROWDSOURCING

This chapter addresses the third and last research question of the dissertation. Based on previous insights, it investigates how decision support systems should be designed in crowdsourcing. It presents the results of a design science research study[6] that was concerned with defining design principles for such decision support systems based on text mining and machine learning technologies. In this chapter, section 7.1 first explains the motivation and objectives of the study in more detail. Second, section 7.2 describes the methodology of the study and elaborates on the design science research approach. Third, sections 7.3 and 7.4 reveal the results and discuss their implications for both theory and practice. Finally, section 7.5 concludes the chapter with a summary of the main findings.

7.1 The Importance of Design Knowledge

In crowdsourcing, organizations use open calls to engage large networks of people and collect their solutions, ideas, or feedback to solve a predefined task (Blohm et al., 2013). The approach offers the opportunity to take advantage of vast amounts of user-generated data and has found widespread adoption in different domains, including innovation management for the ideation of novel products (Blohm et al., 2013), software development for application testing (Leicht, Rhyn, & Hansbauer, 2016), or humanitarian aid for distributing relief supplies (Barbier et al., 2012). However, it represents a latent challenge to review and evaluate crowdsourced data (Barbier et al., 2012; Blohm et al., 2013). Piezunka and Dahlander (2015) studied how 922 organizations leveraged crowdsourced data and found that they often “fail to harness the full potential of crowdsourcing due to inadequate filtering mechanisms” (p. 876). Google, for example, required almost three years and 3’000 employees to analyze 150’000 ideas submitted to its *Project 10^{100}* (Blohm et al., 2013) while IBM had to employ 50 executives for several weeks to assess 46’000 ideas during its Innovation Jam (Bjelland & Wood, 2008).

In order to cope with large amounts of user-generated contributions in crowdsourcing, research and practice are increasingly using text mining and machine learning to support their evaluation. The ability of these algorithms to recognize patterns and extract useful information in a fast, scalable, and repeatable way is argued to be a key component for

[6] Parts of this chapter (early versions of sections 7.1 and 7.2) have been published as research in progress in Rhyn and Blohm (2017b). The research has been completed, reformatted, and updated for this dissertation. It includes new data, new design knowledge, an instantiation of the artifact, and a discussion of the findings. A modified version of the content is under review in: Rhyn, M., Leicht, N., Blohm, I., & Leimeister, J. M. (2020). Opening the Black Box: How to Design Intelligent Decision Support Systems in Crowdsourcing. Proceedings of the 15th International Conference on Wirtschaftsinformatik (WI), 1-15. Potsdam, Germany: AIS.

the (semi-)automated analysis of crowdsourced data (Chen et al., 2012). In crowdsourcing, a number of studies have already demonstrated the potential of these algorithms to support the evaluation of ideas (Walter & Back, 2013), the prioritization of software defects (Feng et al., 2015), or the identification of locations in incident reports (Barbier et al., 2012). However, these studies have mostly focused on domain-specific instantiations that demonstrate the technical capabilities (e.g., performance) of the algorithms. They have focused less on design knowledge that guides the deployment and adoption of text mining and machine learning in full-fledged information systems (Zhao & Zhu, 2014). Hence, while the technical development of the algorithms is already advanced, it is still unclear how information systems (IS) based on these algorithms (i.e., intelligent decision support systems) should be designed in crowdsourcing (Abbasi et al., 2016). Appropriate IS designs are crucial for the acceptance and adoption of intelligent decision support systems (W. Wang & Benbasat, 2005). They affect how decision makers work with the systems and how they improve their efficiency and effectiveness (Todd & Benbasat, 1999). Studies also show that, if such systems are not well designed, decision makers are likely to reject their recommendations and refrain from relying on them (W. Wang & Benbasat, 2005). In order to better understand how these systems should be designed, scholars have thus called for studies to "contribute guidelines for design artifacts" that support decision making in these contexts (Abbasi et al., 2016, p. xvii).

In this study, we aim to close this gap and answer the following research question: What design principles should guide the development of intelligent decision support systems (DSS) in crowdsourcing? Design principles are statements that capture abstract design knowledge and prescribe "what and how to build an artifact in order to achieve a predefined design goal" (Chandra et al., 2015, p. 4040). They make a design problem (e.g., designing intelligent DSS in crowdsourcing) more manageable for practitioners and provide researchers with a theoretical foundation to predict and evaluate the use patterns and impacts of the DSS (Markus, Majchrzak, & Gasser, 2002). To develop these design principles, we followed a design science research approach based on Peffers et al. (2007). Our research was conducted over a period of 3.5 years with a cross-industry research consortium comprising 8 organizations (Österle & Otto, 2010). It took part in three design-and-evaluate iterations that included a total of 41 semi-structured interviews, 13 focus group discussions with 53 participants, statistical analyses with training and test data from 676 crowdsourcing projects, and 2 field tests. In these iterations, we (1) defined design requirements with related design principles and design features for intelligent DSS, (2) developed software prototypes for a formative evaluation, and (3) instantiated them in a DSS in organizations for a summative evaluation.

The contribution of this study is threefold. First, we extend existing studies in the decision support field, which have mostly focused on the traditional efficiency-effectiveness framework (Shim et al., 2002; Todd & Benbasat, 1999; W. Wang & Benbasat, 2009), and introduce transparency and control as additional meta-requirements when designing intelligent systems. We find these two requirements to be fundamental for the willingness of decision makers to work with intelligent DSS and rely on their results. Second, for research on crowdsourcing, we define design principles to guide the development of intelligent DSS. We extend existing literature, which has already examined specific instantiations of text mining and machine learning technologies (e.g., Barbier et al., 2012; Feng et al., 2015; Walter & Back, 2013), and capture the necessary design knowledge for their deployment in intelligent DSS. Third, for developers of DSS, we describe specific design features that show how the design requirements and design principles can be addressed.

7.2 Design Science Research Approach

Design science research (DSR) represents a well-established approach in IS research that is concerned with the creation of artifacts seeking to extend the boundaries of human and organizational capabilities (Hevner et al., 2004). These artifacts may range from specific instantiations in the form of implemented software or algorithms to more theoretical contributions in the form of abstract design knowledge (Gregor & Jones, 2007). In this study, we focus on the latter and are concerned with defining design principles for intelligent DSS in crowdsourcing. Design principles are one of the most widely used vehicles to "convey design knowledge that contribute beyond instantiations applicable in a limited use context" (Chandra et al., 2015, p. 4039). Research typically conceptualizes design principles in conjunction with design requirements and design features (Meth, Mueller, & Maedche, 2015). *Design requirements* represent meta-requirements (Walls et al., 1992) which describe the "generic requirements that any artifact instantiated from this design should meet" (Meth et al., 2015, p. 807). *Design principles* can be defined as statements that prescribe how instantiated artifacts should be built in order to meet its requirements (Chandra et al., 2015; Meth et al., 2015). *Design features* represent specific ways to implement design principles in an actual artifact (Meth et al., 2015). Thus, design principles represent the link between overarching design requirements and concrete design features. They are important for IS research and practice on three accounts (Chandra et al., 2015). First, they *abstract* away from specific instantiations (e.g., design features) and capture design knowledge about instances of artifacts that belong to the same class (Sein, Henfridsson, Rossi, & Lindgren, 2011). Second, they communicate essential design knowledge and *prescribe* "what and how to build an artifact in order to achieve a predefined design goal [i.e., a design requirement]" (Chandra et al., 2015, p. 4040). Third, they *contribute* to more comprehensive design theories, e.g., IS designs for intelligent DSS (Gregor & Jones, 2007).

7.2.1 Research Process and Context

In order to systematically develop design requirements, design principles, and design features for intelligent DSS in crowdsourcing, we followed the well-established DSR process proposed by Peffers et al. (2007). This approach synthesizes the common phases of design science research discussed in existing literature (e.g., Hevner et al., 2004; Walls et al., 1992). Figure 8 below provides an overview of our process. As design science research represents an iterative and incremental approach (Hevner et al., 2004), we conducted three design-and-evaluate iterations. The data collection and analysis in these iterations is explained in more detail in section 7.2.2.

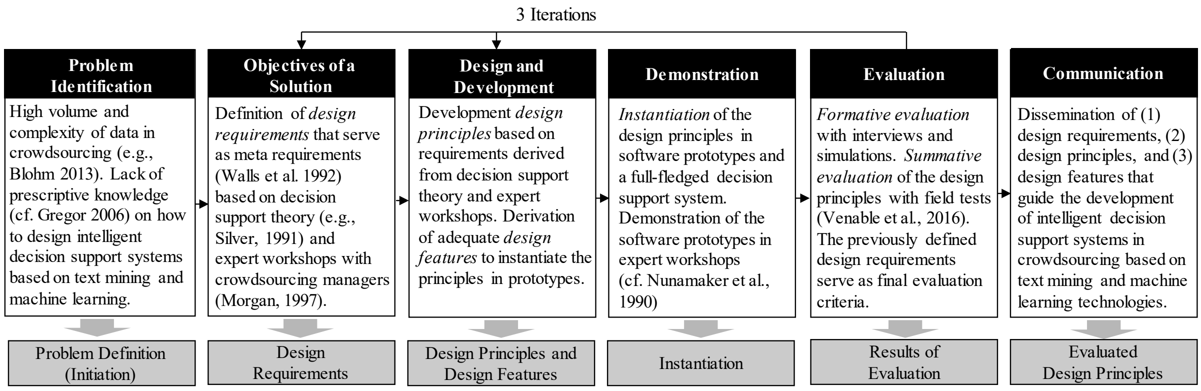


Figure 8: Design Science Research Approach

Source: Own Illustration based on Peffers et al. (2007)

Our research context was a cross-industry research consortium (Österle & Otto, 2010, p. 283) that consisted of 2 financial institutes, 2 insurance companies, 2 industrial corporations, 1 multinational retailer, and 1 public transportation provider. All companies use crowdsourcing for software testing and innovation (CST). This setting was chosen because CST exhibits two characteristics that make it especially well-suited for developing overarching design principles for intelligent DSS. First, CST comprises different types of textual contributions. Functional testing, for example, requires the crowd to contribute short and technical bug reports with ground truth, while usability testing aims to elicit generative feedback and ideas for software features with no ground truth. Second, CST comprises distinct decision making activities. In functional testing, decision makers are required to judge the severity of bug reports and prioritize them. In usability testing, they need to aggregate feedback and select the most requested features for change requests. Thus, CST can be regarded as a "microcosm" (Leicht et al., 2016, p.

3) for crowdsourcing insofar that it integrates a variety of textual contributions and decision making tasks in one unified setting. This should benefit the generalizability of the design principles beyond our research context.

7.2.2 Data Collection and Analysis

The three DSR iterations were conducted over a period of 3.5 years from December 2015 to June 2019. The final set of design requirements, design principles, and design features is based on rich data from a total of 41 semi-structured interviews, 13 focus group discussions with 53 participants, statistical analyses and simulations with data from 676 crowdsourcing projects, and 2 field tests with complete DSS in organizations.

In the first iteration, we aimed at *defining* an initial set of design requirements, design principles, and design features for intelligent DSS in crowdsourcing. For this purpose, we reviewed existing literature on decision making and decision support (e.g., Shim et al., 2002; Todd & Benbasat, 1999), and conducted 4 expert workshops and interviews with 40 participants from our research consortium. We employed moderated focus group discussions (Morgan, 1997) and asked the participants to describe the crowdsourcing process, explicate focal challenges, and outline potential improvements through DSS. We took notes and clustered the responses. To evaluate our results, we conducted 31 semi-structured with independent, external subject-matter experts (e.g., testing experts, QA managers, software developers). These interviews served as a first, formative evaluation[7] (Venable, Pries-Heje, & Baskerville, 2016).

In the second iteration, we focused on *instantiating* the design requirements, design principles, and design features in feasible software prototypes to ensure their technical feasibility. For this purpose, we referred to well-established text mining and machine learning algorithms (e.g., Breiman, 2001). We developed these prototypes in Python and demonstrated them in 9 expert workshops with 16 participants to gather insights on how to configure the algorithms, achieve sufficient performance (i.e., accuracy, sensitivity, specificity; Fawcett, 2006), and visualize the results. To ensure that the prototypes are feasible and achieve the targeted performance, we used simulations with data from 676 crowdsourcing projects conducted by organizations in our research consortium. The training and test data comprised more than 300’000 crowdsourced contributions. The simulations served as a second, formative evaluation[7] with instantiated prototypes (Venable et al., 2016).

[7] For more details on the evaluation, please refer to section 7.3.2.

In the third iteration, we focused on *implementing* a complete DSS in organizations of our consortium. We used insights from existing research on DSS designs (e.g., Silver, 1991) to develop a functional front end, a back end with our text mining and machine learning prototypes, and a database for a web-based DSS in crowdsourcing. We conducted 10 semi-structured interviews during which we demonstrated mockups of the system to experts (e.g., testing experts, QA managers) and gathered their feedback on the functionality of the system and the visualization of the results. As a final, summative evaluation[7], we conducted 2 field tests and implemented the DSS in organizations (Venable et al., 2016).

