Text- and Data-Mining of Ideation.
Instruments for the Management of Crowdsourcing-Platforms.

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

Thomas Pierre Walter

from

Germany

Approved on the application of

Prof. Dr. Andrea Back

and

Prof. Christian Wagner, PhD

Dissertation no. 4249

Difo-Druck GmbH, Bamberg 2014
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, October 21, 2013

The President:

Prof. Dr. Thomas Bieger
Acknowledgements

Special thanks belong to my academic supervisor Prof. Dr. Andrea Back, who gave me the opportunity to research an interesting subject that genuinely intrigues me. Furthermore my thanks regards her openness to explore and apply rather unconventional methods such as text- and data-mining. I would also like to thank Prof. Dr. Christian Wagner for his willingness to co-supervise my thesis, his invitation to become a visiting researcher at the City University of Hong Kong as well as his inspiring suggestions to my research ideas.

My time as a PhD student in St.Gallen was most importantly shaped by my colleagues. I am pleased to thank all those who have become true friends during this doctoral program and who have made my PhD time a joyful experience. Thank you for the many interesting discussions, for the fun also outside the office and the wonderful memories.

Above all, I thank my family, my parents Marlies and Georg and my sister Julia for their unconditional support and care. I also want to express my gratitude to my partner, Anja. During the time she was a consistent sounding board to help me work through a variety of technical as well as professional questions. Thanks for your benevolent helpfulness and patience.

Hong Kong,
December 2012
Contributions

Contribution A:
Thomas P Walter, Crowdsourcing-Plattformen, in Andrea Back, Norbert Gronau, Klaus Tochtermann (Hrsg.) Web 2.0 in der Unternehmenspraxis, 3. Auflage, Oldenbourg, München, p. 73-82.

Contribution B:

Contribution C:

Contribution D:

Contribution E:
Thomas P Walter, Andrea Back, A Text Mining Approach to Evaluate Submissions to Crowdsourcing Contests, in 46th Hawaii International Conference on System Sciences (HICSS), Wailea, HI, USA, January 2013.

Contribution F:
List of Tables/ Tabellenverzeichnis

Part A
Table 1 Crowdsourcing-platforms as an IT artifact.............................. 18
Table 2 Overview of the first set of research questions........................ 24
Table 3 Overview of the second set of research questions...................... 25
Table 4 Results from Database- and Journal-Search.......................... 33
Table 5 Results from keyword search........................................... 33
Table 6 Results from the backward search..................................... 35
Table 7 Research streams detected within publications from the literature review.................................................. 38
Table 8 Research framework....................................................... 46

Part B
Contribution A
Table 1 Bibliografische Informationen für Beitrag A.......................... 73

Contribution B
Table 1 Bibliographical Information for Contribution B..................... 85
Table 2 Characteristics of business models using crowdsourcing as a key resource......................................................... 88
Table 3 Findings of characteristics of business models using within case studies.......................................................... 93
Table 4 Questions of further research derived from case study research.............................................................. 96

Contribution C
Table 1 Bibliographical Information for Contribution C..................... 101
Table 2 Dependent and independent variables and methods of measurement.............................................................. 110
Table 3 OLS Regression Analysis for Modelling the Amount and Quality in Online Idea Contests........................................... 114

Contribution D
Table 1 Bibliografische Informationen für Beitrag D....................... 123

Contribution E
Table 1 Bibliographical Information for Contribution E..................... 136
Table 2 Summary of background literature.................................... 139
Table 3 Metrics of submissions to the crowdsourcing website............ 141
Table 4 Process of a text mining based evaluation of crowdsourcing
Table 5  Selection of submissions made by the text mining approach compared to expert decisions

Table 6  Precision, Recall and F1-Scores for all models

Contribution F

Table 1  Bibliographical information for Contribution F
Table 2  Results from a literature search for applied research methods within the field of ideation
Table 3  Summary of background literature by representative findings and identified research gap
Table 4  Procedure model steps to analyze ideation contest with data mining and SNA techniques
Table 5  Descriptive statistics of the dataset
Table 6  Ideation concept network characteristics, split by two halves of an ideation contest
Table 7  Comparison of concept network metrics for rewarded and unrewarded ideas
List of Figures/ Abbildungsverzeichnis

Part A

Figure 1  Interest in topics over time, determined by term frequencies used in Google queries. The number 100 represents the peak search volume between compared terms and within the compared time. The figure was drawn using www.google.com/trends................................................................. 19

Figure 2  Exemplary database-model of a crowdsourcing-platform................................................................. 20

Figure 3  Exemplary business model of a crowdsourcing-platform, illustrated by using the business model canvas from Osterwalder & Pigneur, 2010)................................................................. 23

Figure 4  Overview to the structure of this dissertation................................................................. 27

Figure 5  Crowdsourcing concept map based on (Walter, 2012)................................................................. 31

Figure 6  Illustration of the literature search process................................................................. 32

Figure 7  Publications by year and output format................................................................. 36

Figure 8  Publication by keyword set and output format................................................................. 36

Figure 9  Publications by research method and research objective................................................................. 37

Figure 10  Procedure model to detect novel ideas in ideation contests by using a text-mining approach (as explained in contribution E)................................................................. 52

Figure 11  Procedure model to analyze and visualize the ideation function of ideation contests by using a data-mining approach (as shown in contribution F)................................................................. 53

Figure 12  Generalized results on the ideation function (as explained in contribution F)................................................................. 55

Part B

Contribution A

Abbildung 1  Crowdsourcing Concept Map................................................................. 74
Abbildung 2  Crowdsourcing Plattform Typen................................................................. 76
Abbildung 3  Entscheidungsbaum für die Auswahl einer geeigneten Crowdsourcing Plattform................................................................. 81

Contribution C

Figure 1  Hypothesis of potential external factors on the outcome of online idea contests................................................................. 109

Figure 2  Effects of external Factors on the average quality of rated ideas in online idea contests................................................................. 115

Figure 3  Effects of external Factors on the amount of submitted ideas in online idea contests................................................................. 116
Contribution D
Abbildung 1  Phasen des Text Mining und beispielhafte Techniken, eigene Darstellung auf Basis von (Feldmann und Sanger 2007, S. 1 sowie Miller 2005, S. 104).......................................................... 127
Abbildung 2  Screenshot der Funktion Import Bearbeiten.......................... 130
Abbildung 3  Screenshot der Dashboard Funktion Sentiment Analyse........ 131
Abbildung 4  Screenshot der Funktion Feedback bearbeiten.................. 131
Abbildung 5  Ergebnisse der Genauigkeit der Text Mining Analysetechniken 132

Contribution E
Figure 1   Typical answer to a crowdsourcing contest, shown as raw text data within the QDA Miner software................................. 142
Figure 2  Excerpt of similarity index matrix showing cosine coefficients between submissions to a crowdsourcing contest............. 144
Figure 3  Dendrogram of clustered submissions within a crowdsourcing contest............................................................. 145

Contribution F
Figure 1   Analogy of an ideation contest as black box.......................... 156
Figure 2  Ideation functions implied by different theories..................... 159
Figure 3  Procedure model........................................................................ 164
Figure 4  Concept maps based on CDM, clustering and MDS after ten, 20, 30 and 40 of 725 ideas (top left to bottom right).................. 171
Figure 5  Concept maps of an ideation contests after N=362 (left) and 725 (right) ideas................................................................. 172
Figure 6  Ideation contest characteristics (ideation function).................. 174
Figure 7  Cluster development of concepts co-occurring with the centered concept of Versioning/Customization at N=2, 5, 10 and 15 ideas using the centered concept........................................ 175
Figure 8  Cluster development of concepts co-occurring with the centered concept of Versioning/Customization at N=25, 35, 45, 55, 65 and 75 (from top left to bottom right)................................. 176
Figure 9  Ideation function of a centered concept within an ideation contest........................................................................... 177
Figure 10 Idea network, based on coding. Rewarded ideas in yellow, edges to 10 nearest neighbors in red......................................... 180
Table of Contents
Acknowledgements ............................................................................................................. 3
Contributions ..................................................................................................................... 4
List of Tables/ Tabellenverzeichnis ....................................................................................... 5
List of Figures/ Abbildungsverzeichnis ............................................................................... 7
Table of Contents ............................................................................................................... 9
Summary ............................................................................................................................. 14
Kurzfassung ......................................................................................................................... 15

Part A.................................................................................................................................. 16

1 Exposition ......................................................................................................................... 17
   1.1 Introduction and Motivation ..................................................................................... 17
   1.2 Detailed Problem Description and Research Questions ........................................... 22
   1.3 Contribution and Audience .................................................................................... 26
   1.4 Thesis Structure ....................................................................................................... 27

2 Research Foundation ....................................................................................................... 28
   2.1 Revisiting the Literature Review on Crowdsourcing ............................................... 29
   2.2 Scoping the Literature Review .............................................................................. 29
   2.3 Conceptualizing Relevant Search Terms ................................................................ 30
   2.4 The Literature Search Process ................................................................................ 31
   2.5 Statistical Literature Stock Analysis ........................................................................ 35
   2.6 Literature Streams ................................................................................................... 37
   2.7 Defined Research Gaps ........................................................................................... 38

3 Research Framework ....................................................................................................... 39
   3.1 An Explanatory Research Approach ....................................................................... 39
   3.2 A Multi-Method Research Design .......................................................................... 41
      3.2.1 Case Study Research ....................................................................................... 42
      3.2.2 Structural Equation Modeling ......................................................................... 42
      3.2.3 Text- and Data-Mining .................................................................................. 43
      3.2.4 Social Network Analysis (SNA) ..................................................................... 44
   3.3 Development of the Research Framework ................................................................ 45

4 Research Results ............................................................................................................... 46
   4.1 Critical Success Factors of Crowdsourcing-Platforms ............................................ 46
      4.1.1 Types and Characteristics of Crowdsourcing-Platforms .................................. 47
      4.1.2 Crowdsourcing-Platforms and Collective Intelligence ..................................... 48
      4.1.3 Critical Design Elements of Crowdsourcing-Platforms .................................. 49
4.1.4 Quantity and Quality on Crowdsourcing-Platforms .......................... 49
4.2 Applying Text- and Data Mining and SNA to detect Quality on Crowdsourcing-Platforms ................................................................. 50
   4.2.1 Relevant Text- and Data-Mining Techniques ................................ 51
   4.2.2 Applying Text- and Data-Mining Techniques towards Crowdsourcing-Platforms ................................................................. 52
   4.2.3 Semi-Automated Detection of Quality on Crowdsourcing-Platforms ...... 53
   4.2.4 Measuring the Ideation Function of an Ideation Contest ...................... 54
   4.2.5 Cognitive Abilities of Solver Crowds ........................................... 56
5 Critical Reflection and Closure ................................................................ 57
   5.1 Theoretical Contribution .................................................................... 58
   5.2 Managerial Implications .................................................................... 60
   5.3 Critical Reflection and Final Personal Opinion ....................................... 61
6 Literature ................................................................................................ 63

Part B .......................................................................................................... 72

Contribution A: Crowdsourcing-Plattformen ........................................... 73
   A.1 Einleitung ........................................................................................... 73
   A.2 Das Crowdsourcing Konzept .............................................................. 73
   A.3 Kategorien von Crowdsourcing Plattformen ...................................... 76
      A.3.1 Crowdsourcing Makler ................................................................. 76
      A.3.2 Direktes Crowdsourcing über die Unternehmenswebseite ............ 78
      A.3.3 Verkaufsplattformen für Crowdsourcing Design ......................... 79
      A.3.4 Unternehmensinternes Crowdsourcing ........................................ 79
   A.4 Analyse der Crowdsourcing Plattformen ........................................... 80
      A.4.1 Stärken von Crowdsourcing Plattformen ..................................... 81
      A.4.2 Schwächen und Bedrohungen von Crowdsourcing Plattformen ...... 81
      A.4.3 Möglichkeiten und Chancen von Crowdsourcing Plattformen ...... 82
   A.5 Literatur .............................................................................................. 83

Contribution B: Crowdsourcing as a Business Model: An Exploration of Emergent Textbooks Harnessing the Wisdom of Crowds ....................... 85
   B.1 Introduction ....................................................................................... 85
   B.2 Crowdsourcing as Business Model Innovation .................................... 87
      B.2.1 Business Model Definition as Framework .................................... 87
B.2.2 Crowdsourcing as Key Factor of a Business Model ........................................ 88
B.3 Case Studies ........................................................................................................... 89
  B.3.1 Charles Leadbeater – “We think” ..................................................................... 89
  B.3.2 Osterwalder and Pigneur – “Business Model Generation” ......................... 90
  B.3.3 Crumlish and Malone – “Designing Social Interfaces” .................................. 91
  B.3.4 Williams and Tapscott – “The Wikinomics Playbook” .................................. 91
B.4 Analysis and Shaping Hypothesis .......................................................................... 92
  B.4.1 Book Topics ...................................................................................................... 93
  B.4.2 Crowdsourcing as value proposition ............................................................... 93
  B.4.3 Crowds ............................................................................................................... 94
  B.4.4 Participation Hurdles ....................................................................................... 94
  B.4.5 Incentives .......................................................................................................... 95
  B.4.6 Technical solutions .......................................................................................... 95
B.5 Conclusion ............................................................................................................... 96
B.6 References .............................................................................................................. 97

Contribution C: Towards Measuring Crowdsourcing Success: An Empirical Study on Effects of External Factors in Online Idea Contests ................................. 101
C.1 Introduction ............................................................................................................. 102
C.2 Theoretical Background .......................................................................................... 103
  C.2.1 Classification of relevant terms ........................................................................ 103
  C.2.2 Effects of External Factors on the Outcome of Idea Contests ...................... 105
    C.2.2.1 Open Innovation and Crowdsourcing Research ........................................ 106
    C.2.2.2 Research on FOSS Communities and Electronic Brainstorming .......... 107
    C.2.2.3 Economic and Sociologic Research ......................................................... 107
C.3 Hypothesis and Prediction Model ........................................................................... 108
  C.3.1 Hypothesis ........................................................................................................ 108
  C.3.2 Variables Measurement .................................................................................... 109
  C.3.3 The Model ......................................................................................................... 111
C.4 Data Collection and Descriptive Analysis ................................................................ 111
  C.4.1 The Atizo Platform .......................................................................................... 111
  C.4.2 Descriptive Statistics of Online Idea Contests on the Atizo Platform .......... 112
C.5 Results and Analysis ............................................................................................... 113
C.6 Conclusions and Limitations .................................................................................... 115
C.7 References .............................................................................................................. 117
Contribution D: Controlling von Kundenmeinungen durch Text Mining: Entwicklung einer Lösung zur Analyse von In-App-Feedback .......................... 123
D.1 Kundenfeedback in Apps .................................................................................................................. 124
  D.1.1 Problemstellung: Kundenverhalten bei App-Feedback-Funktionalität ............................. 124
  D.1.2 Anforderungen an das Controlling ....................................................................................... 125
D.2 Grundlagen des Text Mining ......................................................................................................... 126
  D.2.1 Pre-Processing ....................................................................................................................... 127
  D.2.2 Text Mining Analysetechniken ............................................................................................... 128
  D.2.3 Visualisierung ......................................................................................................................... 129
D.3 Entwicklung einer Lösung für das Feedback-Controlling .......................................................... 129
D.4 Ergebnisse ........................................................................................................................................ 132
  D.4.1 Evaluation der Genauigkeit der Lösung ................................................................................. 132
  D.4.2 Bedeutung für das Controlling .............................................................................................. 133
D.5 Fazit .................................................................................................................................................. 134
D.6 Literatur ............................................................................................................................................ 134

Contribution E: A Text Mining Approach to Evaluate Submissions to Crowdsourcing Contests .......................................................................................................................... 136
E.1 Introduction ........................................................................................................................................ 136
E.2 Background ...................................................................................................................................... 138
  E.2.1 Success Patterns of Crowdsourcing ......................................................................................... 138
  E.2.2 Measuring the Quality of Crowdsourcing ............................................................................... 139
  E.2.3 Using Text Mining to Measure the Quality of Crowdsourcing ............................................ 139
E.3 Methodology .................................................................................................................................... 140
  E.3.1 Dataset and descriptive Statistics ............................................................................................ 141
  E.3.2 Pre-Processing Crowdsourcing website data ......................................................................... 142
  E.3.3 Clustering Submissions to Crowdsourcing Contests ............................................................... 143
  E.3.4 A Text Mining Based Evaluation of Submissions to Crowdsourcing Contests .................. 146
E.4 Results ............................................................................................................................................... 146
E.5 Discussion and Conclusion ............................................................................................................. 148
E.6 References ....................................................................................................................................... 150

Contribution F: Data Mining and Social Network Analysis of Ideation Contests: A Repeated Measures Design .............................................................................................................. 155
F.1 Introduction ....................................................................................................................................... 155
Summary

This dissertation deals with the emerging internet phenomenon of crowdsourcing-platforms. Crowdsourcing has attracted worldwide attention of both, the practitioner and the scientific community since it was introduced as the new pool of cheap labor in 2006: Everyday people using their spare cycles to create content, solve problems, even do corporate R&D by using crowdsourcing-platforms in the internet. Growing ever since, today crowdsourcing-platforms offer global workforces (called solvers) to firms (called seekers) which intend to outsource routine tasks, human intelligence tasks as well as creative thinking and brain teasers. At the same time, crowdsourcing-platforms are easy to maintain from a technical perspective. The cost of storing and processing data has decreased dramatically over recent years and hence, the amount of data stored in electronic form has grown at an explosive rate. In the broadest sense this dissertation deals with problems of managing crowdsourcing-platforms, especially when data is increasing rapidly in both, volume and complexity. Today, seekers are more than often faced with an unmanageable amount of data, foremost in form of a broad variety of solvers submissions like ideas, concepts, drafts or suggested solutions. Next to keeping track of all submissions, the task of analysis and evaluation represents the biggest issue of managing crowdsourcing-platforms. Driven by practitioner’s problems, this dissertation carves out strategic as well as technical solutions.

Consequently, this accumulative dissertation is split into two parts. Part A introduces the general motivation, sounds all research questions, the aimed audience and the thesis’ structure. The research questions are backed up by a literature review showing a current research gap. Furthermore Part A explains the chosen research approaches and summarizes all findings of this dissertation. It closes with the theoretical contributions, a managerial impact and a critical reflection. Part B presents all six research papers in full length. All contributions are published or submitted to academic outlets. Basically, contributions A, B and C can be defined as common research on the topic of crowdsourcing. The result can be coined to the issue of crowdsourcing success patterns. Contributions D, E and F address the more specific problem of measuring, evaluating and visualizing contributions made by solvers. Essentially, text- and data-mining methods as well as social network analysis is applied to search for highly valuable ideas, to detect hidden patterns inside ideation contest and to visualize and analyze the ideation function, that is to say the process of knowledge aggregation within brainstorming processes.

In sum the results of this dissertation enlarge the knowledgebase of Information Systems research. It stresses the value of text- and data-mining as valuable and yet underused research methods and illustrates that modern processes of ideation might not be able to be explained by hitherto predominant theories.
Kurzfassung


In Summe trägt diese Dissertation zur Wissensbasis der Wirtschaftsinformatik im weiteren Sinne bei. Die Beiträge betonen den Wertbeitrag von Text- und Data-Mining als wegewisende Methode und zeigen, dass bisher anhand der Theorie zu gruppenbasierten Gestaltungsprozessen nicht mehr zeitgemäss ist.
Part A
1 Exposition

“Having spent a lifetime analyzing the game of chess and comparing the capacity of computers to the capacity of the human brain, I’ve often wondered, where does our success come from? The answer is synthesis, the ability to combine creativity and calculation, art and science, into a whole that is much greater then the sum of its parts.”

(Kasparov, 2007)

1.1 Introduction and Motivation

The process of translating an idea or invention into a good or service that creates value or for which customers will pay. The generic definition of innovation, used by patent offices around the world, endows the situation that innovation does not appear from nowhere, but rather requires a distinct process of translating initial thoughts, ideas, sketches, drafts or concepts into the final goods or services that we later are willing to consume. On account of this, innovation translating processes have been a grateful research objective used by a variety of scholars and in multiple research disciplines. Scholars strive to figure out the genome of producing innovation, the essential determinants of being innovative or creative, ever since. Likewise, throughout centuries the desired research approach is to decrypt innovation translating processes in order to make them repeatable, manageable, measurable and improvable.

The research field of information systems (IS) plays a significant role in this endeavor. IS is the study of complementary networks of hardware and software (information technology, IT) that people and organizations use to collect, filter, process, create, and distribute data. In general, IS focuses upon processing information within organizations, especially within business enterprises, and sharing the benefits with modern society. In the particular case of the aforementioned innovation translating processes, IS has been keen to provide IT artifacts which support enterprise R&D processes. IT artifacts are defined as the application of IT to enable or support a task embedded within a business structure that itself is embedded within a corporate context. As (Benbasat & Zmud, 2003) state, IS scholars and IS practitioners strive to increase their collective understandings of how IT artifacts are conceived, constructed, and implemented, how IT artifacts are used, supported, and evolved, and how IT artifacts impact (and are impacted by) the contexts in which they are embedded.

The research objective of this dissertation is crowdsourcing-platforms, which constitute a modern IT artifact. A crowdsourcing-platform is a web-based platform on which an enterprise (in this context called seeker) can make tasks of corporate R&D accessible to the public. Usually a crowd of voluntarily participating internet users (called solvers) tries to solve these tasks by submitting thoughts, ideas, sketches, drafts, concepts or solutions to the crowdsourcing-platform. In many cases
Crowdsourcing is conducted in form of a tournament (and therewith often called ideation contest). This refers to the situation, that the seeker usually pays a reward to the best solvers within an ideation contest. Table 1 conceptualizes the IT artifact of crowdsourcing-platforms from an enterprise’s perspective using the framework of (Benbasat & Zmud, 2003).

| Information Technology | • A web-based crowdsourcing-platform to support enterprise R&D processes.  
| Task Structure | • Ideation messaging threading by software as a service.  
| • Database of current and historical crowdsourcing events.  
| • The crowd is connected via a Social Network Service (SNS) and accesses the platform via multiple software devices.  
| Task | • Outsourcing corporate R&D to a crowd (a large and merely unknown workforce of non-experts in the internet).  
| Task Context | • Extending the formal enterprise R&D process and selecting an appropriate crowdsourcing-platform.  
| • Defining the task to be outsourced as well as start date, duration, reward, of the ideation contest.  
| • Evaluating all ideas that solvers submitted to the ideation contest, selecting winning ideas and allotting rewards.  
| • Receive the ownership of intellectual property (ideas) and assimilate results from ideation contest into the corporate R&D process.  
| • Enterprise and divisional values and norms.  
| • Industry and firm business conditions.  
| • External and internal jolts.  

*Table 1 Crowdsourcing-platforms as an IT artifact.*

Crowdsourcing-platforms are by no means the first appearance of an IT artifact which intention it is to leverage corporate R&D. Eventually best known beforehand are electronic meeting systems (Nunamaker, Dennis, Valacich, Vogel, & George, 1991) and electronic brainstorming systems (EBS) respectively (Dennis & Valacich, 1993). With EBS, members used computers to interact and exchange ideas. As in verbal brainstorming (Osborn, 1957), electronic interaction enabled ideas from one user to stimulate new ideas in others. Electronic interaction also exhibits characteristics of nominal group brainstorming, in that participants submit ideas anonymously and cannot be blocked from contributing ideas, as it did sometimes occur in verbal brainstorming. Technically spoken, crowdsourcing-platforms do exactly the same as EBS. Hence, the question is why practitioners as well as researchers recently speak about crowdsourcing and no longer about EBS? The answer is manifold but still simple. First, terms used in IS research are subject to natural trends and fluctuations. A comparison by the term frequencies used within Google searches shows that
crowdsourcing \(^{(1)}\) gained popularity ever since the term was coined by Wired magazine columnist Jeff Howe in 2006 (Howe, 2006). Recently it outdistances open innovation \(^{(2)}\), coined by Henry Chesbrough in 2003 (Chesbrough, 2003) as well as brainstorming in general \(^{(3)}\) and ideation contest \(^{(4)}\).

Second, in contrast to the hitherto prevailing EBS, crowdsourcing does merely not account for technological, but rather for cultural change. As part of the much-cited paradigm shift towards the Web 2.0, technological and cultural changes came in unison. Based on the foundation of the Internet’s open technology standards like HTML, CSS and AJAX, web-based software like weblogs, wikis or social network services (SNS) were developed (Musser & O’Reilly, 2007). Moreover, the term Web 2.0 relates to changes in the ways software developers and end-users use the Web, and a key characteristic being user generated content. And as (Musser & O’Reilly, 2007) state, this trend towards networked applications is accelerating. While Web 2.0 has initially taken hold in consumer-facing applications (like facebook, twitter, wordpres, Wikipedia, flickr or del.icio.us) the infrastructure required to build these applications, and the scale at which they are operating, means that, much as PCs took over from mainframes in a classic demonstration of the “innovator dilemma” hypothesis stated by (Christensen, 1997), web applications can and will move into the enterprise space (the latter was coined as Enterprise 2.0 by Andrew McAfee in 2007 (McAfee, 2009)). Crowdsourcing-platforms are built upon the technology that is allotted to the Web 2.0 paradigm. In reverse the corporate usage of a crowdsourcing-platform can be called Enterprise 2.0 per se. Revisiting the initial definition of crowdsourcing by (Howe, 2006) emphasizes this point:

Simply defined, crowdsourcing represents the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call. This can take the form of peer-production (when the job is performed collaboratively), but is also often
undertaken by sole individuals. The crucial prerequisite is the use of the open call format and the large network of potential laborers.

The database model of an exemplary crowdsourcing-platform, illustrated in Figure 2, gives reasoning to the process, which is triggered by the mentioned open call. Platform users are connected (c.f. cs_platform.contact in Figure 2) alike in a SNS. Next to maintaining a profile on the crowdsourcing platform, their main intention is to post ideas (cs_platform.idea) to various ideation contests that are offered on the platform (cs_platform.ideation_contest). Additionally users can comment on other user’s ideas (cs_platform.comment). Considering the attributes shows that ideation contests include detailed task descriptions, distinct timeslots, specified rewards and seekers which can be allotted to certain branches. User attributes include their demographic information (age, gender, language, profession, interests and past experience) and idea attributes include the information who submitted an idea at what time, in which ideation contest, as well as how this idea was rated by other users and rewarded by the ideation contest’s seeker.

Figure 2 Exemplary database-model of a crowdsourcing-platform.

In simple terms, this database model narrows down the research objective and motivation to this dissertation. Simplicity, accessibility, ease of use and the rise of Web 2.0 in general lead to a rising popularity of crowdsourcing-platforms in public. Basically, this is beneficial for platform providers and seekers as they can tap onto more ideas by a larger variety of solvers. However, the rising amount of data that is created within a single ideation contests constitute a problem for seekers and the
platform provider. For instance, ideation contests often achieve up to 1000 ideas, an amount that can be similar up to 500 pages of plain text data. Usually expert committees have to rate all ideas, select winners and adjust monetary rewards. However, they often are unable to cope with this tremendous quantity and complexity of data. Understandably, they mostly are unwilling to go through such large amounts of data manually. From a practitioner’s perspective, the difficulties embodied in understanding, managing, analyzing and evaluating data from ideation contest put the efficiency of crowdsourcing-platforms into question. No practitioner will invest in crowdsourcing-platforms if it requires a significant amount of time and manpower to filter the relevant, high quality information from a pile of text that is as long as this complete dissertation. Hence, the classic problem, which motivates this dissertation, is to find ways of searching for the needle in the haystack or, at best, avoiding the situation of searching for the needle at all. Taking the practitioner’s perspective, this motivation comprises two aspects:

1) How can we avoid being overflooded with data when we decide to join a crowdsourcing-platform?
2) If large amounts of unstructured data are considered a necessary evil of crowdsourcing-platforms, how can we make use of IS to structure, analyze and control that data, foremost to detect high quality, so that we can reimburse essentials to our R&D?

Those two questions are trigger and drive of this dissertation. The first question starts before an enterprise’s decision to join a crowdsourcing-platform. Enterprises need to know about the types of platforms, their benefits and requirements (c.f. Table 1). If a platform is chosen, enterprises need a basic understanding of how to set up the ideation contests, foremost which effects might be caused by adjustable design parameters of a crowdsourcing-platform (c.f. Figure 2). The second question embodies the need for a decision support system (DSS) that facilitates the management of results taken from an ideation contest. Practitioner’s wishful thinking describes a system that takes care of evaluating all submissions, detecting and selecting high quality and allocating all rewards in accordance to this quality.