7.3 Design Knowledge for Intelligent DSS in Crowdsourcing

7.3.1 Design Requirements, Design Principles, and Design Features

Design principles communicate design knowledge on how to build an artifact to achieve a predefined design goal (Chandra et al., 2015). We refer to such design goals as design requirements (Meth et al., 2015; Walls et al., 1992). In decision support theory, existing research typically describes two primary objectives of decision makers: maximizing decision quality and minimizing effort (Shim et al., 2002; Todd & Benbasat, 1999). In practice, the decision makers in our workshops and interviews described similar goals and outlined two major issues in crowdsourcing that need to be addressed: (1) the quantity of contributions and (2) the complexity of their content. The former makes the evaluation time-consuming (e.g., "*it is not possible to manually process and evaluate all data*"; Innovation Manager, IT Services). The latter induces a high information load and makes the evaluation error-prone (e.g., "*it is definitely possible that a business analyst will reject [a change request] at a later stage because I made a mistake*"; Test Manager, Retail Bank). Thus, as a first important insight, we find that intelligent DSS should at least aim to increase (1) the efficiency and (2) the effectiveness of decision making in order to be useful in crowdsourcing. Importantly, however, the workshops and interviews revealed two additional requirements that have received much less attention in existing DSS research: maintaining transparency and control during decision making. For decision makers, intelligent DSS often represent a black box if they are not well explained (e.g., "*I would not blindly trust automated reports. I always want to know what is going on. I want to have enough control to be able to intervene*", Test Manager, Insurance). Transparency can be defined as a "mechanism to expose decision making" (Theodorou, Wortham, & Bryson, 2017, p. 233). Siau & Wang (2018) explain that for intelligent DSS, it is crucial to be able to understand "how they are programmed and what function will be performed in certain conditions. [A DSS] should be able to explain/justify its behaviors and decisions" (p. 51). Control, on the other hand, refers to "the degree of actual influence over the nature of the decision made" (Tyler, Rasinski, & Spodick, 1985, p. 72). It involves authority over the procedure through which a decision is made. Thus, as a second important insight, we argue that the design of an intelligent DSS in crowdsourcing should be considerate of two additional meta-requirements that revolve around (3) maintaining a sufficient degree of transparency and (4) maintaining a sufficient degree of control by its user. Figure 9 depicts the design requirements with theoretical and practical sources.

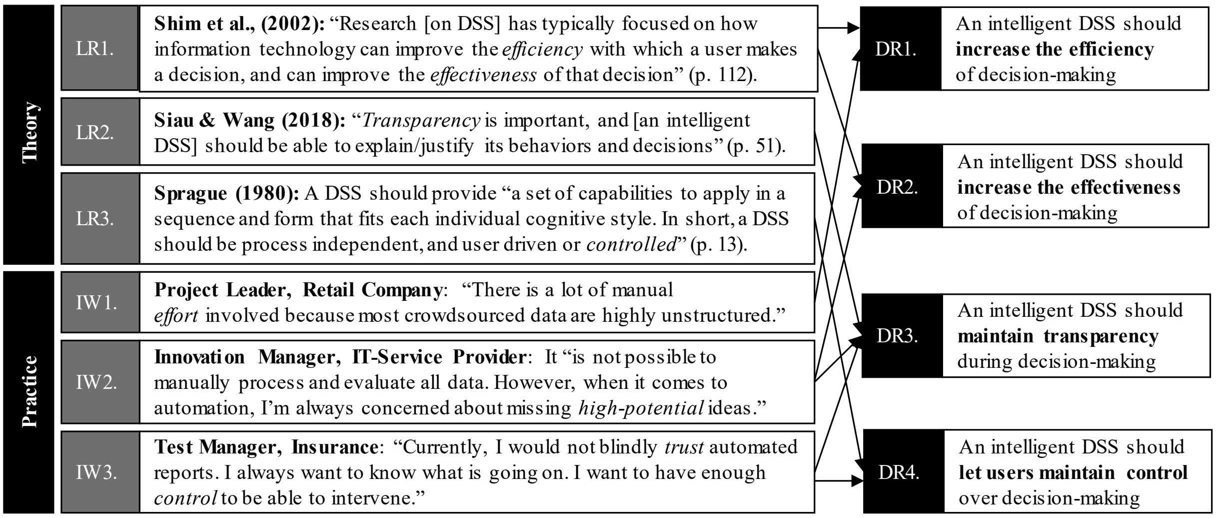


Figure 9: Design Requirements for Intelligent DSS in Crowdsourcing

Source: Own Illustration

Guided by these design requirements (DR), we developed design principles (DP) and design features (DF) for intelligent DSS in crowdsourcing. The workshops and interviews revealed 4 design principles that make it possible to reduce the manual effort and information load in crowdsourcing: an *omission* of irrelevant contributions (DP1), a *consolidation* of redundant contributions (DP2), a *prioritization* of important contributions (DP3), and an *indication* of the recommended decision (DP4). DSS that follow these design principles help to increase the efficiency (DR1) and effectiveness (DR2) of decision making in crowdsourcing. In crowdsourcing, it is possible to instantiate these principles with spam filters (DF1), triaging systems that group similar feedback (DF2), duplicate detection with sentiment analysis (DF3), and recommendations (i.e., probabilities for successful implementation; DF4). However, given that these features are part of intelligent DSS and build upon text mining and machine learning algorithms, it is crucial to maintain transparency (DR3) and control (DR4). The workshops and interviews revealed three additional design principles to address these design requirements:

a *translation* of machine outputs in human understandable actions (DP5), an *explanation* of operations leading to recommended actions (DP6), and a potential *adapation* of operations and rules (DP7). To instantiate these principles, the DSS should communicate the results in actionable and easy interpretable statements instead of abstract values or outputs (DF5), include popups and tooltips to explain the results (DF6), and allow the user of the DSS to configure the system and control the workflow (DF7). Figure 10 provides an integrated overview of our findings.

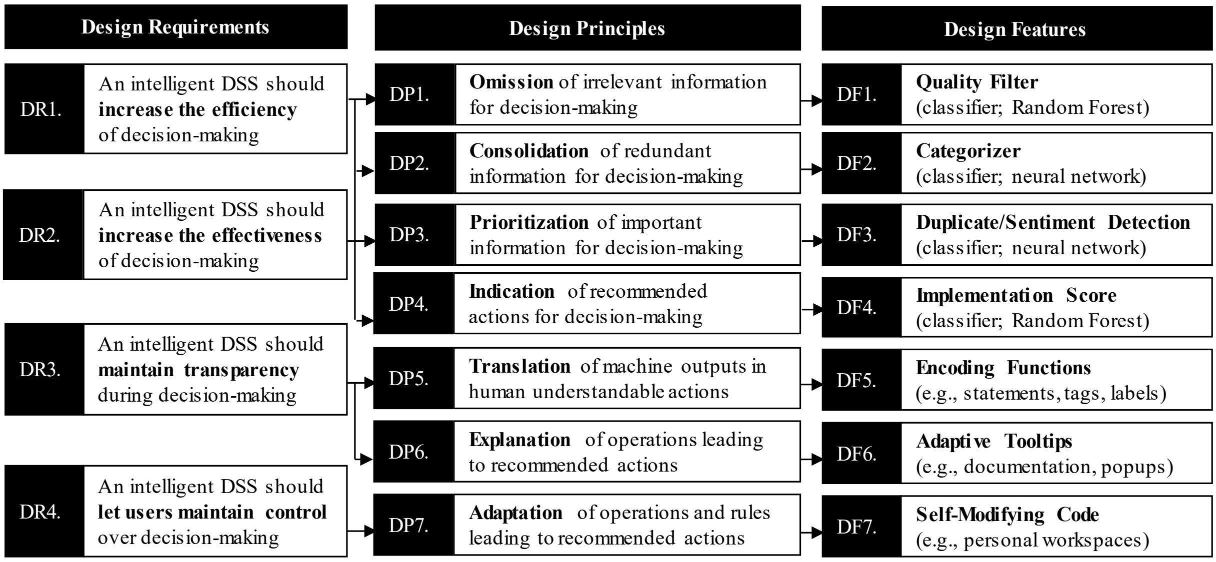


Figure 10: Overview of Design Requirements, Principles, and Features

Source: Own Illustration based on Meth et al. (2015)

Figure 11: Front End of the DSS with Design Features

Source: Own Illustration

For the final instantiation of our design principles and design features, we developed a complete DSS (see Figure 11). Exhibit A shows the system's analytics dashboard (Holsapple et al., 2014) that allows decision makers to visualize key performance indicators and monitor trends. It gives access to aggregated, high-level data and aims to provide a better understanding of the crowdsourced data. The interface is tile-based and offers the decision makers control over the appearance and the order of the algorithms' results. The underlying functions are explained in tooltips. Exhibit B shows a drill-down view into low-level data. These views are accessible through the dashboard and allow decision makers to search or scan for specific information in crowdsourced data once interesting patterns or trends have been identified. The DSS prioritizes important contributions, collapses duplicates, and offers recommendations in this view. Based on the decision makers' actions, verified labels are generated to improve the models in the back end.

7.3.2 Evaluation Results

For the evaluation (see Table 13), we followed the framework proposed by Venable et al. (2016). This evaluation framework creates a bridge between evaluation goals (*formative* or *summative*) and evaluation strategies (*artificial* or *naturalistic*) in DSR.

	Iteration 1	*Iteration 2*	*Iteration 3*
Goal	Relevancy (formative)	Feasibility (formative)	Usefulness (summative)
Strategy	Qualitative (artificial)	Quantitative (artificial)	Mixed (naturalistic)
Method	Interviews	Simulations	Field Tests
Data	31 semi-structured interviews	300'000 training and test data from 676 projects	Usage data; 10 semi-structured interviews

Table 13: Overview of Evaluation

Source: Own Illustration based on Venable et al. (2016)

First, we aimed to evaluate whether the requirements, principles, and features are relevant for DSS designs or whether they need to be adapted. The results of our 31 semi-structured interviews with independent subject-matter experts are consistent with findings from our workshops and from theory. They confirm that decision makers are looking for ways to increase their efficiency (e.g., "*faster reaction times";* Test Manager, Financial Services; DR1, DP1-4) and effectiveness (e.g., "*categorize feedback to examine the effectiveness of new app releases and updates*"; Technical Project Manager, Mobile App; DR2, DP1-4). They also emphasized that transparency (e.g., "*transparent and comprehensive results*", Interaction Designer, Marketing Services; DR3, DP5-6) and control (e.g., a "*human-centered approach"* with as much user control as possible; Project Manager, Utilities Provider; DR4, DP7) would be imperative for any intelligent DSS used in their jobs. Thus, we find support that our requirements, principles, and features capture relevant components for intelligent DSS designs.

Second, we aimed to ensure that the principles and features are technically feasible and can be instantiated. The results of our simulations with data from 676 crowdsourcing projects in the consortium are listed in Table 14. The performance measures show that classification algorithms are capable to achieve the decision makers' minimum requirements of 75% for the accuracy of DF1-DF4. We also note that the sensitivity and specificity are sufficiently high. The implementation of DF5-7 is possible with standard web-technologies for front end design (e.g., HTML, CSS, Javascript).

	Requirement (Acc.)	Accuracy	Sensitivity	Specificity
DF1	0.75	0.76	0.75	0.77
DF2	0.75	0.86	0.76	0.87
DF3	0.75	0.95	0.83	0.95
DF4	0.75	0.72	0.76	0.69

Table 14: Results of the Formative Evaluation in Second DSR Iteration

Source: Own Illustration

As a final, summative evaluation to ensure that our principles and features are useful and address the requirements with a DSS, we conducted 2 field tests with the DSS in organizations. We collected usage data and interviewed decision makers. For the evaluation of the efficiency (DR1) and the effectiveness (DR2), we followed Sproles' (2001). We were interested in the time reduction to process crowdsourced data (i.e., "how well the solution does what it actually does"; Sproles, 2001, p. 146) and asked the decision makers whether the DSS supports the evaluation (i.e., "the capability of a solution to meet the needs of a problem"; Sproles, 2001, p. 146). Pre-support processing was reported to take around 8 hours for 221 contributions. Post-support processing with the DSS took 4 hours (-50%) for the same data. In the second case, the reduction was reported to amount to -20%. The decision makers stated that the reduction is substantial and explained that DP2 and DP3 are particularly effective to support the evaluation. The decision makers confirmed that the tooltips (DP6) and the translated labels (DP5) offered the necessary explanation to understand the DSS's operations and interpret the reliability of the results (DR3). They also explained that the opportunity to change parameters (e.g., thresholds; DP7) and the authority over the final decisions gives them sufficient influence and control.

7.4 Discussion of the Design Knowledge

Taken together, the results of our study offer a number of important insights for the design of intelligent DSS in crowdsourcing. A key insight from the first iteration of our DSR study is that the traditional efficiency-effectiveness framework as discussed in existing decision support research (Shim et al., 2002; Todd & Benbasat, 1999; W. Wang

& Benbasat, 2009) does not sufficiently capture the decision makers' requirements for intelligent systems. Surprisingly, however, it is not primarily *trust* that is important for decision makers as discussed in much related literature (e.g., Siau & Wang, 2018; L. Wang, Jamieson, & Hollands, 2009). Trust is generally defined as "the willingness of a party to be vulnerable to the actions of another party […] irrespective of the ability to *monitor or control* that other party" (Mayer, Davis, & Schoorman, 1995, p. 712). However, our results show that decision makers do want to monitor and control intelligent systems. We find *transparency* and *control* to be crucial for the willingness of decision makers to work with intelligent DSS and rely on their results. They might also serve as crucial antecedents for trust in intelligent DSS (cf. Siau & Wang, 2018). Thus, we argue that maintaining transparency and control should be regarded as important meta-requirements for the design of intelligent DSS in crowdsourcing.

Second, insights from our study allow us to understand and explain the mechanisms through which efficiency, effectiveness, transparency, and control can be addressed. We find that, in crowdsourcing, increased efficiency and effectiveness can be explained by a reduction of both manual effort and information load in processing highly unstructured data (DP1-4) (Häubl & Trifts, 2000; Silver, 1991). However, even if the DSS is able to efficiently and effectively automate tasks, it is still important to give the users of the system a form of "decision control" (Tyler et al., 1985). Here, we find that transparency and control are related to (and achievable by) an understanding of the system's functions and an adequate representation of its results (DP5-7) (Siau & Wang, 2018). Our study provides exemplary features for DSS designs.