This dissertation will not fulfill the requirements of this vision. Even it is a longstanding dream of the IS community to have algorithms that are capable of automatically reading and obtaining information, that are capable to understand human language, we will not have such possibilities in the near future. Coming back to the initial quotation taken from (Kasparov, 2007), many scholars think it will require a full simulation of how the mind works before we can write programs that read and understand the way people do. This dissertation intends to count for one small step on our way towards the realization of this wishful thinking. It sounds and recombines a variety of existing data analysis techniques to facilitate corporate usage of crowdsourcing-platforms.
1.2 Detailed Problem Description and Research Questions

Juxtaposing the importance of contribution quantity and quality in crowdsourcing-platforms motives this thesis’ research questions. As the quality vs. quantity discussion is by no means novel regarding online platforms in general, somewhat different perspectives have to be considered when dealing with crowdsourcing-platforms or online ideation contests. Following the rules of a network economy (Shapiro & Varian, 1997), the persistence of web-platforms in general is reliant on the amount of users actually using it. The cost of setting up a platform as well as storing and processing data has decreased dramatically in the past, and, as a result, the amount of web-platforms has grown at an explosive rate. All platforms aim for extreme economies of scale, trying to leverage direct and cross-sided network effects (Rochet & Tirole, 2003) in order to promptly reaching the required critical mass of users. Platforms which fail in achieving this sub-step will be eliminated from the market in the long run.

Another reason why it is so important for nearly every online-platform to maintain a large user base is, that users and their contributions often are included in the value proposition of a platform’s business model (Osterwalder & Pigneur, 2010). For example, in SNS, like facebook, LinkedIn or QQ, the network size itself constitutes the central element of the value proposition. Following Metcalfe’s law (Shapiro & Varian, 1997), the value of a network is proportional to the square of the number of connected users ($n^2$). However, with firms creating company profiles, organizations offering to join sub-groups and game developers offering to play games inside platforms, even Metcalfe’s law underestimates the value created by a group-forming SNS as it grows. In other words, adding up all the potential two-person groups, three-person groups, and so on that users, firms, organizations and game developers could form, the number of possible groups equals $2^n$ (and the n, to a certain degree, currently equaling 955 million for the instance of facebook). Hence, following (Reed, 2001), also the value of such SNS increase exponentially, in proportion to $2^n$. On the contrary, quality of participation is not perceived to be of high relevance. In other words, the activity of the users of a SNS is seldom used within the value proposition and neither being of fundamental interest for funders and investors. The fact that by 2012 at least 8.7 percent of facebook’s 955 million registered accounts (over ten percent of ca. 600 million accounts in the case of twitter) are fake, duplicate, nonhuman or spam accounts (Alexa, 2012; Internetworldstats, 2012) is rather worth a side note than being central issue.

Likewise, the value propositions of weblogs and wikis are solely based on quantitative criteria, habitually the amount of user generated content (posts, comments or entries). In that, weblogs and wikis augment the previous value proposition by the results of the user’s activity on the platform. Following the web 2.0 pattern of data being the next intel inside (Musser & O’Reilly, 2007) the value proposition of such platforms is
variety and content richness. For example, the success of Wikipedia is mainly reasoned by the amount of articles the platform comprehends. According to Wikipedia’s own statistics, the platform grows by three times the speed that a single person could read if he or she follows a ratio of 600 words a minute for 16 hours a day (Wikipedia, 2012). Reading the current English incarnation (which encompasses round about 4 million articles) at that rate would take over ten years, and by the time one would be done, so much would have changed with the parts they had already read that they would have to start over.

In a similar vein, quantity criteria are used within the value proposition of crowdsourcing-platforms. The tenet such platforms are following and thus, are offering as a value proposition can be coined as *the larger the fundament, the higher and more stable the top*. In other words, crowdsourcing-platforms, which offer the outsourcing of firm’s ideation processes, base their value proposition on tapping an endless variety of ideas. Often crowdsourcing-platforms supplement this value proposition by offering *the wisdom of crowds*. Following the web 2.0 pattern of *harnessing a collective intelligence* (Musser & O’Reilly, 2007), the rational is that as long as high quantity is achieved, the likelihood of high quality being directly or indirectly included is guaranteed as well. Hence, *fast access to a diversity of valuable ideas and the wisdom of crowds* can be put as *value propositions* within the business model of crowdsourcing-platforms, illustrated in Figure 3.

![Figure 3: Exemplary business model of a crowdsourcing-platform, illustrated by using the business model canvas from (Osterwalder & Pigneur, 2010).](image-url)
In comparison, all mentioned quantity based arguments are in agreement: The larger the number of users of a SNS the higher its allotted value, the larger the amount of articles in Wikipedia the higher its utility for the reader and the larger the amount of submissions towards an ideation contests the higher the chance that high quality can be directly extracted from it or aggregated out of it. In case of the latter, this rational is depicted by the circular relationship (1) between the value proposition, the customer segments and the revenue streams within figure 3. Revenues which a crowdsourcing-platform provider receives from firms are based on their ability to warrant a diversity of valuable ideas. Consequently it is a key activity of ideation contests to foster the process of ideation technically, which in turn is influenced by the crowdsourcing-platform’s know-how of design patterns and characteristic crowd behavior. This know-how therewith is stated as the key resource within the business model of an ideation contest.

A crowdsourcing-platform that is missing these key resources and hence, is unable to provide such key activities, will not produce the value proposition to the customer and will be exposed to a variety of critics (c.f. relationship (2) in Figure 3). Thus, understanding the abilities and limits of the crowdsourcing-platform is crucial for their providers. But as crowdsourcing-platforms are rather a new and also changing phenomenon such understanding is not yet carved into stone. From the just described problems the first set of research questions that is to be addressed in this dissertation can be formulated, and hence, is depicted in Table 1.

<table>
<thead>
<tr>
<th>No</th>
<th>Research Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.a</td>
<td>What types of crowdsourcing platform can be separated and what are their characteristics?</td>
</tr>
<tr>
<td>I.b</td>
<td>How are crowdsourcing-platforms related to the theory of a collective intelligence?</td>
</tr>
<tr>
<td>I.c</td>
<td>What are the critical design elements of crowdsourcing-platforms?</td>
</tr>
<tr>
<td>I.d</td>
<td>What are the effects of a crowdsourcing-platform’s design elements on the quantity and quality of ideation?</td>
</tr>
</tbody>
</table>

Table 2 Overview of the first set of research questions.

Basically, this first set of research question addresses the theoretical understanding of the crowdsourcing concept. The managerial impact is comprehended by deepening the knowledge base (seen as key resource) about the new phenomena. This includes knowledge about the differences between various types of crowdsourcing platforms, common user behavior in regards with crowdsourcing and the impact of technical design elements in platform or contest design.

The second set of research questions focuses the aforementioned issue of participation or contribution quality. As illustrated in Figure 3, the main value proposition of crowdsourcing is that it can be used to create a diversity of valuable ideas, or even a
form of collective intelligence. However, the theoretical argumentation of this proposition is mainly based on anecdotal evidence from situations that moreover might not even be related to the instance of modern online ideation contests. The most frequently cited examples that research papers recall, are mentioned in the 2004 book *The wisdom of crowds* written by James Surowiecki (Surowiecki, 2004). Surowiecki refers Francis Galton's surprise that the crowd at a county fair accurately guessed the weight of an ox when their individual guesses were averaged and moreover, the average was closer to the ox's true butchered weight than any of the separate estimates made by cattle experts. A similar incidence is describing the lifeline to “ask the audience” in the popular TV-show “Who Wants to Be a Millionaire?”.

With both examples referring to the law of large numbers, the key to success lies in aggregating large quantities of diverse and independent opinions. But the aggregation (or filtering) of a big dataset is considerably more difficult when the underlying data is not structured. Such is the case in crowdsourcing and ideation contests in particular, on which data is increasing rapidly in both volume and complexity. As a result, the analysis of ideation contests usually requires large amounts of time. As a consequence, firms are running the risk of missing the benefits of crowdsourcing. The above outlined problems motivate the second set of research questions, which is to be addressed by this dissertation and which is depicted in Table 2.

<table>
<thead>
<tr>
<th>No</th>
<th>Research Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>II.a</td>
<td>What types of text- and data-mining methodologies are relevant in terms of crowdsourcing and ideation contest?</td>
</tr>
<tr>
<td>II.b</td>
<td>How can text mining methodology be technically applied to support quality evaluation processes within crowdsourcing platforms?</td>
</tr>
<tr>
<td>II2.c</td>
<td>To what extent can quality measurement of crowdsourcing be automated by applying text- and data- mining methodologies?</td>
</tr>
<tr>
<td>II.d</td>
<td>How can the process of ideation technically be measured and analyzed by using text- and data- mining methodologies?</td>
</tr>
<tr>
<td>II.e</td>
<td>What cognitive abilities of solvers can be detected within ideation contests and how are they visible in intermediate and final results of an ideation contest?</td>
</tr>
</tbody>
</table>

*Table 3 Overview of the second set of research questions.*

Like the first set of research questions, the second set also has a theoretical and managerial impact. From the theoretical perspective this dissertation takes an explanatory approach. Rather than exploring for new theory, the thesis aims to detect the patterns of accepted ideation and crowdsourcing theory within online ideation contests. As mentioned, this acceptance, however, hitherto is based primarily on the strength of the theoretical argument, and the citation of the aforementioned commonplace examples, rather than rigorous empirical surveys.
The managerial input of the second set of research questions addresses the platforms' key activities (c.f. Figure 3). As mentioned, theory states high quantities as requirement for platform existence (network size) as well as high quality (amount of ideas). On the one hand, large quantity of unstructured data slows down the evaluation process at best, but inhibits it completely at worse. As a result, high quality that might be embedded in such unstructured data will drain away due to the current lack of IS in supporting filtering and aggregation. On the other hand methods of text- and data-mining are known as solid methods for precisely this kind of problem. Text-mining is the semi-automated process of extracting patterns (useful information and knowledge) from large amounts of unstructured data sources (such as ideas within an ideation contest). It works by transposing words and phrases in unstructured data into numerical values which can then be linked with structured data in a database and analyzed with data-mining techniques (Turban, Sharda, & Delen, 2011). To sum up the managerial input of the second set of research questions, the idea is to explore if and how methods of text- and data-mining can be applied to crowdsourcing-platforms in general, and the quality of ideation in particular.

The separation of the research questions within both sets of questions thereby reflects the necessity to address different audiences (researchers and practitioners), and also already leads to the structure of this thesis. Both aspects are explained during the following two sections.

**1.3 Contribution and Audience**

The results that are presented in this dissertation address two different audiences, namely researchers and practitioners. In case of the latter, a separation between providers of crowdsourcing-platforms and innovation seeking enterprises can be taken. Enterprises may benefit slightly more from the first set of research questions, for example when they are exploring crowdsourcing for the first time or facing the decision to choose a particular crowdsourcing-platform or setting up the parameters of an ideation contests. In return, both audiences can benefit from the second set of research questions. As mentioned, the current lack of clear methods to analyze and evaluate crowdsourcing is a major concern within the practitioner community. The methodologies and procedures that are developed and tested following the second set of research questions may help practitioners to overcome the current problem of filtering signals from noise (*finding the needles in the haystacks*) as well as developing a deeper understanding of the ideation process itself. Consequently, in all contributions the text-and data-mining approaches are documented exhaustively in order to be reproducible by the practitioner community. Additionally, all approaches are evaluated using datasets from practice, which emphasizes the importance of relevance and reliability within the practitioner community.
This dissertation is also intended for a broad audience in IS- and innovation research. Using the conference or journal classifications of crowdsourcing or ideation contests, manifold stakeholders within the research community are present. The terms are frequently associated and mentioned within the call for paper to open innovation-, web 2.0- or social media-tracks in most valued IS conferences. On the other hand the terms are increasingly included in non IS-based research fields such as Marketing, Innovation Management, Management Science or even Biology.

Additionally this dissertation has a substantial relation to the research field of information retrieval. The contribution does not refer to new information retrieval methodologies (like text- or data-mining algorithms) but rather new ways of applying and recombining those methodologies towards a new space. In that manner the dissertation again is of explanatory nature. The focus is to find out how much of the initially stated desire to automate the management of an ideation contest can be achieved by combining algorithms and methods that already exist within the field of information retrieval, rather than exploring or adding new algorithms to the field.

1.4 Thesis Structure

This dissertation is structured in Part A and Part B: Part A provides an overview of the whole dissertation, while part B consists of six research papers that address the research questions in detail. The structure is illustrated in Figure 4.

![Figure 4 Overview to the structure of this dissertation.](image)
Section 1 of Part A explains the overall motivation for conducting research on crowdsourcing, provides the detailed problem description and develops research questions. On this basis, two audience groups of this dissertation are deduced. Following the relevant conceptual foundation, the research gap within IS-research, addressed by this dissertation, is presented in section 2, while the chosen research paradigm and research methods are described in section 3. Section 4 summarizes the findings of the six individual contributions. The final section 5 summarizes findings to create the big picture of this dissertation. Research results, implications for practice and research are presented as well as an outline of opportunities for further research. The section closes with a critical reflection and personal opinion.

Part B consists of six articles that present the research results in full detail. Five articles already have been published, thereof three in proceedings of peer-reviewed international IS conferences (contributions B, C and E), one in a peer-reviewed edited book (contribution A), and one in a practitioner-oriented, but ranked scientific journal (contribution D). A final contribution (contribution F) is currently under review within a peer-reviewed, ranked scientific journal. Because contributions have been published in or are submitted to different academic outlets, articles vary in scope, language, and level of analysis. In order to achieve a consistent presentation, the formats of all contributions have been unified, foremost represented by a unified layout (fonts, font size, spacing, tables and figures). It occurred that original tables employed smaller font sizes or otherwise different formats and the redesigned tables exceeded the herein given page dimensions. In this case, tables were split and continued on the following pages. As suggested, a consolidated table of contents as well as lists of figures and tables (keeping the original content declaration) covering Part A and part B of the thesis is included to improve readability and traceability of content.