Third, insights from our study demonstrate that the instantiation of these design principles in intelligent DSS is both technically feasible and economically viable to support decision makers in crowdsourcing. From a technical perspective, our prototypes show that, even with traditional machine learning approaches (e.g., random forest algorithms; Breiman, 2001), it is possible achieve performance measures that are sufficient for practical use in crowdsourcing. From an economical perspective, our implemented DSS was able to reduce to manual workload of decision makers by up to 50%, which may lead to considerable savings in terms of cost and time. In line with existing research (Chen et al., 2012), we see organizations greatly benefitting from these technologies in crowdsourcing.

7.4.1 Implications for Theory

For research on decision support, we introduce transparency and control as two additional meta-requirements for intelligent systems. Existing decision support research has mostly focused on the traditional efficiency-effectiveness framework (e.g., Shim et al., 2002; Todd & Benbasat, 1999; W. Wang & Benbasat, 2009). Recently, however, with a shift towards large-scale, unstructured data collected from crowds (Holsapple et al., 2014), scholars in the DSS field emphasized that "this shift necessitates reconsidering guidelines for the design product and design process" (Abbasi et al., 2016, p. xvi) of DSS. Our study addresses this call and shows that increased efficiency and effectiveness are not sufficient for decision makers working with intelligent DSS. Instead, we find that transparency and control serve as key components for the adoption of intelligent DSS. Thus, we provide more "nuanced design requirements" (Abbasi et al., 2016, p. xvii) as requested in related literature.

For research on crowdsourcing, we capture the theoretical design knowledge for instantiating DSS based on text mining and machine learning. Prior studies have already examined isolated instantiations of these technologies and exemplified their capabilities in domain-specific applications (e.g., Barbier et al., 2012; Feng et al., 2015; Walter & Back, 2013). We extend these studies and provide a set of design principles that guide the deployment and adoption of text mining and machine learning in intelligent DSS. The design principles represent the link between overarching design requirements and concrete design features. They explain how efficiency, effectiveness, transparency, and control in decision making can be increased.

7.4.2 Practical Contributions

Our study also offers a number of practical contributions for developers of DSS and managers of crowdsourcing initiatives or platforms. For developers, we describe specific design features that show how the design principles can be instantiated in order to meet the four design requirements of intelligent DSS in crowdsourcing. These features include a quality/spam filter, a categorization system, a sentiment analysis in combination with duplicate recognition, and recommendations based on predictive models for the decision makers. Based on our results, we also urge developers to translate raw machine outputs into more human understandable actions and to pay attention to documentation and tooltips to explain the functionality of the underlying algorithms.

Second, for managers of crowdsourcing initiatives or platforms, our findings show that adequate intelligent DSS may drastically increase the efficiency and effectiveness of the evaluation of user-generated contributions by crowds. Abbasi et al. (2016), for example, emphasize that "IS research needs to not only contribute to the design but also examine the feasibility and effectiveness of such IT artifacts for different stakeholders" (p. viii). Decision makers that used our DSS were able to reduce the required time to process the contributions by 50%. Based on our findings, we recommend managers of crowdsourcing initiatives to make use of such systems and implement text mining and machine learning algorithms on crowdsourcing platforms. We found that these technologies are both technically feasible and economically viable to support decision makers in evaluating large amounts of crowdsourced contributions.

7.4.3 Limitations and Outlook

As with all research, there are limitations to the findings presented in this study. First, our study focused specifically on intelligent DSS that deal with textual data in CST. While we believe that our design requirements also translate to other intelligent DSS beyond crowdsourcing, we cannot claim that our design principles and features capture universal design knowledge for all other forms of intelligent DSS. We urge future research to investigate design principles and features for other contexts and examine similarities or differences between them (e.g., study the mutability of our principles; Gregor & Jones, 2007).

Second, we followed Chandra et al. (2015) and captured design knowledge in the form of design requirements, design principles, and design features. We acknowledge that the conceptualization of design requirements, design principles, and design features represents only the first step toward a more comprehensive understanding of IS designs for intelligent DSS. We see great potential in future research to extend our study and delve deeper into principles of implementations for DSS (e.g., methods or processes for organizational adoption; Gregor & Jones, 2007) and use patterns (e.g., testable propositions; Gregor & Jones, 2007).

Third, we focused on system design rather than system use of intelligent DSS. Thus, an interesting avenue for future research is to study in more detail how different designs of intelligent DSS affect performance in organizations and how decision makers work with these systems. We strongly believe that these technologies will represent fundamental components for future DSS designs and thus justify further research.

7.5 Conclusion of Chapter

In crowdsourcing, it represents a challenge to process textual contributions. Research already examined the technical capabilities of text mining and machine learning to support decision makers. Yet, it remained unclear how to design intelligent DSS based on these algorithms. We addressed this gap with a DSR approach and developed design requirements, design principles, and design features to guide the development of intelligent DSS in crowdsourcing. Our study shows that intelligent DSS based on these principles are feasible to support decision makers in evaluating crowdsourced data.

8 SYNTHESIS OF THE FINDINGS

The overarching objective of this dissertation was to study decision making in crowdsourcing and examine how text mining and machine learning can provide decision support. Based on the findings presented in the dissertation, it is possible to draw a comprehensive picture of emerging patterns of decision making in crowdsourcing (RQ1), understand how text mining and machine learning can be used to support decision making and process crowdsourced data (RQ2), and outline adequate design principles for decision support systems in crowdsourcing based on these technologies (RQ3). The following sections integrate and discuss these findings.

8.1 The Nature of Decision Making in Crowdsourcing

With regard to the characteristics of decision making processes in crowdsourcing (RQ1), the dissertation reveals two important differences compared to traditional organizational settings. First, in crowdsourcing, decision makers have the opportunity to freely source and prospect new data. This may lead to better-informed decisions. Compared to traditional decision making in organizations, they are no longer limited to existing databases nor do they have to build upon extant knowledge in their organizations. Instead, they can access vast pools of user-generated content and behavioral data from crowds at any stage of decision making. In this way, crowdsourcing may drastically facilitate the collection of data and enables decisions to be grounded on actual data from large networks of people. Such developments have frequently been described as a shift towards more open, data-driven decision making that draws upon actual information about people's behavior, opinions, or choices (Abbasi et al., 2016; Lycett, 2013; Sharma et al., 2014). Second, however, this new opportunity to freely source and prospect data does have a downside. The characteristics of the data and the way in which they need to be processed during decision making become more complex. In traditional, controlled environments, it is possible to ensure data quality through clear specification of input and output requirements. In crowdsourcing, with distributed and diverse information sources, "the traditional process of information requirements determination is practically unachievable" (Lukyanenko et al., 2014, p. 687). Thus, working with crowdsourced data is often challenging and requires sophistical data-processing activities before they can be used.

With these differences in the way that data can be sourced and the way that data need to be processed, the nature of decision making changes in crowdsourcing. In existing literature, these changes were not well understood (Sharma et al., 2014). Thus, the first research question of the dissertation aimed to explore the patterns of decision making that emerge in crowdsourcing. The findings to answer this research question were presented in chapter 3. They provided an in-depth understanding of these patterns and their characteristics. The findings reveal two important insights.

A first important insight is that decision making processes in crowdsourcing can be described sequences of five distinct phases. These phases represent germane episodes of (1) sourcing, (2) validating, (3) consolidating, and (4) evaluating data to (5) choose adequate courses of actions for solving an organizational problem. On the one hand, these findings make it possible to better understand how decision makers leverage crowdsourced data. In this way, they may help to identify critical activities or phases during crowdsourcing projects in terms of time and effort and offer a more granular, data-centric perspective on decision making than the traditional phase theorem (cf. Simon, 1960). On the other hand, the phases allow for a flexible and modular analysis of their alignment during decision making. Given the opportunity to work with actual data from large and diverse networks of people, decision makers in crowdsourcing are often found to alternate between phases. This leads to the second important insight.

A second important insight is that decision making processes in crowdsourcing do not always represent a predetermined sequence of phases but rather follow different patterns. The findings presented in chapter 3 reveal that there are four dominant patterns of decision making in crowdsourcing. These patterns are the result of an interaction between the structure of the decision problem and the decision maker's mode of acquiring information. Decision making follows sequential patterns when decision makers use crowdsourced data to efficiently "inform" decisions or gradually "solve" decision problems. In these cases, decision makers employ a goal-oriented search and follow predefined routines. Decision making follows more recursive and iterative patterns when decision makers use crowdsourcing to "explore" new options or better "understand" them. In these cases, data act as drivers for the decision making process and decision makers employ a dynamic and open scanning for information. These patterns contrast early models of decision making, which either assumed that decision making is mostly linear (e.g., Simon, 1960) or does generally not exhibit any recurring structure (e.g., Witte, 1972). They are more in line with more recent with findings presented by Mintzberg et al. (1976) and Boonstra (2003), who argue that it is possible to "identify general patterns and a basic logic in decision-making, although the processes are not always predetermined, linear and explicit" (p. 197).

8.2 Mechanisms to Support Decision Makers in Crowdsourcing

The findings described in chapter 3 not only provide a better understanding of decision making in crowdsourcing, they also offer valuable insights with regard to how decision making can be supported (RQ2). In related literature, the degree to which decision making can – and should – be supported in these situations is highly debated. Markus (2015), for example, note a "large-scale shift to decision automation" (p. 58) in recent years, while Sharma et al. (2014) emphasize that "human insights are still involved in 'accepting' the insights generated via machine learning as being valid and useful" and "in 'deciding' to deploy them to run operations in an unguided manner" (p. 436). Arnott and Pervan (2012), on the other hand, argue that decision support mechanisms should ultimately "support decision-making, not replace the person in the decision-making process" (p. 925). The results of the dissertation show that there is high potential to *combine* automated data-processing procedures with human judgement from decision makers to fully leverage crowdsourcing and capture value from crowdsourced data. Text mining and machine learning provide the technological means to support decision making through the (semi-)automated processing of unstructured data. Chapters 4 to 6 studied in more detail how crowds generate valuable contributions and how text mining and machine learning can be used to facilitate their evaluation. Thus, they show how the patterns of decision making identified in chapter 3 can potentially be supported.

In cases where decision making patterns resemble a goal-oriented search for information (e.g., when analyzing and verifying defect reports in crowdsourced software testing), there is great potential to support decision makers by automating repetitive data-processing activities for increased efficiency in decision making. Chapter 4 built upon these findings and showed how text mining and machine learning can be used to automate the time-consuming validation of crowdsourced data through a filtering mechanism. The results suggest that it is possible to explain and predict the quality of contributions in crowdsourcing based purely on textual features of crowdsourced contributions. These results provide evidence for the relationship between a contribution's quality and its contextual and representational characteristics. Thus, they show that is possible to train machine learning algorithms to analyze contextual and representational characteristics of crowdsourced contributions and make predictions about their quality. In addition, the models and variables presented in chapter 4 describe how to operationalize these contextual and representational characteristics. Four variables (i.e., length, specificity, readability, and spelling) have been shown to work well with a random forest algorithm to automatically assess and classify textual contributions. This may provide the foundation for partially automating the evaluation of textual data in crowdsourcing.

In cases where decision making patterns reflect a more dynamic exploration of data (e.g., when deciding on new products during innovation campaigns), there is great potential to support decision makers in identifying valuable contributions and extracting novel insights from crowdsourced data. Chapters 5 and 6 built upon these findings and studied the emergence and recombination of knowledge in crowds to potentially identify valuable contributions more easily. The findings show that valuable contributions often result from collaborative efforts in a crowd and emerge from a combination of different ideas and comments. Valuable contributions initially emerge from novel ideas provided by individuals in the crowd who are not highly connected. These individuals have unique informational benefits and can serve as knowledge brokers that introduce new knowledge to a crowd. Individuals who are highly connected may add to these contributions by integrating and transferring complex and local knowledge through discussions. In this way, the findings offer important insights with regard to the origin of valuable contributions and the way they are developed by individuals in the crowd.

Taken together, chapters 4 to 6 show the potential of text mining and machine learning to provide decision support by analyzing the characteristics of both crowdsourced data and their contributors. They contribute the empirical foundations to develop sophisticated decision support systems that integrate these algorithms and models to facilitate the evaluation of large amounts of contributions on crowdsourcing platforms.

8.3 Building Information Systems for Decision Support

Based on the decision making patterns described in chapter 3 and the empirical foundations on how to support decision making with text mining and machine learning developed in chapters 4 to 6, an important question that arises for IS research is how to design related decision support systems (RQ3). Design principles are one of the most widely used vehicles to convey such design knowledge (Chandra et al., 2015). They are typically conceptualized in conjunction with design requirements and design features (Meth et al., 2015). Thus, the DSR study in chapter 7 was concerned with investigating design requirements, design principles, and design features for decision support systems in crowdsourcing that build upon text mining and machine learning technologies.

One of the most important findings described in chapter 7 is that decision makers are often hesitant to trust text mining and machine learning technologies in crowdsourcing. For decision makers, transparency and control are two fundamental requirements when it comes to using intelligent decision support systems based on these technologies. That is, decision makers want to be able to understand how the underlying algorithms work and intervene if they do not agree with their results or recommendations. Thus, when designing intelligent decision support systems based on text mining and machine learning technologies in crowdsourcing, focusing purely on increasing the efficiency and effectiveness in decision making as suggested in most existing research (e.g., Shim et al., 2002; Todd & Benbasat, 1999; W. Wang & Benbasat, 2009) is not sufficient to ensure an acceptance and adoption of such systems – even if text mining and machine learning are technically able to address these two requirements.