However, the corpus itself, that is to say the essential bodies of text, of the original publications have not been changed. Therewith each contribution can be read by its own, including contribution based references in the citation style required by the contribution’s output format. A table summarizing the bibliographical information (title, author(s) and affiliation(s), publication type, year, publication status, and the suggested correct full citation) was conducted and put in front of each particular contribution.

2 Research Foundation
Following the standardized procedure of a dissertation, the research foundation was elaborated during the preliminary study to the dissertation (ger. “Vorstudie”). This preliminary study was handed in by February of 2011 and the suggested research plan accepted in April 2011. A main component of the preliminary study to the dissertation
is to conduct a rigorous literature review within the research field of the dissertation. The goal of this is to demonstrate knowledge about a particular field of research (i.e. crowdsourcing in this case), including vocabulary, theories, key variables and phenomena, and its methods and history (Randolph, 2009). In best case, a literature review of academic publications is able to create clarity about the state of the art of research, whenever a certain topic spread uncontrollably due to raising popularity in practice and academia. To the best of my knowledge, such circumstances were given by the time of the preliminary research study, foremost due to the situation of new terms like crowdsourcing or ideation contests immingling with hitherto existing terms like open innovation, collective intelligence or user-centric innovation. The following section revisits the literature review and summarizes its essential findings on academic literature on crowdsourcing by the time of early 2011 the preliminary study. It’s intention is not to copy but to revisit the review from today’s perspective. Thus, it provides the research fundament of this dissertation and gives reason to the sets of research questions depicted in Table 2 and Table 3. Furthermore the literature review provides reasoning to the chosen research approach explained in section 3 and supports the critical reflection within section 5.

2.1 Revisiting the Literature Review on Crowdsourcing
A literature review should not only be a brief collection of papers, published within a certain domain. Moreover, it should uncover the sources relevant to a topic under study and, thus, make a vital contribution to the rigor and relevance of research (Vom Brocke et al., 2009). Relevance refers to the avoidance of investigating what is already known (Baker, 2000) and rigor to the effective use of the existing knowledge base (Hevner, March, Park, & Ram, 2004). To conduct a literature review which fulfills requirements of rigor and relevance, the review process follows suggestions of (Webster & Watson, 2002). Furthermore, various methodological guidelines, e.g. given by (Baker, 2000; Cooper, 1985; Levy & Ellis, 2006) are taken as advice and are referenced at their according passage. Hence, the following sections will contain the literature review process, including scoping, conceptualization of relevant topics and a detailed description of the search process. Findings from the literature search process are analyzed and synthesized with the aim to provide an overview of relevant research streams and research gaps.

2.2 Scoping the Literature Review
(Cooper, 1985) summarized the characteristics of literature reviews by providing a taxonomy. Following this, literature reviews differentiate in their focus, goal, organization, perspective, audience and degree of coverage. It is suggested to apply this taxonomy to initially set the scope, before starting the literature review itself. In
order of doing so, the following literature review intends to be of exhaustive and rigorous nature (coverage). In the center of interests are research outcomes, underlying theories and applied research methods (focus). A neutral perspective is taken, as the goal is not to criticize, but to detect key issues and integrate them to the already known knowledge base. Therefore, the organization of the literature review follows a conceptual approach (meaning the search process is organized using known concepts and keywords). As the intention is to contribute to the IS knowledge base, specialized scholars from this research field and practitioners are seen as core audience of the literature review.

2.3 Conceptualizing Relevant Search Terms
After defining the literature reviews’ scope, it is recommended to give a broad conception of what is known about the topic and to have a common understanding of basic terms within the research field (Torraco, 2005; vom Brocke et al., 2009).

As mentioned in section 1.1. and 1.2 (especially recall Table 1, Figure 2 and Figure 3), crowdsourcing is the central concept of this dissertation and therewith also the main term within the literature review. However, as the term crowdsourcing has essentially emerged from the practitioner community, hitherto predominant used terms within the IS research community also are of relevance. In other words, research papers on crowdsourcing do not constitute a yet unexplored field. In fact, a plurality of partly overlapping or constitutive theories can be identified dealing with aspects of ideation generation processes or collective intelligence. Hence, the next step determines the relation between the terms crowdsourcing and other pertinent terms.

Following the suggested literature review procedures (Baker, 2000; Torraco, 2005), concept maps should be used to illustrate those relations. Concept maps are graphical tools for organizing and representing knowledge, mainly in form of a diagram, showing the relationships among concepts (Novak & Cañas, 2008). In contrast to an entity relationship model, a concept maps does not follow a mandatory declaration. Concepts simply are represented as boxes or circles, and connected with labeled arrows in a downward-branching, hierarchical, or relational structure. Figure 5 illustrates the concept map of crowdsourcing and related terms.

The concept map shows that crowdsourcing is not a merely new concept. Crowdsourcing follows the open innovation paradigm, but leverages the use of modern IS. Usually crowdsourcing-platforms are set up online. On crowdsourcing-platforms ideation contests are used to provide the tasks to a solver crowd. The Web 2.0 pattern of “harnessing collective intelligence” is often demanded by seeking enterprises, however, neither not all solver crowds are able to form a collective intelligence, nor does the definition of crowdsourcing itself require them to do so. In other words, crowdsourcing is not attached to innovation or ideation per se, but can also refer to
simple tasks like clickwork. Crowdsourcing and open innovation attempts have the same goal, that is to say to produce user centric (or user based) innovation (also often referred to as customer co-creation). However, open innovation rather addresses a strategy or state of mind, than clear guidelines, and especially does not require the use of IS.

![Crowdsourcing concept map based on (Walter, 2012).](image)

**2.4 The Literature Search Process**

Pledging for rigorous searches, (Vom Brocke et al., 2009) suggest a four-step search process: *journal search, database search, keyword search* and *backward/forward search*. Each phase is composed of a search task and an evaluation task (Levy & Ellis, 2006) and needs to be documented in order to ensure reliability. Figure 6 illustrates this procedure schematically. Initially, a *databases* and *journal search* is necessary to define sources and scope for subsequent queries. In the next step, defined *keywords* can be queried via selected databases and keywords. Findings are stored in a paper stock, which is evaluated during each step of the literature search process. Evaluation describes the process of limiting the number of articles to those relevant to the topic. The sum of findings within the keyword search creates the basis of *forward* and *backward search*. The *backward search* is reviewing literature that is cited in articles derived from the keyword search. *Forward search* specifies the process of reviewing articles, which have been published subsequent and have cited the articles derived from the keyword search (Webster & Watson, 2002). The sum of papers, found via *keyword, forward* or *backward search* define the paper stock, which represents the basis for the following literature analysis.
In order to fulfill the relevance criteria (Webster & Watson, 2002), only peer-reviewed journals or conference proceedings are included. The database search resulted in three relevant databases, that are to be precise: EBSCOhost, The Association of Information Systems electronic Library (AISeL) and ScienceDirect. Those three databases cover all top ranked journals as well as conference-proceedings within IS research. Therewith searching in additional databases, for example Proquest, JSTOR, Wiley Inter-Science, The Social Science Research Network (SSRN) or Google Scholar would only result in duplicate publications.

<table>
<thead>
<tr>
<th>DB</th>
<th>Short description and exemplary journals and conferences</th>
</tr>
</thead>
<tbody>
<tr>
<td>AISeL</td>
<td>Library covering journals as well as conference proceedings, both sponsored or associated with the AIS, e.g. the journals: Journal of the AIS (JAIS), Communications of the AIS (CAIS), Business &amp; Information Systems Engineering (BISE) and conferences: ICIS, ECIS, PACIS, AMCIS, MCIS, HICSS, Wirtschaftsinformatik Proceedings or BLED Proceedings.</td>
</tr>
</tbody>
</table>
The identified databases (and hence, all included journals) are now queried, using distinct sets of keywords which are extracted from the conceptualization of the crowdsourcing (c.f. Figure 5). To be precise, five sets of keywords are used:

1) ideation contest, idea contest OR idea tournament
2) innovation contest, innovation market OR innovation tournament
3) crowdsourcing, crowd sourcing OR crowd-sourcing.
4) user generated innovation, user centric innovation OR customer co-creation
5) the wisdom of crowds, collective intelligence OR swarm intelligence

Queries are set to titles (T), abstracts (A) or allotted keywords (K) within the advanced search masks of corresponding databases. Queries are combined by a logical OR and set in quotation marks, meaning only the exact spelling of keywords (e.g. not crowd wisdom, but only “the wisdom of crowds”) are found.

<table>
<thead>
<tr>
<th>Keyword Set</th>
<th>1)</th>
<th>2)</th>
<th>3)</th>
<th>4)</th>
<th>5)</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>query by</td>
<td>T</td>
<td>A</td>
<td>K</td>
<td>T</td>
<td>A</td>
<td>K</td>
</tr>
<tr>
<td>Σ papers</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>unique papers</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 5 summarizes the results from this keyword search numerically, that is to say by the amounts of publications per set of keywords and query type (in T, A or K). Since neither the coverage of journals within databases nor the usage of keywords within paper titles, abstracts or allotted keywords are mutually exclusive, double counts have to be removed manually in order to get the net hits (unique papers). Clearing such duplications from the findings reduces 154 absolute hits to 86 unique publications (39 by title, 35 by abstract and 12 by keyword query).
In accordance with the steps of literature review, backward and forward search have to be undertaken as final steps within the search procedure. Following suggestions of (Torraco, 2005; vom Brocke et al., 2009; Webster & Watson, 2002) to proceed with the forward search, manual reading through abstracts of publications which cited one of the 86 papers detected in the keyword search led to five additional findings. In contrast, the backward search is conducted purely schematic, explicitly copying all references, given within the 86 unique papers from the keyword search. In total, 3865 references were cited within the 86 papers from the keyword search. Deleting duplicates leads towards 2882 unique references (detected via the backward search), which states that only approximately one fourth of all references were cited at least twice, whereas three fourth of all citations are unique. Also only 16 of the 86 initial references are inside references, meaning that they are found in backward search, but also already have been found via keyword search. In order finalize the literature search and proceed with the literature analysis a clear cut among our total paper stock has to be taken, as the total of nearly 3000 references states far too much to analyze. Therefore, the 39 most frequently cited references, detected by the backward search are selected and hence, complete the final literature stock. The amount of 39 publications represented the cut of at least 4 out of the 86 possible citations. Table 6 summarizes the backward search schematically.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Surowiecki (2004)</td>
<td>1</td>
<td></td>
<td>25</td>
</tr>
<tr>
<td>Von Hippel (2005)</td>
<td>1</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>Howe (2006)</td>
<td>1</td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>O’Reilly (2005)</td>
<td>1</td>
<td></td>
<td>12</td>
</tr>
<tr>
<td>Howe (2008)</td>
<td>1</td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>Tapscott &amp; Williams (2008)</td>
<td>1</td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>Chesbrough (2003)</td>
<td>1</td>
<td></td>
<td>8</td>
</tr>
</tbody>
</table>
Table 6 Results from the backward search.

To sum it up, a total of 125 publications are detected by the sequence of database-, journal-, keyword-, forward- and backward-search. It is the purpose of the following literature analysis to constructively concentrate the information detected in this pile of 125 publications and describe what patterns can be found (Webster & Watson, 2002).

2.5 Statistical Literature Stock Analysis
The literature search led towards 125 relevant publications (journal papers, books and papers from conference proceedings). Following, the dataset will be described statistically as a basis to define literature streams and research gaps respectively. As forward and backward search already suggested, amounts of publications are accelerating since the introduction of crowdsourcing in 2006. Likewise, conference proceedings represent the strongest growing publication type. The increasing amounts show, that the queried search terms are no longer practitioner’s buzzwords, but increasingly are also used by scientific research. These trends are presented within Figure 7.

Output formats are distinguished further by analyzing what set of keywords have lead towards the hits. The terms included in keyword set 1) and 2) were not yet established by early 2011 within the domain of IS research. In contrast set 3) and 5) are the dominantly used terms in IS research. The tendency towards more conference proceedings for keyword sets 3) follows a general tendency towards more conference proceedings and the dominant presence of these categories. In other words, by early 2011 IS research emphasized the new phenomena of crowdsourcing over previously existing terms. This led to a plurality of conference proceedings. This situation is illustrated in Figure 8.
In order to identify possible research gaps, the research methodologies applied within the 125 different publications are compared against the addressed research objectives. Recalling Figure 5, seekers, solvers and the intermediary crowdsourcing-platform are representative research objectives. Figure 9 illustrates this comparison.

Solvers are the dominant research objective. Within case studies or even without clearly announced research method, solver’s motivation to participate, their demographics and their produced results are described and analyzed. A second dominant combination of research method and research objective is given by design science approaches towards crowdsourcing-platforms. In this instance, the intention of publications is to deduce design patterns or design elements of platforms in order to foster the ideation process.
Such publications usually start by collecting all possible characteristics a crowdsourcing-platform can choose within certain steps. For example, the rating mechanism for ideas can be crowd-based, seeker-based or expert-based, and can be technically supported by a Likert-scale, a like button, a qualitative measurement, a Boolean selection and so forth. Such characteristics are combined in taxonomies or morphological boxes to provide artifacts to support crowdsourcing-platform design. Usually a definite instance is measured to evaluate the approach.

### 2.6 Literature Streams

Aiming to find all research gaps within a certain field is a helpless venture, because implicit that would mean to already know all potential research streams. Hence, it is suggested to finalize the literature analysis by defining obvious research streams and comparing them to the considered own research approaches to discuss whether the own research intention would fill a gap (Baker, 2000; Levy & Ellis, 2006; Randolph, 2009). Based on the literature analysis, five divergent combinations of research method and research objective lead to five clear research streams. Table 7 summarizes these literature streams and recalls representative publications.

The research streams are scaled from rather being descriptive and qualitative (stream 1) towards being rather empirical or quantitative (stream 5). Qualitative research is represented by streams 1, 2 and 3, whereas streams 4 and especially 5 represent quantitative approaches.

![Figure 9 Publications by research method and research objective.](image-url)
<table>
<thead>
<tr>
<th>Research Stream</th>
<th>Representative References</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Explaining or exploring crowdsourcing as a new phenomenon or concept by descriptive examples or case study research.</td>
<td>(Adams &amp; Ramos, 2010; Brabham, 2008, 2009, 2010; Corney et al., 2010; Dearstyn, 2010; Nam, 2010; Shang &amp; Chen, 2010; Walter &amp; Back, 2010)</td>
</tr>
<tr>
<td>2 Define and analyze design patterns of crowdsourcing-platforms.</td>
<td>(Bullinger, Neyer, Rass, &amp; Moeslein, 2010; K.-T. Chen, Chang, Wu, Chang, &amp; Lei, 2010; Ebner, Leimeister, &amp; Krcmar, 2009; Lee &amp; Chang, 2010; Leimeister et al., 2009; Müller, Thorig, &amp; Oostinga, 2010; Rouse, 2010; Wagner &amp; Back, 2008)</td>
</tr>
<tr>
<td>3 Create methods, models or frameworks (applying design science methodology).</td>
<td>(Gans &amp; Stern, 2003; Lin &amp; Davis, 2010; Matthing, Sandén, &amp; Edvardsson, 2004; Morgan &amp; Wang, 2010; Schut, 2010; Yassin, 2009)</td>
</tr>
<tr>
<td>4 Surveying the solver crowd and using descriptive statistics to define trends and hints towards success patterns.</td>
<td>(Baron &amp; Warnaby, 2010; Lakhani &amp; Panetta, 2007; Lakhani &amp; von Hippel, 2003; Luthje, Herstett, von Hippel, &amp; Vonhippel, 2005; Ogawa &amp; Piller, 2006; Piller &amp; Walcher, 2006; Riedl, Blohm, Leimeister, &amp; Krcmar, 2010; Wolfers &amp; Zitzewitz, 2004)</td>
</tr>
</tbody>
</table>

Table 7 Research streams detected within publications from the literature review.

### 2.7 Defined Research Gaps

On first sight, there is no clear gap between these streams. Although to different extents, qualitative as well as quantitative approaches are already made towards the all research objectives. Success factors, design patters, incentive schemes, models or frameworks are analyzed or developed. As a result, potential research gaps rather emerge from the researcher’s ambitions and applied research methods. In characteristic fashion, the majority of publications (71 out of 125) claims to be exploratory in nature, that is theory grounding. In addition, only 9 publications proclaim themselves as being explanatory, which is proofing or falsifying an existent
theory. Similar ratios are given within the category of a target audience, where a majority addresses specialized scholars only. However, an existent criticism of IS research is, that it is only focusing on the scientific community and neglecting practice (Gill & Bhattacherjee, 2009). Only 23 of 125 publications directly address (or include) a managerial audience. In accordance, a vast majority of papers focuses on highlighting positive aspects of crowdsourcing, foremost case studies on successful implementations of platforms.

As a result the existing research streams (c.f. Table 7) are unsuccessful in answering the motivational questions from practice (stated in section 1.1). Recalling Figure 9, hence, the question is which research methods are missing in order to address the motivational questions. By the time of 2011, obvious structural research approaches like text- or data-mining (Feldmann & Sanger, 2007; Miller, 2005; Tan, Steinbach, & Kumar, 2006) from the field of information retrieval, as well as social network analysis (Scott, 2000; Wassermann & Faust, 1994) were completely missing.

3 Research Framework
The research framework aims to answer the question how this dissertation intends to address the research questions exposed in Table 2 and Table 3. Therefore, initially the understanding of an explanatory research approach is introduced. Afterwards, the idea of a multi-method research design is explained and finally all applied research methods are briefly introduced and juxtaposed to the targeted research questions.

3.1 An Explanatory Research Approach
As mentioned, overall, this dissertation follows an explanatory research approach. As the literature review has shown, such perspective is not very common amongst researchers, especially whenever a topic is still in the phase of rising popularity (c.f. Figure 1). Most researchers in the field of crowdsourcing claim their research to be of exploratory, that is to say of fundamental, nature. This means their belief is that the addressed problem itself has not yet been clearly defined. Hence, exploratory research aims to determine the best research questions, appropriate data collection methods and selection of subjects.

Contrariwise, this dissertation takes the standpoint that these issues are given in the case of crowdsourcing-platforms. Almost seven years after the concept has been introduced, it is my belief that certain problems (c.f. Tables 2 and 3) are well defined, and the data as well as the research subjects are well-rehearsed. Crowdsourcing-platforms have reached a certain maturity. They have overcome the exploration phase, deal with recurrent seekers and solvers and increasingly are questioned about the
quality of their results. In return, they have established clear structures, rules and
guidelines, but by some means still lack in evaluation processes. Such situation is a
typical instance of what Gartner Research describes as the slope of enlightenment
leading towards the plateau of productivity within their well-known hype cycle for
emerging technologies (Landry, 2010).

Hence, the research approach of this dissertation is to actually address problems (and
not to introduce yet another set) in order to leverage crowdsourcing-platforms to the
plateau of productivity. Therefore, an explanatory perspective is taken. Per definition,
explanatory research intends to explain things in detail and not just report them.
Hence, explanatory research draws on existing theory and enriches the reasons behind
them rather than developing new theories (Miles & Huberman, 1994; Yin, 2003). If
there are several explanations for a particular phenomenon, such as for crowdsourcing,
explanatory research should determine which one provides the best answer. If theories
already have been developed, explanatory research focuses on testing their predictions
or principles. If the results consistently agree with a theory, it is assumed to be valid. If
research fails to reproduce the same results as the original theory, in all likelihood the
theory is false and a return to the scientific method will need to be undertaken to find a
better explanation as to the phenomenon (Creswell, 2008; Hair, Black, Babin, &
Anderson, 2006). However, it is not the main goal of explanatory research to come up
with such particular explanations.

As a result, explanatory perspectives might not be considered most fashionable. They
frequently are countered by the allegation of only evaluating, but not really innovating.
However, reasons for taking such perspectives are backed up philosophically, using
Popper’s epistemological philosophy of critical rationalism (Popper, 1959). Popper
argues that scientific theories and any other claims to knowledge can and should be
rationally criticized, and can and should be exposed to tests which may falsify them.
Thus, the strong focus lies in finding empirical evidence to theories, which are often
based on stringent logical argumentation and anecdotal evidence, such as in the case of
crowdsourcing or collective intelligence. Poppers’s critical rationalism further states
that knowledge should be contrastingly and normatively evaluated. Claims are either
falsifiable and thus empirical, or not falsifiable and thus non-empirical.

Condensed, the goal of this dissertation is not to explore the phenomena of
crowdsourcing even deeper in order to further develop underlying theories, but rather
to test the existing theories and assumptions empirically. Comparably, it is also not the
goal to explore new methods, especially designed to analyze crowdsourcing-platforms.
Contrary, the simple goal is to find empirical evidence for existing theories and the
adequacy of yet existing methods. Therefore, I use methods which are well established
within IS research (case study research, structural equation modeling, text mining and
social network analysis) to analyze a research objective (crowdsourcing-platforms) for
which theory and tenets already exist. Taking such research perspective imposes the
researcher to take a rather critical view on a certain topic. This is also true for this dissertation. As a consequence, I will discuss this issue by giving my personal opinion in section 5 of Part A.

3.2 A Multi-Method Research Design
Generally speaking, a multi-method research design describes the researcher’s decision to combine a set of different research methodologies to the same research objective in order to produce reliable and hence, valuable results. This means that in order to achieve a particular research objective, such as sampling, data collection, or data analysis, multiple strategies are used to enhance the validity. Such approach is routinely advocated by literature on research strategies (Creswell, 2008; Hevner et al., 2004; Miles & Huberman, 1994; Yin, 2003). In all research guidelines, suggested within those publications, mixing, enhancing or integrating various methods within a research design (qualitative and/or quantitative) is considered a common feature of good research.

The foremost mentioned reason in advocating the use of multiple methods is to avoid a method biased, that is to say narrow, view of a particular research objective (Creswell, 2008). By using a single method multiple times, researchers might develop such method biased perspective on the addressed field of research. As a consequence, researching a subject from different perspectives or paradigms may help to gain a holistic perspective and using more than one method should help to get a clearer picture of the subject and create more adequate explanations.

The research design of this dissertation is based on using multiple methods. Contribution A and B provide the fundamental understanding of crowdsourcing in regards to this dissertation by using qualitative research approaches, foremost case study research. Contribution C uses a quantitative approach, that is to say structural equation modeling, to test the effects of external factors on the outcomes of crowdsourcing. Contrary, contributions D, E and F separately combine multiple methodologies within the publication. All three contributions make use of text- or data-mining methods in order to structure and analyze data and develop procedure models to analyze and evaluate the quality of crowdsourcing. Moreover, within contributions D, E and F quantitative methods are used subsequently to evaluate the validity of suggested the procedure models.

However, in this dissertation, methods are not only given artifacts that are used as means to an end. As the literature review shows, one central research gap, which is addressed by this dissertation, is the lack of using certain methods, especially text- or data-mining within research on crowdsourcing-platforms. Hence, contributions D, E and F reintroduce those methods, reassemble them and adjust them to the research
objective of crowdsourcing-platforms. Consequently, mainly applied research methods are briefly described in the following.

3.2.1 Case Study Research
Case study research can help to understand new or complex phenomena, whenever an isolated observation is insufficient and it requires a holistic approach that includes the context in which the phenomena occurs (Benbasat, Goldstein, & Mead, 1987; Yin, 2003). Usually the barriers between the phenomenon and its context are blurry within case study research. Such situation was given at the beginning of this dissertation. Crowdsourcing was still an emerging phenomenon of unclear value and with a blurred borderline to web 2.0-platforms in general.

Hence, within contribution B, we attempt convergence of the crowdsourcing phenomenon by using case study research methods, foremost following (Yin, 2003) and (Eisenhardt, 1989). We distinguish case study research as adequate, because it’s major constituting characteristic of analyzing how and why questions. Analyzing authors who partly crowdsourced the composing of textbooks enters the research field and shapes further questions of research as it is suggested in (Eisenhardt, 1989). Following the suggestion of (Yin, 2003), we make use of a theory (business model theory) as a constitutional perspective and develop a framework in order to establish comparability during the following cross case synthesis. (Yin, 2003) calls this analytic strategy developing a case description.

3.2.2 Structural Equation Modeling
Structural equation modeling (SEM) summarizes various statistical techniques for testing and estimating causal relations between latent constructs (Armstrong, 2001). SEM probably is the most frequently used method within recent publications from the field of social sciences. SEM allows both, explanatory as well as exploratory modeling, meaning they are suited to both theory testing and theory development (Hair et al., 2006).

Within contribution C we follow a traditional SEM approach. We start the explanatory modeling of effects by deriving hypothesis from existing literature. These hypotheses are combined in a causal model, which is operationalized using Ordinary Least Squares (OLS) to determine size and strength of relationships between dependent and independent variables. OLS is commonly used approach to operationalize SEM and is also used frequently in studies on crowdsourcing. We test two models against obtained measurement data taken from a crowdsourcing-platform to determine how well the model fits the data. As we find significant evidence for some rather obvious hypothesis, the important result (in regards to the superordinate explanatory research approach) is represented in hypothesis for which we were not able to find significant proof. In contribution C this is given by a missing causal relation between external factors on the quality of crowdsourcing. In other words, the failure to reproduce the
same results as suggested by the literature provides evidence that in all likelihood the theory and basic assumptions on crowdsourcing quality might need reconsideration.

3.2.3 Text- and Data-Mining

Applied in contributions D, E and F, text- and data-mining methods represent the most important method within this dissertation. As mentioned, the reason to apply those methods results from the literature review, which proofed text- and data-mining to be underused methods within the research field of crowdsourcing. Additional, but rather brief literature reviews on the application of those methods, conducted in contributions E and F endorse this situation. Hence, the question rather is, why research has not adopted such well-rehearsed methods, even their potential value proposition to the field seems obvious.

Text mining is the semi-automated process of extracting patterns (useful information and knowledge) from large amounts of unstructured data sources (Turban et al., 2011). Text mining works by transposing words and phrases in unstructured data, such as submissions to crowdsourcing-platforms (Ideas in ideation contests), into numerical values which can then be linked with structured data in a database and analyzed with data mining techniques (Feldmann & Sanger, 2007). In other words, the initial position is not about mining for data but rather drowning in data, and the idea of text mining is to structure those data to detect hidden patterns, hits relevant to a query and highly valuable information (often referred to as signals). Hence, text mining can be described as a combined process or information retrieval + information extraction + information analysis (Feldmann & Sanger, 2007).

Even though the problem of providing natural language access to textual material and responsive techniques of text- and data-mining were already discussed back in the 1980s (Hobbs, Walker, & Amsler, 1982), yet, text-mining is by no means an automated process. Today, a large variety of mining techniques such as language recognition, sentiment detection, term frequency analysis, categorization, clustering, word volatility and many more are included in numerous text mining software packages and hence, can be applied to nearly any corpus of text. However, in order to be able do so, to this very day still the cumbersome, since often manual and thus time-consuming, steps of pre-processing have to be conducted. Such include extracting the raw text (the so called corpus) from its origin and importing it into software for text analysis, structuring the corpus into a tabular format, stemming words, clearing the corpus from so called stop-words, bonding the corpus to given library of words or calculating the Term Document Matrix (TDM). Eventually, despite the common enormous efforts researcher put into text mining, results often lack significance. Reasons for this are manifold, but the core lies in the complexity of the human language. As stated in (Fuller, Biros, & Dursun, 2008), many researchers think it will
require a full simulation of how the mind works before we can write programs that read and understand the way people do.

As mentioned, three contributions of this dissertation (D, E and F), apply text- and data-mining methods. These contributions also represent my personal harmonization with the methods itself. Contribution D summarizes and enhances the results of a student project that I coached. Getting in touch with the method I sampled various text-mining methods, such as language recognition, sentiment detection and categorization, which the student tried to put together into a DSS for an airline. Contribution E is built upon these experiences, foremost that manual reconstruction of text mining methods is a disproportional effort. Hence, we used professional software, that is to say the QDA Miner by Provalis Research, to support text-mining steps. We clustered over 40,000 ideas within over 100 ideation contests on a crowdsourcing-platform in order to semi automatically define most innovative solutions. The results led to contribution F. Again, a clustering of ideas within an ideation contests was conducted. However, this time the purpose was to reconstruct the ideation function, which can be described as the process of knowledge aggregation within an ideation contest. Therefore, we selected a prototypical ideation contest containing 725 ideas. In order to apply clustering over time, all ideas had to be manually recoded by a category (pre-processing) and the clustering itself had to be executed 725 times.