Furthermore, while chapters 4 to 6 showed how text mining and machine learning can be used to automatically measure and examine the characteristics of crowdsourced data and their contributors, they did not prescribe how to design decision support mechanisms based on text mining and machine learning. Chapter 7 extends these findings and outlines overarching design principles on how to address the efficiency, effectiveness, transparency, and control in decision support systems. The guiding design principles in crowdsourcing to achieve increased efficiency and effectiveness are an omission of irrelevant contributions, an aggregation of redundant contributions, a prioritization of important information, and an indication of recommended actions. These principles resonate strongly with the mechanisms and models described in chapters 4 to 6. They relate to a reduction of both manual effort and information load (Häubl & Trifts, 2000; Silver, 1991). As described previously, however, increased efficiency and effectiveness are not the only meta-requirements for the design of intelligent decision support systems in crowdsourcing. It is also important to consider the representation of the system's results and recommendations (transparency) and potential interventions by the decision maker (control). Thus, chapter 7 shows that these objectives can be achieved in crowdsourcing by translating machine outputs into human understandable actions, by clearly explaining the operations that led to the recommended actions, and by adapting the operations and rules to fit user preferences.

Finally, and most importantly, decision support systems that use text mining and machine learning and are based on these design principles have shown to be viable to support decision makers in crowdsourcing. The results presented in chapter 7 show that, from a technical perspective, traditional machine learning algorithms (e.g., random forest algorithms; Breiman, 2001) are able to achieve performance measures that are sufficient for applications in crowdsourcing. From an economical perspective, decision makers that used intelligent decision support systems in crowdsourcing were able to reduce the manual workload by up to 50%. In line with existing research (Chen et al., 2012), this shows the great potential of these technologies to support decision making and realize considerable savings in terms of cost and time in crowdsourcing.

9 THEORETICAL CONTRIBUTIONS

The dissertation offers theoretical contributions for research on decision making, research on crowdsourcing, and research on decision support systems. The following sections discuss the theoretical contributions for each of these research streams.

9.1 Decision Making Patterns for Research on Decision Making

To date, much existing research in the field of decision making built upon the traditional decision making model proposed by Simon (1960) and assumed that "structured (often quantitative) data is intentionally collected to inform specific models and provide predefined input to the decision-making process" (Constantiou & Kallinikos, 2015, p. 45). Comprehensive literature reviews show that this model remains the "most cited conceptualization of the phase theorem of decision making" (Arnott & Pervan, 2014, p. 271). In recent years, however, there has been a move "toward dealing with massive collections of relatively unstructured data" (Holsapple et al., 2014, p. 131). Scholars thus began to question how the nature of decision making changes in this context and have issued multiple calls for research to study decision making in environments that deal with large amounts of unstructured data (e.g., Abbasi et al., 2016; Constantiou & Kallinikos, 2015; Sharma et al., 2014). The dissertation addresses this gap and uses crowdsourcing as an exemplary case to examine how decision making changes in this context. The dissertation yields two important contributions to this research field.

First, the findings presented in this dissertation show that decision making does not always adhere to a systematic and linear structure. This contrasts the traditional phase theorem of decision making (cf. Simon, 1960) and reveals its limitations in more data-driven environments, such as crowdsourcing. A systematic and linear structure in decision making may still occur when decision makers employ a goal-oriented search and have a good understanding of how to source and analyze data (e.g., in crowdsourced software testing). However, there is evidence that, especially when facing unstructured decision problems (e.g., innovation problems), decision making becomes much more flexible and volatile than assumed by the traditional phase theorem. When given the opportunity to freely source and prospect data, decision makers tend to employ a more dynamic and open scanning for information, with decision making processes becoming increasingly recursive and iterative in nature. Thus, in such decision settings, decision making reflects a data-driven "sense-making process" (Lycett, 2013).

Second, in response to the rather deterministic phase theorem (cf. Simon, 1960), some scholars have rejected the idea of distinct phases in decision making and argued that decision making rather consists of a plurality of sub-decisions oftentimes occurring simultaneously while gathering and processing information (e.g., Witte, 1972). However, the results of this dissertation show that even more dynamic and data-driven "sense-making processes" during decision making are not random. Instead, the dissertation shows that an interaction between the structure of the underlying decision problem (Shim et al., 2002; Simon, 1960) and the decision maker's mode of acquiring information (Aguilar, 1967; Huber, 1991; Vandenbosch & Huff, 1997) may evoke different patterns in decision making. Thus, the findings of the dissertation extend findings by Boonstra (2003) and Mintzberg et al. (1976), who have already argued for the existence of different patterns in decision making. The dissertation shows how the phase theorem of decision making can be extended to data-driven contexts, by describing different path configurations of decision making phases rather than a uniform process.

9.2 Determinants of Contribution Quality for Research on Crowds

In research on online crowds and crowdsourcing, there has been an increasing interest to leverage crowdsourced data to uncover new insights on how large networks of people create and share knowledge (e.g., Bayus, 2013; Huang et al., 2014; Schemmann et al., 2016) and how such insights can be used to support decision making processes in organizations (e.g., Hoornaert et al., 2017; M. Li et al., 2016). To this end, Hoornaert et al. (2017) argue that it is crucial to understand the content of crowdsourced data, the contributors of the data, and the crowd in which the contributors are embedded. This understanding not only makes it possible to explain behavior in crowds. It also makes it possible to better identify valuable contributions and support decision makers that engage online crowds. The findings presented in this dissertation contribute to research on online crowds and crowdsourcing on three accounts.

First, the dissertation describes the relationship between the quality of crowdsourced data and their contextual and representational characteristics. It extends previous frameworks for data quality presented in related literature (e.g., Otterbacher, 2009; R. Y. Wang & Strong, 1996) and provides empirical evidence that the quality of textual contributions in crowdsourcing can be explained by four features: the length, the specificity, the readability, and the spelling. These features show that textual contributions in crowdsourcing need to be well-elaborated and precise in order to be useful for decision makers in organizations (contextual characteristics) and that they need to be presented in a clear and easily interpretable manner (representational characteristics). With these insights, the dissertation also contributes the foundation for partially automating the evaluation of textual data in crowdsourcing and supporting decision makers. The length of a contribution has shown to be one of the most effective indicators for explaining and predicting its quality, while the readability and the specificity reveal moderate predictive power. The spelling represents the least important feature for the classification.

Second, the dissertation provides an extended understanding of how large networks of people create and share knowledge with their textual contributions. Much related research so far has suggested that engaging in crowdsourcing is a very efficient and effective approach to span organizational boundaries and elicit new knowledge to solve problems (e.g., Afuah & Tucci, 2012; Jeppesen & Lakhani, 2010). However, the findings of the dissertation outline the limitations of crowdsourcing and highlight challenges for effective decision making. Even on crowdsourcing platforms, increased participation and collaboration have been found to lead to dominant paradigms and local knowledge bases. Thus, simply engaging in crowdsourcing is not sufficient to elicit or find new knowledge. Novel information is likely to be solicited from individual who are not yet highly connected in a crowd. They often serve as knowledge brokers and are able to introduce new and diverse perspectives to the crowd. Immersed users, on the hand, are better capable to integrate and transfer knowledge through discussions. Thus, in line with Hoornaert et al. (2017), the findings underline that it is important to examine not only the content of the contributions, but also the characteristics of their contributors and the crowd in which they are embedded.

Third, the results of the dissertation extend our understanding of collaboration and knowledge recombination in crowdsourcing (e.g., Faraj et al., 2011; Majchrzak & Malhotra, 2016). Existing research (e.g., Katila & Ahuja, 2002; Winter, 1984) has already shown that unique and valuable solutions are often created through a combination of existing knowledge (depth) with new inputs (scope). The findings presented in this dissertation provide a better understanding of these processes in online crowds and show how collaboration changes the manner in which individuals are able to operate within – or challenge – paradigms. The results show that collaboration increases an individual's ability to operate within paradigms and refine existing topics but decreases his or her ability to challenge these paradigms and introduce novel topics. This suggests that individuals change from being "innovators" to "adaptors" in online crowds over time (Kirton, 1976; Nagasundaram & Bostrom, 1994). Paradigm-modifying ideas often emerge from individuals who have not yet intensely collaborated and are able to provide novel information to the crowd. Individuals with many network ties, on the other hand, may refine, integrate, and transfer knowledge.

9.3 Design Knowledge for Research on Decision Support Systems

Over the past years, research on decision support systems has witnessed a fundamental shift toward dealing with massive collections of unstructured data (Holsapple et al., 2014). Decision support systems are becoming increasingly "intelligent" and build upon novel text mining and machine learning technologies to deal with these large-scale, unstructured data (Arnott & Pervan, 2014). As a result, however, the designs and requirements for these intelligent systems are expected to change (Abbasi et al., 2016). Scholars have emphasized that this makes it necessary to reconsider the "guidelines for the design product and design process associated with such artifacts" (Abbasi et al., 2016, p. xvi). The dissertation makes important contributions to this research field on two accounts.

First, most existing research on decision support systems built upon the traditional efficiency-effectiveness framework (Shim et al., 2002; Todd & Benbasat, 1999; W. Wang & Benbasat, 2009) and considered improvements to the efficiency and effectiveness of decision making as the two overarching objectives (or meta-requirements) for the design of decisions support systems. The dissertation extends the traditional efficiency-effectiveness framework and introduces transparency and control as two additional meta-requirements for the design of intelligent decision support systems. It shows that increased efficiency and effectiveness are not sufficient to ensure an acceptance and adoption of intelligent systems – even if the systems are technically able to address these two requirements. The results and recommendations generated by intelligent systems are often perceived as inscrutable due to inadequate representations and lacking interpretability (Siau & Wang, 2018). The dissertation shows that fundamental requirements of decision makers are to be able to understand how the underlying algorithms of intelligent systems work and intervene if they do not agree with their results or recommendations. Transparency and control should be regarded as additional meta-requirements alongside efficiency and effectiveness of intelligent decision support systems.

Second, while the technical foundations of intelligent decision support systems are already advanced and well-understood, research still lacked the prescriptive design knowledge on how to build these systems (Abbasi et al., 2016). Appropriate IS designs are not only crucial for the acceptance and adoption of text mining and machine learning (W. Wang & Benbasat, 2005), they also affect how people work with this technology and improve their efficiency and effectiveness (Todd & Benbasat, 1999). The dissertation addresses this gap and provides prescriptive knowledge in the form of design principles. These design principles serve as a theoretical foundation to evaluate the use patterns and impacts of decision support systems in crowdsourcing (Markus et al., 2002).

10 PRACTICAL IMPLICATIONS

The findings of the dissertation offer a number of practical contributions and provide valuable insights for improving decision making in crowdsourcing, developing text mining and machine learning models for this purpose, and designing related decision support systems. The following sections discuss these practical contributions.

10.1 Improving Decision Making in Crowdsourcing

For organizations that engage in crowdsourcing, the findings of this dissertation show how to improve the efficiency and effectiveness of decision making when facing large amounts of user-generated data. Processing and evaluating user-generated contributions has often been described as a latent challenge and represented one of the most time-consuming and cost-intensive activities in practical applications of crowdsourcing (cf. Barbier et al., 2012; Blohm et al., 2013; Nagar et al., 2016; Walter & Back, 2013). In cases of large projects, such processes could take several weeks (Bjelland & Wood, 2008) or even years (Blohm et al., 2013). With the results of this dissertation, it is possible to provide a number of recommendations on how to address these challenges.

First, chapter 3 offered a better understanding of how decision makers typically process crowdsourced data in practice. With these insights, it is possible to better identify the phases during decision making that represent critical bottlenecks in terms of time and effort and to derive ways to support decision makers during these phases. For structured decision problems, where decision making often resembles a goal-oriented search for specific information (e.g., in software testing), the potential for improving decision making has been found to be especially high by automatically preprocessing data to support the time-consuming and repetitive validation of crowdsourced data. In unstructured cases, where decision making patterns reflect a more dynamic exploration of data (e.g., in product innovation), decision making can be improved by offering in-depth options for experimenting with data and visualizing results. Importantly, these findings show that there is no "one-size-fits-all" solution to improve decision making in crowdsourcing and that organizations should adapt the means for decision support to different, context-specific decision making patterns.

Second, the findings presented in this dissertation show that text mining and machine learning are able to serve as the technological means to support these decision making patterns in crowdsourcing. Organizations that face large amounts of user-generated data in crowdsourcing and have mostly relied on a manual evaluation (e.g., expert reviews) are urged to integrate text mining and machine in their decision making processes. As shown in chapters 4 to 6, text mining and machine learning are capable to make the evaluation of large amounts of crowdsourced data more efficient and effective. They can automate time-consuming and repetitive data-processing steps to reduce manual effort and recommend promising contributions for implementation. This offers the potential to realize both time and cost savings.

10.2 Developing Text Mining and Machine Learning Models

For practitioners that aim to develop these text mining and machine learning applications in crowdsourcing, the dissertation also offers a set of predictors and models that can be used for setting up feasible instantiations in crowdsourcing.

First, the findings presented in chapter 4 provide developers with four textual features that have been found to work well for an assessment of contribution quality in crowdsourcing. They show that it is possible to predict the quality of textual contributions in crowdsourcing based on the contributions' length, their readability, their specificity, and their spelling. These features can be used by developers to instantiate filter mechanisms on crowdsourcing platforms or in decision support systems. Such filter mechanisms may help to reduce manual efforts in crowdsourcing and support the time-consuming and repetitive validation of crowdsourced data.

Second, the findings presented in chapters 5 and 6 offer insights on the origin of innovative contributions in crowdsourcing and provide developers with potential features to automatically identify these types of contributions. They show that it is possible to determine whether a contribution is innovative or not by examining the network size of the contributor in the crowd, the novelty of the contribution's content, the novelty of the comments' content, and the diversity of the comments' content. All features can be calculated by relatively simple measures, such as the contributor's effective network size through social network analysis or the TF-IDF-index from information retrieval.

10.3 Designing Intelligent Decision Support Systems

Finally, based on the findings presented in this dissertation, it is possible to provide practitioners with recommendations on how to develop intelligent decision support systems in crowdsourcing. The dissertation offers the necessary design knowledge to instantiate text mining and machine learning technologies in practice.

First, the findings presented in chapter 7 outline generalized design principles and specific design features for this purpose. The latter include a quality/spam filter, a categorization system, a sentiment analysis in combination with duplicate recognition, and recommendations based on predictive models for decision makers. They show that, even with rather traditional text mining and machine learning algorithms, it is possible to reduce the amount of time to process crowdsourced data for decision making by up to 50%. Thus, based on the findings presented in this dissertation, it is highly recommended for developers of intelligent decision support systems to incorporate text mining and machine learning technologies in software products and make use of their capability to automatically process and evaluate large amounts of user-generated data.

Second, the dissertation also shows that it is important for developers to implement functions that translate raw machine output into human understandable actions. Based on the findings presented in chapter 7, developers are urged to make use of documentation and tooltips to explain how text mining and machine learning algorithms work and how the results of related decision support systems are generated. Such features have been found to be crucial for perceived transparency and control when using intelligent decision support systems in practice.

11 LIMITATIONS AND AVENUES FOR FUTURE RESEARCH

As with all research, there are limitations to the findings presented in the dissertation. The following sections discuss these limitations and outline how future research may address them or extend the dissertation's findings.

11.1 Crowdsourcing and Its Boundary Conditions

First, it must be emphasized that this dissertation examines decision making and decision support technologies in the context of crowdsourcing. In crowdsourcing, data are systematically collected to solve a predefined task. This makes it possible to study in detail how different patterns of decision making emerge for different types of decision problems, how data are processed in different phases of these patterns, and how these phases can be supported by information systems for increased efficiency and effectiveness. While these conditions are in many ways ideal to study decision making and decision support mechanisms, the findings are still limited to the context of crowdsourcing and should be regarded in light of the characteristics and boundary conditions of this particular domain. In other contexts beyond crowdsourcing, different patterns may emerge and different decision support technologies might be necessary depending on the type of data and the manner in which data are collected. The dissertation focuses on textual data that are collected from large networks of people for a predefined task.

Thus, further research is needed to draw a more comprehensive picture of decision making patterns and decision support mechanisms beyond crowdsourcing. Based on the findings presented in this dissertation, there are three major avenues that would be particularly interesting for future research: (1) What are the fundamental characteristics of "data-driven" decision making that remain stable across different industries and different domains? (2) How and to what extent do these decision making patterns change over time? (3) What are the roles of both humans and intelligent systems in decision making? Addressing these questions may help provide a more general understanding of decision making and decision support mechanisms in data-driven environments.

11.2 Focus on the Design of Decision Support Technologies

Second, besides examining decision making patterns in crowdsourcing, the dissertation is in large parts concerned with mechanisms through which text mining and machine learning technologies may improve these patterns. It focuses particularly on the design of related decision support systems and less on the actual usage of such systems and how decision making processes may change as a result of the usage of such systems. However, Sharma et al. (2014) emphasize that valuable insights from decision making processes "do not emerge automatically out of mechanically applying analytical tools to data. Rather, insights emerge out of an active process of engagement between analysts and business managers using the data and analytic tools to uncover new knowledge" (p. 435). Thus, it is not only crucial to understand decision making processes and design intelligent decision support systems around them, but also to study the way in which these decision support mechanisms are used and how they may lead to improved performance in organizations (Sharma et al., 2014).

Thus, future research may continue to delve deeper into the use patterns (cf. Markus et al., 2002) of such decision support systems and the resulting changes of decision making processes in organizations. The dissertation offers the necessary design knowledge to develop such systems. Yet, there is still a need to examine the impact of these systems in more detail. Another important avenue for future research is to study the role of transparency (Siau & Wang, 2018) and control (Silver, 1990) in intelligent decision support systems. As described in chapter 7, efficiency and effectiveness only represent two of four meta-requirements for the design of intelligent decision support system. They are mostly based on a traditional understanding of decision support systems (cf. Shim et al., 2002). The results of the dissertation, however, indicate that transparency and control are two additional, crucial meta-requirements for the design of more recent, intelligent systems. Having scratched only the surface of these two meta-requirements, future studies may investigate in more detail how transparency and control in intelligent decision support systems can be improved, how they affect the adoption and usage of such systems, and how they affect performance in decision making.

11.3 Decision Making at an Individual Level

Finally, the dissertation focuses on decision making at an individual level. It does not cover decision making processes at a group level (e.g., Bettenhausen & Murnighan, 1985; De Dreu & West, 2001) or at an organization level (e.g., Cyert & March, 1963; Maitlis & Ozcelik, 2004). However, as also emphasized by related research (e.g., Sharma et al., 2014), individual decision making is only the first step to understanding how organizations may benefit from crowdsourced data and intelligent decision support systems. Individual decision making always takes place in social and organizational structures, processes, routines, hierarchies, and governance mechanisms.

Hence, there is great potential for future research to extend the findings presented in this dissertation and study decision making at a group level or an organizational level. There have been multiple calls in recent years from both scholars in the fields management research (e.g., George, Haas, & Pentland, 2014) and IS research (e.g., Agarwal & Dhar, 2014; Chen et al., 2012; Sharma et al., 2014) to investigate how organizations create business value through data-driven decisions and why some organizations succeed while others fail (Sharma et al., 2014). In particular, an important question that requires more thorough examination is how "existing organizational structures, routines and decision-making processes influence the ability of managers and analysts to generate insights from data" (Sharma et al., 2014, p. 435). Thus, based on the findings presented in this dissertation, future research may address these questions and study processes and structures for data-driven decision making at an organizational level.

12 CONCLUSION

Crowdsourcing represents a powerful approach for organizations to systematically collect data from large networks of people. While research already made great strides in recent years to develop the technological foundations for processing large amounts of user-generated data, it remained mostly unclear how these new data sources and technologies affect decision making in organizations. The objectives of this dissertation were to identify patterns of decision making that emerge in crowdsourcing, understand how decision making in crowdsourcing can be improved with text mining and machine learning, and capture the necessary design knowledge to develop decision support systems in crowdsourcing based on these technologies. To accomplish these objectives, the dissertation was organized in three research streams. The first research stream aimed to describe common patterns of decision making in crowdsourcing. It was based on an exploratory interview study that aimed to offer a better understanding of how the structure of decision problems, the characteristics of the available data, and the way in which such data can be generated in crowdsourcing affect decision making. The second research stream aimed to examine how decision making in crowdsourcing can be improved with text mining and machine learning. Statistical analyses were used to better understand how crowds create valuable contributions for organizations and how decision makers can identify and process these contributions more efficiently and effectively. Finally, the third research stream followed a design science research approach. It was concerned with integrating the previous findings and capturing design knowledge to develop decision support systems in crowdsourcing. Taken together, the dissertation provides a number of important theoretical contributions. First, it illustrates the limitations of traditional decision making models in data-driven environments, such as crowdsourcing, and describes four common patterns of decision making that emerge when decision makers have access to large-scale, user-generated data. Second, the dissertation provides the empirical foundations to increase the efficiency and effectiveness of decision making in crowdsourcing by offering a better understanding of how crowds generate valuable contributions and how decision makers may process these contributions with text mining and machine learning technologies. Third, the dissertation provides prescriptive design knowledge in the form of design requirements and design principles for the development of decision support systems in crowdsourcing. For practitioners, the dissertation offers recommendations on how to improve the efficiency and effectiveness of decision making in crowdsourcing, how to leverage text mining and machine learning technologies in this context, and how to instantiate the technologies.

References

Abbasi, A., Sarker, S., & Chiang, R. H. L. (2016). Big Data Research in Information Systems: Toward an Inclusive Research Agenda. *Journal of the Association of Information Systems*, *17*(2), 1–32.

Afuah, A., & Tucci, C. L. (2012). Crowdsourcing as a Solution to Distance Search. *Academy of Management Review*, *37*(3), 355–375.

Agarwal, R., & Dhar, V. (2014). Big Data, Data Science, and Analytics: The Opportunity and Challenge for IS Research. *Information Systems Research*, *25*(3), 443–448.

Agichtein, E., Castillo, C., Donato, D., Gionis, A., & Mishne, G. (2008). Finding High-Quality Content in Social Media. *Proceedings of the 2008 International Conference on Web Search and Data Mining*, 183–193. Palo Alto: ACM.

Aguilar, F. (1967). *Scanning the Business Environment*. New York: MacMillan.

Ahuja, G., & Lampert, C. M. (2001). Entrepreneurship in the Large Corporation: A Longitudinal Study of How Established Firms Create Breakthrough Inventions. *Strategic Management Journal*, *22*(6–7), 521–543.

Alba, J. W., & Hutchinson, J. W. (1987). Dimensions of Consumer Expertise. *Journal of Consumer Research*, *13*(4), 411–454.

Allahbakhsh, M., Benatallah, B., Ignjatovic, A., Motahari-Nezhad, H. R., Bertino, E., & Dustdar, S. (2013). Quality Control in Crowdsourcing Systems: Issues and Directions. *IEEE Internet Computing*, *17*(2), 76–81.

Amabile, T. M. (1982). Social Psychology of Creativity: A Consensual Assessment Technique. *Journal of Personality and Social Psychology*, *43*(5), 997–1013.

Amabile, T. M. (1996). *Creativity in Context*. Boulder, CO: Westview Press.

Anderson, P. A. (1983). Decision Making by Objection and the Cuban Missile Crisis. *Administrative Science Quarterly*, *28*(2), 201–222.

Ariav, G., & Ginzberg, M. J. (1985). DSS Design: A Systemic View of Decision Support. *Communications of the ACM*, *28*(10), 1045–1052.

Arnott, D., & Pervan, G. (2012). Design Science in Decision Support Systems. *Journal of the Association for Information Systems*, *13*(11), 923–949.

Arnott, D., & Pervan, G. (2014). A Critical Analysis of Decision Support Systems Research Revisited: The Rise of Design Science. *Journal of Information Technology*, *29*(4), 269–293.

Arun, R., Suresh, V., Veni Madhavan, C. E., & Narasimha Murty, M. (2010). On Finding the Natural Number of Topics with Latent Dirichlet Allocation: Some Observations. *Advances in Knowledge Discovery and Data Mining, PAKDD 2010*, 391–402. Hyderabad: Springer.

Barbier, G., Zafarani, R., Gao, H., Fung, G., & Liu, H. (2012). Maximizing Benefits from Crowdsourced Data. *Computational and Mathematical Organization Theory*, *18*(3), 257–279.

Barki, H., & Pinsonneault, A. (2001). Small Group Brainstorming and Idea Quality: Is Electronic Brainstorming the Most Effective Approach? *Small Group Research*, *32*(2), 158–205.

Bauer, D. J., Preacher, K. J., & Gil, K. M. (2006). Conceptualizing and Testing Random Indirect Effects and Moderated Mediation in Multilevel Models: New Procedures and Recommendations. *Psychological Methods*, *11*(2), 142–163.

Bayus, B. L. (2013). Crowdsourcing New Product Ideas over Time: An Analysis of the Dell IdeaStorm Community. *Management Science*, *59*(1), 226–244.

Bertaux, D. (1981). From the Life-History Approach to the Transformation of Sociological Practice. In D. Bertaux (Ed.), *Biography and Society: The Life History Approach in the Social Sciences* (pp. 29–45). Bevery Hills, CA: SAGE Publications.

Bettenhausen, K., & Murnighan, J. K. (1985). The Emergence of Norms in Competitive Decision-Making Groups. *Administrative Science Quarterly*, *30*(3), 350–372.

Bjelland, O. M., & Wood, R. C. (2008). An Inside View of IBM's "Innovation Jam." *MIT Sloan Management Review*, *50*(1), 32–40.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, *3*(2003), 993–1022.

Blohm, I., Bretschneider, U., Leimeister, J. M., & Krcmar, H. (2010). Does Collaboration Among Participants Lead to Better Ideas in IT-based Idea Competitions? An Empirical Investigation. *Proceedings of the 43rd Annual Hawaii International Conference on System Sciences (HICSS)*, 1–10. Honolulu: IEEE.

Blohm, I., Leimeister, J. M., & Krcmar, H. (2013). Crowdsourcing: How to Benefit from (Too) Many Great Ideas. *MIS Quarterly Executive*, *12*(4), 199–211.

Blohm, I., Riedl, C., Füller, J., & Leimeister, J. M. (2016). Rate or Trade? Identifying Winning Ideas in Open Idea Sourcing. *Information Systems Research*, *27*(1), 27–48.

Blohm, I., Zogaj, S., Bretschneider, U., & Leimeister, J. M. (2018). How to Manage Crowdsourcing Platforms Effectively? *California Management Review*, *60*(2), 122–149.

Bonabeau, E. (2009). Decisions 2.0: The Power of Collective Intelligence. *MIT Sloan Management Review*, *50*(2), 45–52.

Boonstra, A. (2003). Structure and Analysis of IS Decision-Making Processes. *European Journal of Information Systems*, *12*(3), 195–209.

Borgatti, S. P. (1997). Structural Holes: Unpacking Burt's Redundancy Measures. *Connections*, *20*(1), 35–38.

Boudreau, K. J., & Lakhani, K. R. (2013). Using the Crowd as an Innovation Partner. *Harvard Business Review*, *91*(4), 60–69.

Bourreau, M., Gensollen, M., & Moreau, F. (2012). The Impact of a Radical Innovation on Business Models: Incremental Adjustments or Big Bang? *Industry & Innovation*, *19*(5), 415–435.

Braha, D., & Reich, Y. (2003). Topological Structures for Modeling Engineering Design Processes. *Research in Engineering Design*, *14*(4), 185–199.

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32.

Brynjolfsson, E., Geva, T., & Reichman, S. (2015). Crowd-Squared: Amplifying the Predictive Power of Large-Scale Crowd-Based Data. *MIS Quartely*, *40*(4), 941–961.

Brynjolfsson, E., & Hitt, L. M. (1998). Beyond the Productivity Paradox. *Communications of the ACM*, *41*(8), 49–55.