3.2.4 Social Network Analysis (SNA)

Social network analysis (SNA) is a process of quantitative and qualitative analysis of a social network. SNA measures and maps the flow of relationships and relationship changes between knowledge-possessing entities. Simple and complex entities include websites, computers, people, their online accounts groups or organizations (Carrington, Scott, & Wassermann, 2005; Scott, 2000; Wassermann & Faust, 1994). Recently often used to explain the diffusion of innovations (Rogers, 2003), SNA itself origins in social sciences from the 1970s. Probably most renowned articles are Granovetter’s theory about the strength of weak ties (Granovetter, 1973) and Burt’s thoughts on cohesion versus structural equivalence (Burt, 1987).

Similar to text- and data-mining, SNA offers a large variety of network-specific metrics. Two perspectives dominate SNA: the socio-centered and ego-centered perspective. The socio-centered perspective means to analyze an entire social network, including all edges between a defined network of nodes. In such case SNA offers metrics like the networks density, the centrality of nodes, the in-betweenness of nodes, so called structural holes, the calculation of shortest paths inside the network and many more. The goal is to detect patterns of ties that indicate cohesive social groups, central actors that may be paramount to the integration of the social network, and asymmetries that may reflect social prestige or social stratification (Scott, 2000). When analyzing an ego-centered network, which means to include all edges in relation to a centered
node, SNA offers in- and out-degrees, prestige, ranking and local structures. The question is whether nodes influence one another through their edges and/or whether nodes adjust their edges to the characteristics of similar nodes (Carrington et al., 2005).

We make use of SNA in contribution F of this dissertation. In an socio-centered perspective we analyze the changes of an ideation concept network within an ideation contest over time. We also apply a ego-centered perspective to analyze the development of a single concept over time. Furthermore we take a best versus the rest comparison between rewarded and unrewarded ideas using the betweenness-centrality of nodes to detect their likelihood of being good boundary spanners, that is to say connecting or bridging two separately mentioned ideas.

### 3.3 Development of the Research Framework

Reintroducing the research questions and juxtaposing them with the recently mentioned research methods enables the development of this thesis' research framework. The framework builds upon the thesis structure (c.f. Figure 4) and outlines the courses of action as well the chosen approaches within all contributions of this dissertation. Hence, Table 8 shows the mainly addressed research questions (RQ), applied methods, datasets and evaluation techniques for all contributions.

<table>
<thead>
<tr>
<th>Paper</th>
<th>RQ</th>
<th>Method</th>
<th>Dataset</th>
<th>Evaluation</th>
</tr>
</thead>
</table>
| A     | I.a, I.b  
and I.c | State of practice analysis (desk research)  
(Desk research) | Ca. 50 different crowdsourcing-platforms | -                                                                         |
| B     | I.b and  
I.c | Case study research, developing a case study framework  
(Case study framework) | 4 case studies | -                                                                         |
| C     | I.c and  
I.d | Structural equation modeling (ordinary least squares)  
(Structural equation modeling (ordinary least squares)) | 84 different ideation contest from a crowdsourcing-platform  
(84 different ideation contest from a crowdsourcing-platform) | Fleiss’ kappa to define inter-rater reliability, significance values of constructs and $R^2$ of models |
| D     | II.b  
and II.c | Text mining (language recognition, categorization, sentiment analysis).  
(Text mining (language recognition, categorization, sentiment analysis)) | 1269 in-app-feedbacks (short text messages) from an airline’s app  
(1269 in-app-feedbacks (short text messages) from an airline’s app) | Cohen’s kappa to compare manual and semi-automatic detection |
| E     | II.a,  
II.b  
and II.c | Text Mining (clustering).  
(Text Mining (clustering)) | 112 different ideation contests containing 42’448 ideas  
(112 different ideation contests containing 42’448 ideas) | Precision, Recall and F1 scores |
Table 8 Research framework.

| F | II.c, II.d, and II.e | Data mining (concept categorization, clustering, MDS) and SNA (density, in-betweenness centrality) | Exemplary ideation contests containing 725 ideas, coded by 75 categories | Krippendorff’s alphas to compare coding validity, Mann Whitney U-Test to compare network characteristics of rewarded and unrewarded ideas |

As mentioned, evaluation plays an important part within the explanatory, multi-method research approach. Often suggested by methodologists, the evaluation is a crucial component of every research process. The common pledge is that the utility, quality, and efficacy of any IT (or design) artifact must be rigorously demonstrated via well-executed evaluation methods (Hevner et al., 2004). The evaluation then provides feedback information and a better understanding of the problem in order to improve both, the quality of the product and the process. The idea is, to provide other researchers with the possibility to continue the build-and-evaluate loop that has been started within this dissertation. Hence, especially in contributions D, E and F, results that are produced by applying text or data-mining technology and SNA, are evaluated using statistics.

4 Research Results
At its core, this thesis consists of six individual research publications (contributions A – F) that have been published or submitted to academic outlets and that are presented in full length in part B of this document. All findings presented in this dissertation result from research that has been conducted over the past three years. Except Contribution A, which has been put in front as an introduction to crowdsourcing-platforms, the order of contributions B to F does also reflect the timeline of their actual publication. Therewith, the dissertation can be read and understood as from the general to the specific. The following section 4.1 summarizes the findings to the first set of research questions (I.a – I.d, c.f. Table 2) and section 4.2 summarizes findings to the second set of research questions (II.a – II.e, c.f. Table 3). Both sections can be understood as a shortcut, an overview and summary of the results, without an in-depth argumentation.

4.1 Critical Success Factors of Crowdsourcing-Platforms
The work, which led towards contributions A, B and C, started in 2009. During that time, to many enterprises the value of crowdsourcing-platforms was not yet clear. Next to the issues of reliability and intellectual property rights, enterprises had to face a more and more cluttered field of crowdsourcing-platforms. Besides intermediaries like
InnoCentive some enterprises started crowdsourcing campaigns on their own (e.g. like Dell, IBM or Starbucks) and new kinds of platforms for design, brainstorming or clickwork (like Amazon’s Mechanical Turk Platform) emerged. Hence, the motivation of the first three contributions is to straighten out what kind of crowdsourcing-platforms can be used for which purposes, what enterprises can expect when addressing crowds with an open call and how ideation contests have to be structured in order to receive valuable results. Hence, one might see this section as the 101 to crowdsourcing-platform usage.

4.1.1 Types and Characteristics of Crowdsourcing-Platforms

Contribution A shows, that as the term crowdsourcing was being hyped (c.f. Figure 1) over recent years, also crowdsourcing-platforms sprawled in the web. Today, a variety of instances can be found. Contribution A suggest to differentiate by at least four different kinds of platforms:

- **Crowdsourcing intermediaries** that broker solution seeking enterprises and solution offering crowds (like InnoCentive or Atizo).
- Enterprises that are running direct crowdsourcing campaigns from the company website (like Starbucks or Dell).
- **C2C-platforms**, where crowds sell and buy content that has been produced by the crowd (like threadless or Spreadshirt).
- Enterprises using the crowdsourcing technology to run internal crowdsourcing campaigns (such as documented in the cases of Intel or the Commerzbank).

As contributions A and B show, technically, most crowdsourcing-platforms are based on standardized wiki- or SNS-software. Contribution A further shows, that additional differentiation has to be made within each group of platforms, e.g. crowdsourcing intermediaries can further be distinguished by their required depth of elaboration (from brainstorming towards drafts, concepts until in-depth product or service solutions), their addressed solver skill level (e.g. brainstorming for a marketing campaign vs. the solving R&D problems of a pharmaceutical company), their reward or process structure (e.g. effort based payment vs. ideation contests) and so forth. Solving this cluttered market situation, contribution A develops a decision tree, which intends to support enterprises in choosing the most convenient crowdsourcing-platform.

To sum it up, as by 2013, manifold crowdsourcing-platforms exist. Crowdsourcing should not be underestimated and definitely not be perceived as bibelot or yet another web 2.0 tool. As contribution F later reintroduces, some serious and highly valued inventions and innovative products already resulted from crowdsourcing. In regards to this dissertation, mainly crowdsourcing intermediaries are of interest. To be precise, contributions C, E, and F deal with data from a crowdsourcing-platform that offers ideation contests from a large variety of branches and segments (from engineering to
marketing). It is my personal opinion that the profile of this particular platform is prestigious to draw conclusions for a broad range of crowdsourcing-platforms.

### 4.1.2 Crowdsourcing-Platforms and Collective Intelligence

Contribution A shows, that it depends on the type of crowdsourcing-platform whether the requirements of a collective intelligence actually can be fulfilled. For instance, platforms which solely offer ideation contests on new marketing concepts are certainly far from producing a collective intelligence. In such instances, neither the crowd is organized in a decentralized structure (but in a social network), nor are suggestions hidden (but visible and, following the facebook example, likable). Furthermore, the question exist, whether tasks from fields such as marketing, that is to say tasks where the attributed values mainly are due to subjective measures, are suitable to the concept of a collective intelligence.

As also advocated in contributions A, C, E and F, because of the defining characteristic of the collective intelligence being superior to the knowledge of a single expert, it would require a plurality of experiments of a the crowd vs. experts nature to truly proof the existence of a collective intelligence. However, to the best of my knowledge such proof does not exist. In contrast, the SWOT-analysis within contribution A shows that today the claim of tapping the wisdom of crowds is mainly used for marketing purposes. Yet, no crowdsourcing-platform has actually succeeded in verifying their crowd to elaborate a collective intelligence. However, the SWOT analysis further shows that, by proper use of platform design elements, technically crowdsourcing-platforms would be able to fulfill some requirements of a collective intelligence, specifically decentralized users, mainly working independently from each other and producing manifold opinions or suggestions on a particular question. However, especially in terms of ideation contests no aggregation mechanism has been found yet, which enables to proof whether a form of collective intelligence is actually hidden within these suggestions.

Contribution B shows one certain strength of crowds. Crowds run by the law of large numbers. That means they are fast and often produce a large variety of content that even experts cannot come up with in such short time periods. This situation certainly being a problem in terms of evaluation (c.f. contribution E), it likewise embeds the clear benefit of crowdsourcing. The case studies of four partly crowdsourced management books illustrate that the crowd was able to submit sophisticated use cases, ideas to the books table of contents as well as industrious and repetitive tasks such as grammar and spell checks within days. Furthermore, the cost level underpriced professionals, editors or correctors by far.

To sum it up, looking at the results from this dissertation, it might be a good idea not to mix up crowdsourcing and the wisdom of crowds as in many ways they not necessarily are aligned with each other. Hence, the issue of collective intelligence was
not taken into further consideration. Even when discussing the quality aspects of crowdsourcing-platforms within contributions C to F, the idea of a collective intelligence is not taken as a requirement or measurement benchmark. It is my personal opinion that the existence of a collective intelligence is nearly unable to proof and hence, would only blur the results of further contributions.

4.1.3 Critical Design Elements of Crowdsourcing-Platforms

Contributions B and C introduce the importance of monetary incentives in crowdsourcing. Admittedly rarely, but existent, contribution B presents a case where crowds, not seeking firms, were even willing to pay for participation in crowdsourcing. This instance being from the early days of crowdsourcing, the situation surely has changed recently. Contribution C shows that higher monetary incentives basically trigger crowd participation. This result is in accordance with basic economic tenets. As contributions C, E and F deal with ideation contest, that is to say tournaments, in terms of monetary incentives also the reward structure plays an important role. Knowing that the distribution of rewards does not follow a winner takes it all form has a significant impact on crowd behavior. To be precise, knowing that a plurality of ideas will be rewarded has a positive effect on the amount of submitted ideas per solver. Contribution C further shows that the amount of received ideas can be influenced by seekers or crowdsourcing-platforms via adjustable design elements. The results suggest that next to higher rewards, longer duration, suggesting a distinctive type of answer as well as the seekers brand-strength and the crowdsourcing-platform’s maturity cause positive effects on the amount of ideas that are submitted in ideation contests.

From the technical perspective, further design elements can be distinguished. Contribution B shows, that if the crowdsourcing-task requires a collaborative production of content wiki-software is chosen, whereas ideation contests technically are rather based on SNS-software. Hence, in case of the latter, seekers or crowdsourcing organizers have to be aware of design elements such as the feedback mechanism, the connections amongst solvers, the comment feature as well as the rating mechanism of ideas (c.f. Figure 2).

4.1.4 Quantity and Quality on Crowdsourcing-Platforms

Considering all possible design elements, the rating mechanism, that is to say the way of measuring ideation quality per idea, within ideation contests maybe is the current key issue of crowdsourcing practice and research. A variety of studies have proven that common design elements which are used to support rating ideas, such as crowd based user feedback, 5-point Likert-scales or like buttons are insufficient in determining high quality. On the contrary, most valued are expert decisions, where a seeker allocates a group of experts from a particular field that go through all ideas and allot high quality ideas with rewards.
This issue is a central result of contribution C. It was not possible to detect significant effects on the quality of ideas, whenever quality was measured by common platform design elements. But like rating mechanisms would not work, neither did pure reliance on expert committees. The reason for this was that the evaluation work of expert committees became more complicated with a rising maturity of platforms. Their task changed from selecting five good ideas among 100 towards selecting key information from 1000 submissions. To visualize such amount, the plain text of 1000 ideas is of similar length than this entire dissertation. As the effects on quantity are determined (c.f. section 4.1.3), the quality issue required a different perspective. The following enumeration recaps the reasons for that:

- Only few significant effects on ideation quality could be measured statistically (e.g. in contribution C).
- Given platform design elements are insufficient to support quality measurement (e.g. also see (Riedl et al., 2010)).
- Expert committees are unwilling to go through piles of text manually (e.g. described in contribution E, or in (L. Chen & Liu, 2012; Poetz & Schreier, 2012; Snow, O’Connor, Jurafsky, & Ng, 2008))
- Expert committees are unable to cope with the task of manually differentiating quality within large piles of text data (see the previous bullet).

Hence, contribution C marks the starting point and foundation of contributions D, E and F. Before we can actually detect drivers of crowdsourcing quality, we have to come up with more sophisticated, that is to say not purely manual, expert based methods of determining the quality of ideation and ideas. Hence, a combined result of contributions A, B and C is that text mining methods, by definition, might be sound approach to the given problem. This leads over to the results of contributions D, E and F.

4.2 Applying Text- and Data Mining and SNA to detect Quality on Crowdsourcing-Platforms

Two driving factors led towards the application of text- and data-mining methods in contributions D, E and F.

- First, per definition, text- and data-mining are comprehensive approaches to the given problem of detecting hidden patterns in large quantity textual data such as the results from an ideation contest (c.f. section 3.2.3).
- Second, text- and data-mining represented under-researched methods within the literature review (c.f. section 2). To this day, at most a few researchers applied text- and data-mining to analyze ideation contests.
One explanation for the latter was found out during the work on contributions D, E and F, namely the fact that applying text- and data-mining methods requires countless hours of diligent but routine pieces of work. Contribution D initially explores text-mining methods and explores a variety of techniques. Contribution E takes these findings to the application of text-mining on ideation contests and contribution F provides in-depth usage of data-mining techniques when providing a repeated measure design to analyze ideation functions within ideation contests.

4.2.1 Relevant Text- and Data-Mining Techniques

Throughout contributions D, E and F a large variety of text-and data-mining techniques are applied, specifically stop word clearance, tokenization, stemming, n-grams, lemmatization, language recognition, word frequencies, Term Document Matrices (TDM) and Concept Document Matrices (CDM) in pre-processing, sentiment analysis, 2nd order clustering using k-means algorithm, Cosine or Ochia’s coefficients and the elbow-criterion and categorization as mining techniques and multidimensional scaling, network maps and dendrograms in visualization. However, at this point I cannot claim that this is an exhaustive list of all relevant text- and data-mining techniques, helpful to analyze ideation contests.

Contribution C introduces the importance of pre-processing. After crawling and extracting information from crowdsourcing-platforms, raw data usually comes in numerous, rather short, text fragments, presented in tabular format. In order to subsequently run text-mining techniques, contributions D, E and F suggest sound pre-processing as essential. The raw text data has to be cleared from information cluttering stop words, other words stemmed to their root and often lemmatization can help to group together the different inflected forms of a word so they can be analyzed as a single item. If ideas can be submitted in different languages, language recognition, based on tokenization is required as well. However, contribution D shows, that the shorter the text passages, the less reliant such techniques work. Automatic sentiment detection is also proven as rather prone to errors during contribution D and hereafter left out from application. TDM and CDM can be used in each form, binary, by frequency or inverse frequency. Contributions D, E and F suggest that the form depends on the researcher’s intention or research question. E.g. in contribution E the inverse form produces most valuable results in detecting novel ideas.

The same conclusion is drawn for the application of text- mining and subsequent visualization techniques. During all contributions, simple word or concept frequency lists provide a fast and good overview to the dataset. In terms of clustering or categorization, contributions E and F show that chosen techniques have to be in accordance with the characteristics of the dataset. For instance, in contribution E the distance measurement between ideas is based on Cosine coefficients as the underlying data is based on word frequencies. Contrary, in contribution F the coding of ideas by
concepts results in a binary CDM, that is to say a concept only can or cannot be coded to an idea. Hence, Ochiai’s coefficient, the binary form of the cosine coefficient, is used to measure distances. In terms of a final visualization of text- and data mining, contribution D shows simple frequency diagrams, whereas contributions E and F use more elaborate network diagrams and dendograms. However, it is my personal opinion that again, there is no single best way of visualizing results, but rather appropriate ways for specific situations. For example, contribution F shows that network diagrams work well with smaller sample sizes, but produce a cluttered, bewildering and hence, useless picture when we have to deal with more than 100 nodes.

4.2.2 Applying Text- and Data-Mining Techniques towards Cowdsourcing-Platforms

Following the literature on text- and data-mining like (Feldmann & Sanger, 2007; Miller, 2005) and using a specialized text mining software (like the Provalis Research QDA Miner with the WordStat extension package in contributions E and F) does not leave much space for drifting away from a well-rehearsed procedure. Usually and likewise applied in contributions D, E and F, text- and data-mining follows a superordinate three step approach, to be precise pre-processing (1), processing or actual data mining (2) and visualization (3).

During all steps, manual implementations are disproportionate, and using a specialized text-mining software product is highly recommended. Contributions E and F introduce two instantiated procedural models, both based on the same standardized three step approach. Figure 10 illustrates the procedure model derived in contribution E and Figure 11 the same from contribution F.

![Figure 10 Procedure model to detect novel ideas in ideation contests by using a text-mining approach (as explained in contribution E).](image)

To be able to enter both procedure models, basic knowledge in web crawling and a structured query language are required. Most specialized text mining software require raw data to be available in tabular format (e.g. .tsv, .tab or .csv) in order to be importable. Hence, usually the database dump (c.f. the exemplary data model illustrated in Figure 2), has to be crawled from the crowdsourcing-platform and the raw text data (c.f. cs_platform_idea:description in Figure 2) has to be extracted using MySQL statements. Depending on the research question, further information should be extracted as well. For instance, as contribution F deals with the ideation function, that
is to say the aggregation of ideas over time, every idea’s timestamp (c.f. cs_platform_idea:created in Figure 2) was extracted as well.

Figure 11 Procedure model to analyze and visualize the ideation function of ideation contests by using a data-mining approach (as shown in contribution F).

Once again dependent on the precise research question, the following steps of pre-processing, mining and visualization can result in a variety of instances. E.g. in contribution F, concept categories instead of the raw text is used for clustering. Therefore, within the previous step of pre-processing, reading and coding all ideas manually, at best using a given taxonomy, is required. Furthermore, scientific rigorousness suggests to conduct coding by using at least two coders and controlling their level of agreement by calculating kappa values or Krippendorff’s alpha (Armstrong, 2001; Creswell, 2008; Menard, 2008; Yin, 2003).

As a comparison of Figure 10 and Figure 11 shows, the essential step of mining also can and has to be implemented in a variety of instances. For example, contribution E analyses the results of an ideation contest, mining (i.e. clustering) is executed on final text corpora of every ideation contests. Whereas contribution F intends to analyze the changing characteristics of cluster over time and hence, introduces a repeated measures design. As clustering itself is conducted by the software, the process has to be triggered manually (725 times in the case of contribution). To sum it up, applying text- and data- mining is by no means an automated process, but requires a lot of manual input and repetitive human tasks. As section 4.2.1 already suggested there are more techniques available than useful within the context of ideation contests. Hence, the procedure models developed in contributions E and F can only serve as a blueprint to a very small set of platforms. However, the general patterns, used in both procedure models can be seen as a standardized approach.

4.2.3 Semi-Automated Detection of Quality on Crowdsourcing-Platforms

The question whether quality can be detected semi-automatically is studied in contribution E. Contribution E develops a procedure model to semi-automatically detect winning ideas in ideation contests. The prefix semi refers to the mentioned fact, that as explained in section 4.2.2, text-mining procedures always require huge manual input.

As winning ideas usually are defined by expert committees, the idea was to simulate their (i.e the expert’s) decision making process by using a text-mining approach. Quite
commonly such committees base their decision on corporate idea or ideation quality criteria, explicitly factors like novelty, feasibility, elaboration and strategic relevance. Contribution E shows that foremost the criterion of novelty is suitable to be detected using a text-mining approach. In general an idea is novel, whenever it is unique and rare, original and not yet expressed by anybody, not related with other ideas, revolutionary and radical, with ability to surprise, imaginary and comes unexpected. The approach of contribution E is to match those criteria with solvers abilities to formulate, that is to say literally express, their ideas in unique, rare, uncommon, unrelated and hence, novel ways. Therefore contribution E uses two different TDMs as well as two different selection measures to evaluate four different models. All models are evaluated using the dataset of over 42'448 ideas in 112 ideation contests. By comparing the real-world expert decisions and text-mining based decision, Precision, Recall and F₁-scores are calculated. The results show that all models score rather high in precision and rather low in recall. This means all models tend to have a low amount of false positives, but unfortunately also a rather high amount of false negatives. In other words, selected submissions are mostly included in the experts picks, but experts mostly pick further submissions on top. The inverse relation between Precision and Recall in all models can be described as rather typical. One can always increase Recall by simply retrieving more documents. As final statement, my personal opinion is that semi-automated detection of quality is possible on crowdsourcing-platforms, but it is also due to some limitations.

As shown in section 4.1.3 and 4.1.4 the results are limited to the specific platform used for evaluation. As mentioned throughout contributions A, B and C, crowdsourcing-platforms use the given design elements in different ways. The second limitation once again refers to the prefix semi. It is my personal opinion that a fully automated decision making process on high quality ideas is not possible (meaning not credible) yet. Hence, results from contribution E rather suggest using text-mining passed approaches as part of a crowdsourcing DSS. Following the procedure model, experts are able to gain a quick overview to all topics addressed in an ideation contest. Furthermore, they are pointed towards ideas that literally stand out. Such ideas might not necessarily be rewarded, but the text-mining approach makes certain not to miss out on them when crawling through large piles of noised data.

4.2.4 Measuring the Ideation Function of an Ideation Contest
The question how the ideation function can be measured and analyzed by applying data-mining techniques is analyzed in contribution F. The developed procedure model introduces a repeated measures design to analyze and illustrate the aggregation of concepts and ideas throughout an ideation contests.

Once again, the explanatory research approach (c.f. section 3.1) plays an important role. Various theories such as (Briggs & Reinig, 2007; Gladwell, 2000; Janis, 1973;
MIT Center of Collective Intelligence, 2012; Osborn, 1957) predict an ideation function, a curve which describes the aggregation of ideation quality over time. Following those theories, contribution F intends to find a replication of one of those curves. However, none of the existing theories could be proven within the case of real world ideation contests. Contradictory, contribution F suggests a digressive growth rate of ideation quality, measured mainly by the introduction of new concepts within an ideation contest. This result is illustrated in Figure 12.

![Figure 12](image)

**Figure 12 Generalized results on the ideation function (as explained in contribution F).**

The interesting results derived in contribution F, come with comparing the growth rates of new concepts (c.f. curve (1) in Figure 12) with further ideation contests describing characteristics, foremost the s-curve of solvers joining the contests and submitting ideas (c.f. curve (2) in Figure 2). The comparison between curves offers in depth insides to an ideation contest and illustrates strength and weaknesses at the same time.

Contrary to theoretical conjectures that the value of ideas in ideation will increase over time, we find that increasing time itself changes the context of ideation. Ideas which are submitted during the second half of an ideation contests (c.f. phases B and C in Figure 12) find utterly different starting conditions and hence, are hardly comparable with ideas that are submitted in the early stage (phase A). In terms of new concepts we find a digressive collinear ideation curve within the real world data. The reason for this is that an early brainstorming phase (phase A) is highly productive and forecloses the chance of later ideas to find a concept that has not yet been mentioned. Albeit this seems to be frustrating in terms of ideation novelty, late ideas still give the impression of being highly useful in terms of feasibility. As the nature of ideation changes within the second half, ideas base on existing concepts and enhance them, mostly by recombining and drawing analogies. As a result, ideas are more elaborate, feasible and with higher potential in strategic relevance. However, these advantages at the same time lower the signal to noise ratio. Nevertheless, phase C (in Figure 12) also introduces a certain freerider problem as the degree of additional value proposition approaches zero during the end of the ideation contest.
4.2.5 Cognitive Abilities of Solver Crowds

Research question II.e (in Table 3) represents the final question of this dissertation. Simply put, the question is: How smart are solvers on crowdsourcing-platforms, individually as well as put together as a group? To be honest, this dissertation will not be able to provide a final answer to those questions. However, combining the findings of contributions A to F, it offers some tendencies and suggestions.

Ideation contest represent a very special form of crowdsourcing. Their tournament characteristic has clear impact on the individual solver’s behavior (also see contribution C). On crowdsourcing-platforms, solvers are not forced or supported in collaborating on given tasks and thus, develop individual strategic behaviors like posting early or late within a contest. Even though, such strategic behavior violates collective intelligence’s requirement of individual, unbiased decision making, at the same time crowdsourcing-platforms offer the wisdom of crowds as a value proposition. To what degree the results of ideation contests finally match the defining characteristics of a collective intelligence cannot be measured in the case of ideation contest. To causal limitations sum this up:

- First, in ideation contest environments organizers habitually do not have knowledge about a single best solution against which ideas can be measured.
- Second, in ideation contest environments there are very few methods to aggregate ideas or submissions to group solutions. Mechanisms to add or multiply the novel information or knowledge carried within ideas have not yet been detected.

Contributions E and F help to lower the complexity of that situation. Contribution E shows, that outstanding ideas are quite regularly included in ideation contests and furthermore can be detected by a text-mining based analysis. Contribution F provides an in-depth view of the crowd’s aggregation mechanism by a data-mining based simulation and visualization. In this case, the early and highly productive brainstorming phase in the beginning of an ideation contest (phase A in Figure 3) is most impressive. A large variety of concepts and ideas usually is submitted within a day. Furthermore, the results of contribution F show that certain concepts among those early movers are able to constitute an opinion leadership, defined by dominant ideas that are adopted more often than others. In the presented case of an ideation contest on the motorbike of the future, a representative dominating concept was an electrical driven cabin scooter.

Dealing with predominant concepts, solvers that enter during later stages (phase B or C) have to scrutinize previous ideas to be able to differentiate themselves. Therewith their task is certainly more labor-intensive. Eventually, their only chance to stand out is by following one of the following two strategies:
• The first strategy is to clearly differentiate to the hitherto submitted ideas by choosing an extreme position. Usually such ideas represent a new way of thinking or a new concept direction.
• The second strategy is to enhance the predominant concept by a particularly clever combination of concepts. Enhancement can be achieved by either adding a smaller, less powerful concept on top of an existing concept (representing a new node and consequently also a new edge in a concept network) or connecting two existing concepts which have been mentioned separately.

To sum up, the crowd includes a lot of creative thinkers, but also freerider and copycats. The crowd is able to adopt existing ideas and develop them further. It is hard to intimidate the crowd with a large pile of existing ideas. Crowds will find a loophole to add a detail or change the direction of meaning. Crowds can be described as playful, following the belief that any bad idea is still better than no idea at all. It is my personal opinion that every interference with the crowd or guidance during the stage of ideation will not enhance their cognitive abilities. Guided by results of contributions B, C, E and F, my opinion is that most platform design elements do not leverage the process of ideation. A creative crowd works best when it is left alone after a task is given.