Burt, R. S. (1995). *Structural Holes: The Social Structure of Competition*. Cambridge, Massachusetts: Harvard University Press.

Carlile, P. R. (2002). View of Knowledge and Boundaries: Boundary Objects in New Product Development. *Organization Science*, *13*(4), 442–455.

Carlile, P. R., & Rebentisch, E. S. (2003). Into the Black Box: The Knowledge Transformation Cycle. *Management Science*, *49*(9), 1180–1195.

Chandra, L., Seidel, S., & Gregor, S. (2015). Prescriptive Knowledge in IS Research: Conceptualizing Design Principles in Terms of Materiality, Action, and Boundary Conditions. *Proceedings of the 48th Hawaii International Conference on System Sciences (HICSS)*, 4039–4048. Hawaii: IEEE.

Chatman, J. A., Polzer, J. T., Barsade, S. G., & Neale, M. A. (2007). Being Different Yet Feeling Similar: The Influence of Demographic Composition and Organizational Culture on Work Processes and Outcomes. *Administrative Science Quarterly*, *43*(4), 749–780.

Chen, H., Chaing, R. H. L., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*, *36*(4), 1165–1188.

Chesbrough, H. W. (2003). The Era of Open Innovation. *MIT Sloan Management Review*, *44*(3), 35–41.

Chiu, C. M., Liang, T. P., & Turban, E. (2014). What Can Crowdsourcing Do for Decision Support? *Decision Support Systems*, *65*(2014), 40–49.

Christensen, C. M. (1997). *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Cambridge, Massachusetts: Harvard Business School Press.

Cohen, J. (1968). Weighted Kappa: Nominal Scale Agreement with Provision for Scaled Disagreement or Partial Credit. *Psychological Bulletin*, *70*(4), 213–220.

Coleman, M., & Liau, T. L. (1975). A Computer Readability Formula Designed for Machine Scoring. *Journal of Applied Psychology*, *60*(2), 283–284.

Constantiou, I. D., & Kallinikos, J. (2015). New Games, New Rules: Big Data and the Changing Context of Strategy. *Journal of Information Technology*, *30*(1), 44–57.

Cook, G. J. (1993). An Empirical Investigation of Information Search Strategies with Implications for Decision Support System Design. *Decision Sciences*, *24*(3), 683–698.

Corbin, J., & Strauss, A. (1990). Grounded Theory Research: Procedures, Canons, and Evaluative Criteria. *Qualitative Sociology*, *13*(1), 2–21.

Corney, J. R., Torres-Sánchez, C., Jagadeesan, A. P., & Regli, W. C. (2010). Outsourcing Labour to the Cloud. *International Journal of Innovation and Sustainable Development*, *4*(4), 294–313.

Cronin, M. A., & Weingart, L. R. (2013). Representational Gaps, Information and Conflict in Processing, Diverse Teams Functionally. *Academy of Management Review*, *32*(3), 761–773.

Cyert, R., & March, J. G. (1963). *A Behavioral Theory of the Firm*. Englewood Cliffs, NJ: Prentice Hall.

Dahl, D. W., & Moreau, P. (2002). The Influence and Value of Analogical Thinking During New Product Ideation. *Journal of Marketing Research*, *39*(1), 47–60.

De Dreu, C. K. W., & West, M. A. (2001). Minority Dissent and Team Innovation: The Importance of Participation in Decision Making. *Journal of Applied Psychology*, *86*(6), 1191–1201.

Dean, D. L., Hender, J. M., Rodgers, T. L., & Santanen, E. L. (2006). Identifying Quality, Novel, and Creative Ideas: Constructs and Scales for Idea Evaluation. *Journal of the Association for Information Systems*, *7*(10), 646–698.

Debortoli, S., Müller, O., Junglas, I. A., & vom Brocke, J. (2016). Text Mining for Information Systems Researchers: An Annotated Tutorial. *Communications of the*

AIS, 1–30.
Dell. (2017). IdeaStorm. Retrieved March 15, 2017, from http://www.ideastorm.com/
Doan, A., Ramakrishnan, R., & Halevy, A. Y. (2011). Crowdsourcing Systems on the World-Wide Web. *Communications of the ACM*, *54*(4), 86–96.
Dosi, G. (1982). Technological Paradigms and Technological Trajectories. A Suggested Interpretation of the Determinants and Directions of Technical Change. *Research Policy*, *11*(3), 147–162.
Dunbar, K. (1998). Problem Solving. In W. Bechtel & G. Graham (Eds.), *A Companion to Cognitive Science* (pp. 289–298). London: Blackwell.
Durward, D., Blohm, I., & Leimeister, J. M. (2016). Crowd Work. *Business & Information Systems Engineering*, *58*(4), 281–286.
Eppler, M. J., & Mengis, J. (2004). The Concept of Information Overload: A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines. *The Information Society*, *20*(5), 325–344.
Faraj, S., Jarvenpaa, S. L., & Majchrzak, A. (2011). Knowledge Collaboration in Online Communities. *Organization Science*, *22*(5), 1224–1239.
Fawcett, T. (2006). An Introduction to ROC Analysis. *Pattern Recognition Letters*, *27*(8), 861–874.
Feldman, R., & Sanger, J. (2007). *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Cambridge: Cambridge University Press.
Felin, T., & Zenger, T. R. (2014). Closed or Open Innovation? Problem Solving and the Governance Choice. *Research Policy*, *43*(5), 914–925.
Feng, Y., Chen, Z., Jones, J. A., Fang, C., & Xu, B. (2015). Test Report Prioritization to Assist Crowdsourced Testing. *Proceedings of the 10th Joint Meeting on Foundations of Software Engineering (ESEC/FSE)*, 225–236. Lombardy: ACM.
Fleming, L. (2001). Recombinant Uncertainty in Technological Search. *Management Science*, *47*(1), 117–132.
Fleming, L., & Sorenson, O. (2004). Science as a Map in Technological Search. *Strategic Management Journal*, *25*(8–9), 909–928.
Förster, J., Friedman, R. S., & Liberman, N. (2004). Temporal Construal Effects on Abstract and Concrete Thinking: Consequences for Insight and Creative Cognition. *Journal of Personality and Social Psychology*, *87*(2), 177–189.
Franke, N., Poetz, M. K., & Schreier, M. (2014). Integrating Problem Solvers from Analogous Markets in New Product Ideation. *Management Science*, *60*(4), 1063–1081.
Frisk, J. E., Lindgren, R., & Mathiassen, L. (2014). Design Matters for Decision Makers: Discovering IT Investment Alternatives. *European Journal of Information Systems*, *23*(4), 442–461.
Gallupe, R. B., Dennis, A. R., Cooper, W. H., Valacich, J. S., Lana, M., & Nunamaker, J. F. (1992). Electronic Brainstorming and Group Size. *The Academy of Management Journal*, *35*(2), 350–369.
Garfield, M. J., Taylor, N. J., Dennis, A. R., & Satzinger, J. W. (2001). Modifying Paradigms - Individual Differences, Creativity Techniques, and Exposure to Ideas in Group Idea Generation. *Information Systems Research*, *12*(3), 322–333.
Geiger, D., & Schader, M. (2014). Personalized Task Recommendation in Crowdsourcing Information Systems - Current State of the Art. *Decision Support Systems*, *65*(2014), 3–16.

George, G., Haas, M. R., & Pentland, A. (2014). Big Data and Management. *Academy of Management Journal*, *57*(2), 321–326.

Gerrity, T. P. (1971). Design of Man-Machine Decision Systems: An Application to Portfolio Management. *Sloan Management Review*, *12*(2), 59–75.

Ghose, A., & Ipeirotis, P. G. (2011). Estimating the Helpfulness and Economic Impact of Product Reviews: Mining Text and Reviewer Characteristics. *IEEE Transactions on Knowledge and Data Engineering*, *23*(10), 1498–1512.

Ghose, A., Ipeirotis, P. G., & Li, B. (2012). Designing Ranking Systems for Hotels on Travel Search Engines by Mining User-Generated and Crowdsourced Content. *Marketing Science*, *31*(3), 493–520.

Gieryn, T. F., & Hirsh, R. F. (1983). Marginality and Innovation in Science. *Social Studies of Science*, *13*(1), 87–106.

Gioia, D. a., Corley, K. G., & Hamilton, A. L. (2013). Seeking Qualitative Rigor in Inductive Research: Notes on the Gioia Methodology. *Organizational Research Methods*, *16*(1), 15–31.

Glaser, R., & Strauss, A. (1967). *Discovery of Grounded Theory: Strategies for Qualitative Research*. New York: Routledge.

Goldenberg, J., Lehmann, D. R., & Mazursky, D. (2001). The Idea Itself and the Circumstances of Its Emergence as Predictors of New Product Success. *Management Science*, *47*(1), 69–84.

Goldenberg, J., Mazursky, D., & Solomon, S. (1999). Toward Identifying the Inventive Templates of New Products: A Channeled Ideation Approach. *Journal of Marketing Research*, *36*(2), 200–210.

Gorry, G. A., & Scott Morton, M. (1971). A Framework for Management Information Systems. *Sloan Management Review*, *13*(1), 55–70.

Granovetter, M. (1973). The Strength of Weak Ties. *American Journal of Sociology*, Vol. 78, pp. 1360–1380.

Granovetter, M. (1983). The Strength of Weak Ties: A Network Theory Revisited. *Socialogical Theory*, *1*(1), 201–233.

Gregan-Paxton, J., & John, D. R. (1997). Consumer Learning by Analogy: A Model of Internal Knowledge Transfer. *Journal of Consumer Research*, *24*(3), 266–284.

Gregor, S., & Jones, D. (2007). The Anatomy of a Design Theory. *Journal of the Association for Information Systems*, *8*(5), 312–335.

Griffiths, T. L., & Steyvers, M. (2004). Finding Scientific Topics. *Proceedings of the National Academy of Sciences*, *101*(1), 5228–5235.

Guest, G., Bunce, A., & Johnson, L. (2006). How Many Interviews Are Enough? An Experiment with Data Saturation and Variability. *Field Methods*, *18*(1), 59–82.

Hansen, M. T. (1999). The Search-Transfer Problem: The Role of Weak Ties in Sharing Knowledge across Organization Subunits. *Administrative Science Quarterly*, *44*(1), 82–111.

Hargadon, A. B., & Bechky, B. A. (2006). When Collections of Creatives Become Creative Collectives: A Field Study of Problem Solving at Work. *Organization Science*, *17*(4), 484–500.

Hargadon, A. B., & Sutton, R. I. (1997). Technology Brokering and Innovation in a Product Development Firm. *Administrative Science Quarterly*, *42*(4), 716–749.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). New York: Springer.

Hatchuel, A., & Weil, B. (2009). C-K Design Theory: An Advanced Formulation. *Research in Engineering Design*, *19*(4), 181–192.

Häubl, G., & Trifts, V. (2000). Consumer Decision Making in Online Shopping Environments: The Effects of Interactive Decision Aids. *Marketing Science*, *19*(1), 4–21.

Hautz, J., Hutter, K., Füller, J., Matzler, K., & Rieger, M. (2010). How to Establish an Online Innovation Community? The Role of Users and Their Innovative Content. *Proceedings of the 43rd Hawaii International Conference on Systems Sciences (HICSS)*, 1–11. Kauai: IEEE.

Helfat, C. E. (1994). Evolutionary Trajectories Firm R & D in Petroleum. *Management Science*, *40*(12), 1720–1747.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly*, *28*(1), 75–105.

Hoch, S. J., & Schkade, D. A. (1996). A Psychological Approach to Decision Support Systems. *Management Science*, *42*(1), 51–64.

Holmström, J., Ketokivi, M., & Hameri, A.-P. (2009). Bridging Practice and Theory: A Design Science Approach. *Decision Science*, *40*(1), 65–87.

Holsapple, C., Lee-Post, A., & Pakath, R. (2014). A Unified Foundation for Business Analytics. *Decision Support Systems*, *64*(2014), 130–141.

Hoornaert, S., Ballings, M., Malthouse, E. C., & Van den Poel, D. (2017). Identifying New Product Ideas: Waiting for the Wisdom of the Crowd or Screening Ideas in Real Time. *Journal of Product Innovation Management*, *34*(5), 580–597.

Hotho, A., Nürnberger, A., & Paaß, G. (2005). A Brief Survey of Text Mining. *LDV Forum - GLDV Journal for Computational Linguistics and Language Technology*, *20*(1), 19–62.

Howe, J. (2006). The Rise of Crowdsourcing. *Wired Magazine*, *14*(6), 1–5.

Huang, Y., Singh, P. V., & Srinivasan, K. (2014). Crowdsourcing New Product Ideas Under Consumer Learning. *Management Science*, *60*(9), 2138–2159.

Huber, G. P. (1991). Organizational Learning: The Contributing Processes and the Literatures. *Organization Science*, *2*(1), 88–115.

Jacoby, J. (1977). Information Load and Decision Quality: Some Contested Issues. *Journal of Marketing Research*, *14*(4), 569–573.

Jain, A. K. (2010). Data Clustering: 50 Years Beyond K-Means. *Pattern Recognition Letters*, *31*(8), 651–666.

Jeon, J., Croft, W. B., Lee, J. H., & Park, S. (2006). A Framework to Predict the Quality of Answers with Non-Textual Features. *Proceedings of the 29th Annual International ACM SIGIR Conference*, 228–235. Seattle, Washington: ACM.

Jeppesen, L. B., & Lakhani, K. R. (2010). Marginality and Problem-Solving Effectiveness in Broadcast Search. *Organization Science*, *21*(5), 1016–1033.

Kanter, R. M. (1988). When a Thousand Flowers Bloom: Structural, Collective, and Social Conditions for Innovation in Organization. *Research in Organizational Behavior*, *10*(1988), 169–211.

Katila, R., & Ahuja, G. (2002). Something Old, Something New: A Longitudinal Study of Search Behavior and New Product Introduction. *Academy of Management Journal*, *45*(6), 1183–1194.

Katila, R., Chen, E. L., & Piezunka, H. (2012). All the Right Moves: How Entrepreneurial Firms Compete Effectively. *Strategic Entrepreneurship Journal*, *6*(2012), 116–132.

Kenny, D. A., Korchmaros, J. D., & Bolger, N. (2003). Lower Level Mediation in Multilevel Models. *Psychological Methods*, *8*(2), 115–128.

Kim, S.-M., Pantel, P., Chklovski, T., & Pennacchiotti, M. (2006). Automatically Assessing Review Helpfulness. *Proceedings of the 2006 Conference of Empirical Methods in Natural Language Processing (EMNLP 2006)*, 423–430. Sydney: Association for Computational Linguistics.

Kirton, M. (1976). Adaptors and Innovators: A Description and Measure. *Journal of Applied Psychology*, *61*(5), 622–629.

Kirton, M., & De Ciantis, S. M. (1986). Cognitive Style and Personality: The Kirton Adaption-Innovation and Cattell's Sixteen Personality Factor Inventories. *Personality and Individual Differences*, *7*(2), 141–146.

Kittur, A., Nickerson, J. V., Bernstein, Michael, S., Gerber, E. M., Shaw, A., Zimmermann, J., … Horton, J. J. (2013). The Future of Crowd Work. *Proceedings of the 16th ACM Conference on Computer Supported Cooperative Work (CSCW)*, 1–17. San Antonio: ACM.

Kristensson, P., Gustafsson, A., & Archer, T. (2004). Harnessing the Creative Potential Among Users. *Journal of Product Innovation Management*, *21*(1), 4–14.

Kuechler, B., & Vaishnavi, V. (2008). On Theory Development in Design Science Research: Anatomy of a Research Project. *European Journal of Information Systems*, *17*(5), 489–504.

Kuzel, A. (1992). Sampling in Qualitative Inquiry. In B. Crabtree & W. Miller (Eds.), *Doing Qualitative Research* (pp. 31–44). Thousand Oaks, CA: SAGE Publications.

Lamkanfi, A., Demeyer, S., Soetens, Q. D., & Verdonckz, T. (2011). Comparing Mining Algorithms for Predicting the Severity of a Reported Bug. *15th European Conference on Software Maintenance and Reengineering (CSMR)*, 249–258. Oldenburg: IEEE.

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159–174.

Langley, A. (1999). Strategies for Theorizing From Process Data. *Academy of Management Review*, *24*(4), 691–710.

Langley, A., Smallman, C., Tsoukas, H., & Van De Ven, A. H. (2013). Process Studies of Change in Organization and Management: Unveiling Temporality, Activity, and Flow. *Academy of Management Journal*, *56*(1), 1–13.

Larkin, J., McDermott, J., Simon, D., & Simon, H. A. (1980). Expert and Novice Performance in Solving Physics Problems. *Science*, *208*(4450), 1335–1342.

Laursen, K. (2012). Keep Searching and You'll Find: What Do We Know About Variety Creation Through Firms' Search Activities for Innovation? *Industrial and Corporate Change*, *21*(5), 1181–1220.

Lehrig, T., Krancher, O., & Dibbern, J. (2017). How Users Perceive and Actualize Affordances: An Exploratory Case Study of Collaboration Platforms. *Proceedings of the 38th International Conference on Information Systems (ICIS)*, 1–19. Seoul, South Korea: AIS.

Leicht, N., Blohm, I., & Leimeister, J. M. (2017). Leveraging the Power of the Crowd for Software Testing. *IEEE Software*, *34*(2), 62–69.

Leicht, N., Rhyn, M., & Hansbauer, G. (2016). Can Laymen Outperform Experts? The Effects of User Expertise and Task Design in Crowdsourced Software Testing.

Proceedings of the 24th European Conference on Information Systems (ECIS), 1–11. Istanbul, Turkey: AIS.
Leimeister, J. M., Huber, M., Bretschneider, U., & Krcmar, H. (2009). Leveraging Crowdsourcing: Activation-Supporting Components for IT-Based Ideas Competition. *Journal of Management Information Systems*, *26*(1), 197–224.
Levitt, T. (1963). Creativity is Not Enough. *Harvard Business Review*, *41*(3), 72–83.
Li, M., Kankanhalli, A., & Kim, S. H. (2016). Which Ideas are More Likely to Be Implemented in Online User Innovation Communities? An Empirical Analysis. *Decision Support Systems*, *84*(2016), 28–40.
Li, Q., Magiitti, P. G., Smith, K. G., Tesluk, P. E., & Katila, R. (2013). Top Management Team Attention to Innovation: The Role of Search Selection And Intentsity in New Product Introductions. *Academy of Management Journal*, *56*(3), 893–916.
Liaw, A. (2015). Package "randomForest." Retrieved June 14, 2016, from https://cran.r-project.org/web/packages/randomForest/randomForest.pdf
Lingo, E. L., & O'Mahony, S. (2010). Nexus Work: Brokerage on Creative Projects. *Administrative Science Quarterly*, *55*(1), 47–81.
Liu, J., Cao, Y., Lin, C.-Y., Huang, Y., & Zhou, M. (2007). Low-Quality Product Review Detection in Opinion Summarization. *Proceedings of the 2007 Join Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, 334–342. Prague: Association for Computational Linguistics.
Love, J., & Hirschheim, R. (2017). Crowdsourcing of Information Systems Research. *European Journal of Information Systems*, *26*(3), 315–332.
Lovett, M. C., & Anderson, J. R. (1996). History of Success and Current Context in Problem Solving: Combined Influences on Operator Selection. *Cognitive Psychology*, *31*(2), 168–217.
Lu, Y., Singh, P. V., & Sun, B. (2017). Is a Core-Periphery Network Good For Knowledge Sharing? A Structural Model of Endogenous Network Formation on a Crowdsourced Customer Support Forum. *MIS Quarterly*, *41*(2), 607–628.
Lukyanenko, R., Parsons, J., & Wiersma, Y. F. (2014). The IQ of the Crowd: Understanding and Improving Information Quality in Structured User-Generated Content. *Information Systems Research*, *25*(4), 669–689.
Lüthje, C., Herstatt, C., & Von Hippel, E. (2005). User-Innovators and "Local" Information: The Case of Mountain Biking. *Research Policy*, *34*(6), 951–965.
Lycett, M. (2013). 'Datafication': Making Sense of (Big) Data in a Complex World. *European Journal of Information Systems*, *22*(4), 381–386.
Magee, G. B. (2005). Rethinking Invention: Cognition and the Economics of Technological Creativity. *Journal of Economic Behavior and Organization*, *57*(1), 29–48.
Magnusson, P. R. (2009). Exploring the Contributions of Involving Ordinary Users in Ideation of Technology-Based Services. *Journal of Product Innovation Management*, *26*(5), 578–593.
Maitlis, S., & Ozcelik, H. (2004). Toxic Decision Processes: A Study of Emotion and Organizational Decision Making. *Organization Science*, *15*(4), 375–393.
Majchrzak, A., & Malhotra, A. (2013). Towards an Information Systems Perspective and Research Agenda on Crowdsourcing for Innovation. *Journal of Strategic Information Systems*, *22*(4), 257–268.

Majchrzak, A., & Malhotra, A. (2016). Effects of Knowledge-Sharing Trajectories on Innovative Outcomes in Temporary Online Crowds. *Information Systems Research*, *27*(4), 685–703.

March, J. G. (1991). Exploration and Exploitation in Organizational Learning. *Organization Science*, *2*(1), 71–87.

March, J. G. (1994). *A Primer on Decision Making: How Decisions Happen*. New York: The Free Press.

Marchand, D. A., & Peppard, J. (2013). Why IT Fumbles Analytics. *Harvard Business Review*, *91*(1–2), 104–112.

Markus, M. L. (2015). New Games, New Rules, New Scoreboards: The Potential Consequences of Big Data. *Journal of Information Technology*, *30*(1), 58–59.

Markus, M. L., Majchrzak, A., & Gasser, L. (2002). A Design Theory for Systems That Support Emergent Knowledge Processes. *MIS Quarterly*, *26*(3), 179–212.

Marsh, R. L., Ward, T. B., & Landau, J. D. (1999). The Inadvertent Use of Prior Knowledge in a Generative Cognitive Task. *Memory & Cognition*, *27*(1), 94–105.

Martin, X., & Mitchell, W. (1998). The Influence of Local Search and Performance Heuristics on New Design Introduction in a New Product Market. *Research Policy*, *26*(7–8), 753–771.

Masdeval, C., & Veloso, A. (2015). Mining Citizen Emotions to Estimate the Urgency of Urban Issues. *Information Systems*, *54*(2015), 147–155.

Mayer, R. C., Davis, J. H., & Schoorman, F. . D. (1995). An Integrated Model of Organizational Trust. *Academy of Management Review*, *20*(3), 709–734.

Meth, H., Mueller, B., & Maedche, A. (2015). Designing a Requirement Mining System. *Journal of the Association for Information Systems*, *16*(9), 799–837.

Miles, M., Huberman, M. A., & Saldaña, J. (2014). *Qualitative Data Analysis* (3rd ed.).

Mintzberg, H., Raisinghani, D., & Théorêt, A. (1976). The Structure of “Unstructured” Decision Processes. *Administrative Science Quarterly*, *21*(2), 246–275.

Morgan, D. L. (1997). *Focus Groups as Qualitative Research* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Mumford, M. D., & Gustafson, S. B. (1988). Creativity Syndrome: Integration, Application, and Innovation. *Psychological Bulletin*, *103*(1), 27–43.

Musson, R., Richards, J., Fisher, D., Bird, C., Bussone, B., & Ganguly, S. (2013). Leveraging the Crowd: How 48,000 Users Helped Improve Lync Performance. *IEEE Software*, *30*(4), 38–45.

Myers, M. D., & Newman, M. (2007). The Qualitative Interview in IS Research: Examining the Craft. *Information and Organization*, *17*(1), 2–26.

Nagar, Y., De Boer, P., & Garcia, A. C. B. (2016). Accelerating the Review of Complex Intellectual Artifacts in Crowdsourced Innovation Challenges. *Proceedings of the 37th International Conference on Information Systems (ICIS)*, 1–17. Dublin: AIS.

Nagasundaram, M., & Bostrom, R. P. (1994). The Structuring of Creative Processes Using GSS: A Framework for Research. *Journal of Management Information Systems*, *11*(3), 87–114.

Nagelkerke, N. J. D. (1991). A Note on a General Definition of the Coefficient of Determination. *Biometrika*, *78*(3), 691–692.

Nelson, R., & Winter, S. G. (1982). *An Evolutionary Theory of Economic Change*. Cambridge, Massachusetts: Harvard University Press.

Niekler, A., & Jähnichen, P. (2012). Matching Results of Latent Dirichlet Allocation for

Text. *Proceedings of the 11th International Conference on Cognitive Modeling (ICCM)*, 317–322. Groningen, Netherlands: ICCM.

Nørgaard, M., & Hornbæk, K. (2009). Exploring the Value of Usability Feedback Formats. *International Journal of Human-Computer Interaction*, *25*(1), 49–74.

Oishi, S., Kesebir, S., & Diener, E. (2011). Income Inequality and Happiness. *Psychological Science*, *22*(9), 1095–1100.

Österle, H., & Otto, B. (2010). Consortium Research. *Business & Information Systems Engineering*, *2*(5), 283–293.

Otterbacher, J. (2009). "Helpfulness" in Online Communities: A Measure of Message Quality. *Proceedings of the 27th International Conference on Human Factors in Computing Systems (CHI '09)*, 955–964. Boston: ACM.

Owyang, J. (2015). The State of Crowdsourcing in 2015. Retrieved June 27, 2017, from https://en.eyeka.com/resources/reports#CSreport2015

Parikh, M., Fazlollahi, B., & Verma, S. (2001). The Effectiveness of Decisional Guidance: An Empirical Evaluation. *Decision Sciences*, *32*(2), 303–332.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The Adaptive Decision Maker*. Cambridge: Cambridge University Press.

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, *24*(3), 45–77.

Perry-Smith, J. E., & Shalley, C. E. (2003). The Social Side of Creativity: A Static and Dynamic Social Network Perspective. *The Academy of Management Review*, *28*(1), 89–106.

Piezunka, H., & Dahlander, L. (2015). Distant Search, Narrow Attention: How Crowding Alters Organizations' Filtering of Suggestions in Crowdsourcing. *Academy of Management Journal*, *58*(3), 856–880.

Poetz, M. K., & Schreier, M. (2012). The Value of Crowdsourcing: Can Users Really Compete with Professionals in Generating New Product Ideas? *Journal of Product Innovation Management*, *29*(2), 245–256.

Porter, M. F. (1980). An Algorithm for Suffix Stripping. *Program*, *14*(3), 130–137.

Powell, W. W., Koput, K. W., & Smith-Doerr, L. (1996). Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology. *Administrative Science Quarterly*, *41*(1), 116–145.

Power, D. J. (2008). Decision Support Systems: A Historical Overview. In *Handbook on Decision Support Systems* (pp. 121–140). Berlin, Heidelberg: Springer.

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A General Multilevel SEM Framework for Assessing Multilevel Mediation. *Psychological Methods*, *15*(3), 209–233.

Ransbotham, S., Kane, G. C., & Lurie, N. H. (2012). Network Characteristics and the Value of Collaborative User-Generated Content. *Marketing Science*, *31*(3), 387–405.

Rhyn, M., & Blohm, I. (2017a). A Machine Learning Approach for Classifying Textual Data in Crowdsourcing. *Proceedings of the 13th International Conference on Wirtschaftsinformatik (WI)*, 1171–1185. St. Gallen: AIS.

Rhyn, M., & Blohm, I. (2017b). Combining Collective and Artificial Intelligence: Towards a Design Theory for Decision Support in Crowdsourcing. *Proceedings of the 25th European Conference on Information Systems (ECIS)*, 1–11. Guimarães:

AIS.
Rhyn, M., Blohm, I., & Leimeister, J. M. (2017). Understanding the Emergence and Recombination of Distant Knowledge on Crowdsourcing Platforms. *Proceedings of the 38th International Conference on Information Systems (ICIS)*, 1–21. Seoul, South Korea: AIS.
Riedl, C., Blohm, I., Leimeister, J. M., & Krcmar, H. (2013). The Effect of Rating Scales on Decision Quality and User Attitudes in Online Innovation Communities. *International Journal of Electronic Commerce*, *17*(3), 7–36.
Riedl, C., & Woolley, A. W. (2017). Teams vs. Crowds: A Field Test of the Relative Contribution of Incentives, Member Ability, and Collaboration to Crowd-Based Problem Solving Performance. *Academy of Management Discoveries*, *3*(4), 382–403.
Robinson, O. C. (2014). Sampling in Interview-Based Qualitative Research: A Theoretical and Practical Guide. *Qualitative Research in Psychology*, *11*(1), 25–41.
Rogstadius, J., Vukovic, M., Teixeira, C. A., Kostakos, V., Karapanos, E., & Laredo, J. A. (2013). CrisisTracker: Crowdsourced Social Media Curation for Disaster Awareness. *IBM Journal of Research and Development*, *57*(5), 1–13.
Rosenkopf, L., & Almeida, P. (2003). Overcoming Local Search Through Alliances and Mobility. *Management Science*, *49*(6), 751–766.
Rosenkopf, L., & Nerkar, A. (2001). Beyond Local Search: Boundary-Spanning, Exploration, and Impact in the Optical Disk Industry. *Strategic Management Journal*, *22*(4), 287–306.
Rothaermel, F. T., & Alexandre, M. T. (2009). Ambidexterity in Technology Sourcing: The Moderating Role of Absorptive Capacity. *Organization Science*, *20*(4), 759–780.
Runeson, P., Alexandersson, M., & Nyholm, O. (2007). Detection of Duplicate Defect Reports Using Natural Language Processing. *Proceedings of the 29th International Conference on Software Engineering (ICSE)*, 499–508. Minneapolis: IEEE.
Salton, G., & Buckley, C. (1988). Term Weighting Approaches in Automatic Text Retrieval. *Information Processing & Management*, *24*(5), 513–523.
Satzinger, J. W., Garfield, M. J., & Nagasundaram, M. (1999). The Creative Process: The Effects of Group Memory on Individual Idea Generation. *Journal of Management Information Systems*, *15*(4), 143–160.
Schemmann, B., Herrmann, A. M., Chappin, M. M. H., & Heimeriks, G. J. (2016). Crowdsourcing Ideas: Involving Ordinary Users in the Ideation Phase of New Product Development. *Research Policy*, *45*(6), 1145–1154.
Schenk, E., & Guittard, C. (2011). Towards a Characterization of Crowdsourcing Practices. *Journal of Innovation Economics & Management*, *7*(1), 93–107.
Schick, A. G., Gordon, L. A., & Haka, S. (1990). Information Overload: A Temporal Approach. *Accounting, Organizations and Society*, *15*(3), 199–220.
Schulten, M. B., & Schaefer, F. (2015). Affective Commitment and Customer Loyalty in Crowdsourcing: Antecedents, Interdependencies, and Practical Implications. *The International Review of Retail, Distribution and Consumer Research*, *25*(5), 516–528.
Sebastiani, F. (2002). Machine Learning in Automated Text Categorization. *ACM Computing Surveys*, *34*(1), 1–47.

Sein, M. K., Henfridsson, O., Rossi, M., & Lindgren, R. (2011). Action Design Research. *MIS Quarterly*, *35*(1), 37–56.

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.

Shim, J. P., Warkentin, M., Courtney, J. F., Power, D. J., Sharda, R., & Carlsson, C. (2002). Past, Present, and Future of Decision Support Technology. *Decision Support Systems*, *33*(2), 111–126.

Siau, K., & Wang, W. (2018). Building Trust in Artificial Intelligence, Machine Learning, and Robotics. *Cutter Business Technology Journal*, *31*(2), 47–53.

Silk, E. M., Daly, S. R., Jablokow, K. W., & McKilligan, S. (2019). Incremental to Radical Ideas: Paradigm-Relatedness Metrics for Investigating Ideation creativity and Diversity. *International Journal of Design Creativity and Innovation*, *7*(1–2), 30–49.

Silver, M. S. (1988). User Perceptions of Decision Support System Restrictiveness: An Experiment. *Journal of Management Information Systems*, *5*(1), 51–65.

Silver, M. S. (1990). Decision Support Systems: Directed and Nondirected Change. *Information Systems Research*, *1*(1), 47–70.

Silver, M. S. (1991). Decisional Guidance for Computer-Based Decision Support. *MIS Quarterly*, *15*(1), 105–122.

Simon, H. A. (1960). *The New Science of Management Decision*. New Jersey: Prentice Hall.

Simon, H. A. (1979). Rational Decision Making in Business Organizations. *The American Economic Review*, *69*(4), 493–513.

Simon, H. A. (1990). Invariants of Human Behavior. *Annual Review of Psychology*, *41*(1), 1–19.

Simula, H., & Ahola, T. (2014). A Network Perspective on Idea and Innovation Crowdsourcing in Industrial Firms. *Industrial Marketing Management*, *43*(3), 400–408.

Smith, W. K. (2014). Dynamic Decision Making: A Model of Senior Leaders Managing Strategic Paradoxes. *Academy of Management Journal*, *57*(6), 1592–1623.

Spence, M. T., & Brucks, M. (1997). The Moderating Effects of Problem Characteristics on Experts' and Novices' Judgments. *Journal of Marketing Research*, *34*(2), 233.

Sprague, R. H. (1980). A Framework for the Development of Decision Support Systems. *MIS Quarterly*, *4*(4), 1–26.

Sproles, N. (2001). The Difficult Problem of Establishing Measures of Effectiveness for Command and Control: A Systems Engineering Perspective. *Systems Engineering*, *4*(2), 145–155.

Stephen, A. T., Zubcsek, P. P., & Goldenberg, J. (2016). Lower Connectivity Is Better: The Effects Of Network Structure On Redundancy Of Ideas And Customer Innovativeness In Interdependent Ideation Tasks. *Journal of Marketing Research*, *53*(2), 263–279.

Stol, K.-J., LaToza, T. D., & Bird, C. (2017). Crowdsourcing for Software Engineering. *IEEE Software*, *34*(2), 30–36.

Stuart, T. E., & Podolny, J. M. (1996). Local Search and the Evolution of Technological Capabilities. *Strategic Management Journal*, *17*(1996), 21–38.

Swain, M. R., & Haka, S. F. (2000). Effects of Information Load on Capital Budgeting

Decisions. *Behavioral Research in Accounting*, *12*(2), 171–198.

Tan, A.-H. (1999). Text Mining: The State of the Art and the Challenges. *Proceedings of the 3rd Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD'99)*, 65–70. Beijing, China.

Theodorou, A., Wortham, R. H., & Bryson, J. J. (2017). Designing and Implementing Transparency for Real Time Inspection of Autonomous Robots. *Connection Science*, *29*(3), 230–241.

Thuan, N. H., Antunes, P., & Johnstone, D. (2016). Factors Influencing the Decision to Crowdsource: A Systematic Literature Review. *Information Systems Frontiers*, *18*(1), 47–68.

Todd, P., & Benbasat, I. (1999). Evaluating the Impact of DSS, Cognitive Effort, and Incentives on Strategy Selection. *Information Systems Research*, *10*(4), 356–374.

Tripsas, M., & Gavetti, G. (2000). Capabilities, Cognition, and Inertia: Evidence From Digital Imaging. *Strategic Management Journal*, *21*(10), 1147–1161.

Trope, Y., & Liberman, N. (2010). Construal-Level Theory of Psychological Distance. *Psychological Review*, *117*(2), 440–463.

Tsoukas, H. (2009). A Dialogical Approach to the Creation of New Knowledge in Organizations. *Organization Science*, *20*(6), 941–957.

Tushman, M. L. (1977). Special Boundary Roles in the Innovation Process. *Administrative Science Quarterly*, *22*(4), 587–605.

Tushman, M. L., & Katz, R. (1980). External Communication and Project Performance: An Investigation into the Role of Gatekeepers. *Management Science*, *26*(11), 1071–1085.

Tyler, T. R., Rasinski, K. A., & Spodick, N. (1985). Influence of Voice on Satisfaction With Leaders: Exploring the Meaning of Process Control. *Journal of Personality and Social Psychology*, *48*(1), 72–81.

Uzzi, B., & Spiro, J. (2005). Collaboration and Creativity: The Small World Problem. *American Journal of Sociology*, *111*(2), 447–504.

van Knippenberg, D., De Dreu, C. K. W., & Homan, A. C. (2004). Work Group Diversity and Group Performance: An Integrative Model and Research Agenda. *Journal of Applied Psychology*, *89*(6), 1008–1022.

Vandenbosch, B., & Huff, S. L. (1997). Searching and Scanning: How Executives Obtain Information from Executive Information Systems. *MIS Quartely*, *21*(1), 81–107.

Venable, J., Pries-Heje, J., & Baskerville, R. (2016). FEDS: A Framework for Evaluation in Design Science Research. *European Journal of Information Systems*, *25*(1), 77–89.

Von Hippel, E. (2005). *Democratizing Innovation*. Cambridge, Massachusetts: MIT Press.

Vosoughi, S., Roy, D., & Aral, S. (2018). The Spread of True and False News Online. *Science*, *359*(6380), 1146–1151.

Walls, J. G., Widmeyer, G. R., & El Sawy, O. A. (1992). Building an Information System Design Theory for Vigilant EIS. *Information Systems Research*, *3*(1), 36–59.

Walter, T. P., & Back, A. (2013). A Text Mining Approach to Evaluate Submissions to Crowdsourcing Contests. *Proceedings of the 46th Hawaii International Conference on System Sciences (HICSS)*, 3109–3118. Waikoloa, Hawaii: IEEE.

Wang, L., Jamieson, G. A., & Hollands, J. G. (2009). Trust and Reliance on an Automated Combat Identification System. *Human Factors*, *51*(3), 281–291.

Wang, R. Y., & Strong, D. M. (1996). Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems*, *12*(4), 5–34.

Wang, W., & Benbasat, I. (2005). Trust in and Adoption of Online Recommendation Agents. *Journal of the Association for Information Systems*, *6*(3), 72–101.

Wang, W., & Benbasat, I. (2009). Interactive Decision Aids for Consumer Decision Making in E-Commerce: The Influence of Perceived Strategy Restrictiveness. *MIS Quarterly*, *33*(2), 293–320.

Weimer, M., & Gurevych, I. (2007). Predicting the Perceived Quality of Web Forum Posts. *Proceedings of the 2007 Conference on Recent Advances in Natural Language Processing (RANLP)*, 643–648. Borovets, Bulgaria.

Willcocks, L. P., & Lester, S. (1999). *Beyond the IT Productivity Paradox*. New York: Wiley.

Winter, S. G. (1984). Schumpeterian Competition in Alternative Technological Regimes. *Journal of Economic Behavior and Organization*, *5*(3–4), 287–320.

Witte, E. (1972). Field Research on Complex Decision-Making Processes - The Phase Theorem. *International Studies of Management & Organization*, *2*(2), 156–182.

Ye, H. J., Blohm, I., Bretschneider, U., Goswami, S., Leimeister, J. M., & Krcmar, H. (2016). Promoting the Quality of User Generated Ideas in Online Innovation Communities: A Knowledge Collaboration Perspective. *Proceedings of the 37th International Conference on Information Systems (ICIS)*, 1–16. Dublin, Ireland: AIS.

Zhang, Q., Zeng, D. D., Wang, F.-Y., Breiger, R., & Hendler, J. A. (2016). Brokers or Bridges? Exploring Structural Holes in a Crowdsourcing System. *IEEE Computer*, *49*(6), 56–64.

Zhao, Y., & Zhu, Q. (2014). Evaluation on Crowdsourcing Research: Current Status and Future Direction. *Information Systems Frontiers*, *16*(3), 417–434.

Zogaj, S., Bretschneider, U., & Leimeister, J. M. (2014). Managing Crowdsourced Software Testing: A Case Study Based Insight on the Challenges of a Crowdsourcing Intermediary. *Journal of Business Economics*, *84*(3), 375–405.

Zuchowski, O., Posegga, O., Schlagwein, D., & Fischbach, K. (2016). Internal Crowdsourcing: Conceptual Framework, Structured Review, and Research Agenda. *Journal of Information Technology*, *31*(2), 166–184.

Zwass, V. (2010). Co-Creation: Toward a Taxonomy and an Integrated Research Perspective. *International Journal of Electronic Commerce*, *15*(1), 11–48.

CURRICULUM VITAE

Personal Information

Name	Marcel Rhyn
Date of Birth	26.04.1990
Nationality	Swiss

Education

2016 – 2020	Ph.D. in Management
	University of St.Gallen (HSG)
2013 – 2016	Master of Arts in Business Management
	University of St.Gallen (HSG)
2009 – 2013	Bachelor of Arts in Business Administration
	University of St.Gallen (HSG)
2003 – 2009	Gymnasialmatura
	Kantonsschule Alpenquai Luzern

Professional Experience

2015-2020	Research Associate
	Institute of Information Management, University of St.Gallen
2013	Business Model Innovation Intern
	SAP (Switzerland) Inc., St.Gallen
2012	Research Assistant
	Institute of Technology Management, University of St.Gallen
2011-2012	Web Designer
	Eve's Art GmbH, Zurich
2011	Consulting Intern
	Swissconsult/Inopec AG, Zurich