5 Critical Reflection and Closure
The previous section summarized all results, derived and explained in detail during part B of this dissertation. Hence, during this final section of part A, I will try to construe the impact of those findings for both audience groups (c.f. section 1.3), namely IS scholars and practitioners. Consequently the two remaining questions are:

1) What does this dissertation add towards the knowledge base of IS research in terms of crowdsourcing and ideation contest?
2) How does this dissertation help practitioners in managing crowdsourcing-platforms?

The latter question takes up the motivational questions mentioned in section 1.1. and therewith closes the loop of part A. The first question is addressed in sections 5.1 and the second in 5.2 respectively. Additionally, in both sections limitations towards contributions as well as brief outlooks to potential future research directions are exposed. Part A closes with my personal opinion, reflecting the topic in general, the IS research landscape, and my given results.
5.1 Theoretical Contribution
The theoretical contributions of this dissertation can be summarized under the following heading: Contributions A, B and C rather strengthen the IS knowledgebase on crowdsourcing, while contribution D, E and F extend the IS knowledgebase by applying and evaluating yet seldom used methods to the field. However, both contributions follow the explanatory approach described in section 3.1. Hence, theories as well as applied methods are rather tested and evaluated than invented and developed.

Contributions A, B and C are in line with existing predominant research streams (c.f. section 2.5). Their main contribution is to support the development of basic tenets of crowdsourcing-platforms. Like every new concept, it required multiple studies using a variety of datasets until propositions can be generalized. In such manner, contribution A supports the differentiation amongst platforms by using their comparable characteristics to oppose their strength to their weaknesses. Contribution B helps to understand the various facets of the crowdsourcing phenomena, from a strategic towards a technological perspective. The described examples show, that crowdsourcing does not work from scratch or into the blue. Tasks, technology, time and incentives have to be adjusted to respective contexts. Contribution C picks up those findings and evaluates the effects of a variety of design elements on the outcome of crowdsourcing. Foremost, the contribution strengthens today’s tenets on the effects of monetary rewards and tasks complexities. In accordance with consecutive publications, contribution C finds that both, higher monetary rewards and a lower task complexity will leverage the quantity but not the quality of ideas.

Thus, contribution C represents the bridge towards contributions D, E and F, which deal with the issue of idea or ideation quality within crowdsourcing. As contribution C was not able to find effects on quality at all using hitherto dominant methods of measurement, contributions D, E and F take one step back and question the validity of such methods in general. Within this dissertation, contribution D enters the field of combining text- and data-mining methods towards the field of crowdsourcing. Using real-world feedback text messages that airline passengers sent via the airline’s app, contribution D analyzed how text-mining methods could be applied to aggregate information to a DSS, a crowd controlling dashboard. The contribution points out the variety of pitfalls in applying those methods and provides input to contribution E and F. It is my personal opinion that contribution E and F represent this dissertation’s strongest contributions. The application of text- and data-mining to analyze final and intermediate results of ideation contests provides a hitherto neglected research approach. Not inventing text-or data-mining techniques, but adopting them and recombining them to develop procedure models again is a logical consequence of the explanatory approach (c.f. section 3.1). Hence, the first theoretical contribution of those papers simply is that it is possible to apply those methods, even though it
sometimes requires monotonous and tedious diligence work. The second set of theoretical contributions, made by papers E and F address the research field of crowdsourcing.

Contribution E shows that in over 100 ideation contests, more than often solvers literally stand out due to their phrasing excellence. Even though it does not require rocket science to claim that the verbalization of an idea has impact on readers and examiners, contribution E shows both, how it technically can be proven and how it can strategically be used to analyze ideation contests. In contrast to contribution E, contribution F focusses on analyzing and visualizing the process of ideation next to the final results. The background is, that ideation function have been claimed and proposed by various theorists, but seldom are proven. The reason for this definitely comes with the problem of actually measuring the ideation function itself. Hence, the contributions of paper F once again can be split in two parts. First, the paper shows that and how the ideation function actually can be measured and visualized by applying data-mining and SNA methods. Second, the paper allows us, to draw some limited conclusions on the ideation function itself. Within the exemplary dataset we were not able to find any proof of given theoretical assumptions on the ideation functions. The simple explanation might be, that it is asking too much from 50 year old theories (Osborn, 1957) to fit modern brainstorming incarnations like ideation contests. Consequently, one finding of contribution F is that it requires a lot of analyzing real-world datasets of crowdsourcing-platforms before IS research should try to suggest new theory on crowdsourcing and collective intelligence respectively.

However, none of the contributions presents a final research conclusion. In other words, all contributions have the intention to trigger and foster future research. To the best of my knowledge, next to contributions E and F, not much has been found about the structure and the process of ideation regarding crowdsourcing-platforms. In contrast, IS researchers invest more and more in developing yet another model explaining input-output relations of crowdsourcing (equally does contribution C). In contrast, contributions D, E and F should encourage researchers to mix up methods. The high potential of clustering in detecting hidden patterns in ideation has been highlighted as well as the possibility to actually measure, and not only theorize and philosophize about an ideation curve. Furthermore, following an explanatory approach, the quite simple task should be to know the theoretical possibilities of those methods (i.e. the algorithms and procedures), to analyze and select them, apply them and evaluate their appropriateness. Therefore, it will always require real-world data. It is my personal opinion that a crowdsourcing-theory cannot be built upon experiments in laboratory-type conditions. Hence, often rapid prototyping, a variety of development cycles and a certain playfulness and trial and error mentality will be required and helpful when applying text- and data-mining methods to the research field of crowdsourcing.
5.2 Managerial Implications
Considering recent trends of social media usage in corporate environments and the rising maturity of enterprise 2.0 aligned software tools we can also consider crowdsourcing as an established and increasingly stable trend (Bishop, 2009; Forrester Research, 2012; Landry, 2010). Following the mentioned hype cycle (of Gartner Research), crowdsourcing left the valley of disillusion and is more and more approaching the plateau of productivity. The managerial implications of contributions A to C are foremost intended to support this development.

Contribution A supports the platform selection process and shows practitioners the variety of possibilities in implementing a crowdsourcing campaign. Finally, it intends to help practitioners in avoid commonly made mistakes in designing a crowdsourcing-campaign. Contribution B maybe is of less practical relevance, as analyzed case studies are of slightly different characteristic. However, practitioners might learn about the technical possibilities to implement a crowdsourcing campaign or how to motivate and guide solvers. Therewith, contribution B supports the practitioner’s knowledge base about crowdsourcing in general (a key resource within the business model of crowdsourcing-platforms, c.f. Figure 3). Contribution C follows contributions A and B. The paper illustrates the relevance of understanding effects of given platform design elements on the outcome. Hence, results can be applied in designing future crowdsourcing campaigns, foremost in adjusting monetary rewards, duration or task description in accordance with each other. Furthermore, practitioners are made aware of further external factors, such as the impact of their brand or the platforms maturity on the solver’s behavior.

As mentioned, contributions D, E and F address the task of analyzing and evaluating a manifold, and often cluttered amount of ideas. Currently, this task is commonly performed manually by seeker’s expert committees. Dealing with the final outcome, this task has decisive character about the success or failure of any crowdsourcing campaign. But as more seekers get familiar and comfortable with open innovation strategies and solvers get familiar with posting opinions, quantities will be even larger in the future. In this context, contribution C, D and E support seekers in managing growing amounts of data, deriven from crowdsourcing campaigns, never minding whether 100, 1‘000 or 10‘000 ideas are received. Contribution D shows how information, embedded in crowdsourcing campaigns can be bundled and presented on a performance dashboard. Contribution E provides practitioners with a procedure model to gain a quick overview to all topics addressed in an ideation contest as well as direct insides to literally outstanding ideas. Although by some limitations, the procedure model can support the evaluation process. Finally, Contribution F supports practitioners with an approach to measure the underlying process of ideation. Platform providers can provide seekers with information about the result structure and how people adopted the question and intermediate results. Hence, this information again
can be used to analyze whether the right question was asked and how future ideation contests have to be changed. In the future, IS research should reintroduce its technical component. For instance, research could develop and later analyze a direct or instant integration of text-mining methods on crowdsourcing-platforms. Duplications could be detected instantly by an automated and real-time clustering applying the repeated measures design introduced in contribution F. Therewith, text-mining could become another important design element of crowdsourcing-platforms with definite impact on solver’s behavior. Because as long as we considering demographic change, including digital natives entering the job markets, and more developing countries having strong internet access, quantity values will not fall on crowdsourcing-platforms.

5.3 Critical Reflection and Final Personal Opinion

“It’s important not to overstate the benefits of ideas. Quite frankly, I know it’s kind of a romantic notion that you’re just going to have this one brilliant idea and then everything is going to be great. But the fact is that coming up with an idea is the least important part of creating something great. It has to be the right idea and have good taste, but the execution and delivery are what’s key.”

Sergey Brin in (Kiss, 2009)

Even though it was not stated in relation to crowdsourcing, this disenchanting thought about the value of ideas, expressed by Sergey Brin, the co-founder of Google, largely represents my personal opinion. In order to provide reasoning to this opinion an initial differentiation between various kinds of crowdsourcing-platforms has to be made. Too often, fairly different platforms are tagged using the crowdsourcing label. Platforms for clickwork (like Clickworker.com or Amazon’s Mechanical Turk) are very good tools for enterprises to outsource so called human intelligence tasks. Representative examples are the picture search of a marketing department, the cross-checking of an address lists to support CRM as well as simple translation services. Likewise, social forecasting platforms and prediction markets have demonstrated their capabilities in predicting future developments, foremost when the task is based on estimating and predicting a numerical value or simple up or down?-tendencies. All-encompassing predictions repeatedly are found in cases of stock market predictions or sales planning, but also in Oscar laureates or professional sports such as the Major League Baseball or the National Football League.

However, as mentioned ideation contests represent a different instance of crowdsourcing-platforms. Two key differentiating characteristics are the missing
control value (i.e. the outcome of ideation contests is usually open) and the hitherto missing opportunity to aggregate ideas towards a group solution. On the subject of the latter, it is disenchating how little has been achieved in IS research since EBS were introduced 20 years ago (Dennis & Valacich, 1993). In my opinion, IS research blocks itself by taking crowdsourcing onto a theoretical and even philosophical level. The sad truth is that all existing top crowdsourcing-platforms (e.g. Eli Lilly founded InnoCentive, Dell’s Idea Storm, Amazon’s Mechanical Turk, Google’s Prizes or Starbucks’ MyStarbucksIdea) are products of global enterprises and not based on IS research. In other words, so far IS research had no impact on the development of the actual crowdsourcing-IS. Instead IS research’s main contribution to the field lies in coining terms or taxonomies, outlining research agendas and running countless user-studies, which admittedly apply sophisticated statistical methods. Hence, in my opinion IS research moved away from supporting the development of the IS itself and instead is rather explaining what happened on IS in the past. Also too often, research questions are a pure product of research agendas and miss the practitioner’s actual needs.

Taking the practitioners perspective, the reality is equally blurred. In terms of quality, crowdsourcing-platforms are way worse than their current reputation. To the best of my knowledge, most platforms already follow the unequal distribution of the 90-9-1 rule. Thus, the signal-to-noise-ratio is tremendously low. During countless hours of reading and browsing through crowd-based solutions and ideation contest data, my personal opinion has been disillusioned. The larger those platforms get in terms of unique visitors, page visits and participation frequencies the more they drift away from their initial purpose. This very day, the inspirational spirit, the drive to innovate or improve products and services, is mainly gone on most platforms. For instance, in the beginning of its existence the platform MyStarbucksIdea.com bore new service offers like evening barista classes where customers could learn about the processes behind coffee production or new products like offering a coffee of the day or fair-trade versions of coffees. In the recent two years no such innovations could be accomplished. Instead, the crowd encroaches upon the intention of the platform by using it as a service counter and customer feedback channel. Likewise, suggestions more than often are of repetitive characteristic. Customers ask for a new branch in their own neighborhood, lower prices for certain products or complete a started enumeration, e.g. suggesting an almond latte, whenever hazelnut, walnut and macadamia are already taken.

As this dissertation shows, it would be too easy to blame it all on the crowd. Situations like this, that is to say a cluttered, blurry and hence, nearly unmanageable situation of hundreds of similar ideas, more than often is also a result of a seeker’s lack in understanding the importance of platform design elements. In other words, offering high stakes for very simple tasks, e.g. 10’000 US$ for branding a new product or
finding a catchy URL, can be seen as a free ticket to confusion. In such cases the introduced procedure models might help. However, combining all results of this thesis, early consideration of those situations, knowledge on the effects of design elements (foremost a sound combination of task description, reward and contest duration) and proper methods for analysis and evaluation should foreclose negative effects. Hence, it is my personal opinion, that today often too obvious tasks are being crowdsourced, and that seekers overestimate the capabilities of crowds or underestimate the capabilities of their own employees respectively. Crowdsourcing has potential, but to the best of my knowledge, we are far from getting the best out of it.

To be quite honest, I am not certain to what extend this dissertation can be excluded from the mentioned claims. Following the explanatory approach it was not my intention to build new theory or develop new text- or data-mining algorithms. Instead I tried to make best usage of the already known, theories as well as methods. As the previous sections have shown, text- and data-mining are underused methods and their application towards crowdsourcing can help practitioners a lot. However, they do not seem to be highly accepted in IS research. As only few specialized conferences offer tracks for applying such methods, publishing in highly ranked IS related journals currently seems like tilting at windmills to me.

6 Literature


Forrester Research. (2012). Key Strategies to Capture and Measure the Value of Consumerization of IT.


Part B
A.1 Einleitung

A.2 Das Crowdsourcing Konzept
einen öffentlichen Aufruf an eine große Gruppe. Typischerweise stehen Problemlösung und Ideengenerierung im Zentrum, aber es sind auch repetitive Aufgaben möglich. In der Regel wird dieser Aufruf durch eine Webseite realisiert."


2. Ideen- und Konzeptvorschläge aus der Crowd, bzw. deren Aggregation, können die Qualität internen Vorschläge übertreffen.

![Abbildung 1 Crowdsourcing Concept Map](attachment:image.png)

Konkret findet Crowdsourcing heute oft in Form von Design- oder Ideenwettbewerben, sogenannten Ideation Contests oder Innovation Tournaments [Terwiesch und Ulrich 2009], statt. Hierbei wird nicht jeder Teilnehmer, sondern nur diejenigen mit den vermeintlich besten Beiträgen finanziell entlohnt. Gemeinsam haben alle diese Ansätze jedoch ein suchendes Unternehmen (Seeker genannt), welches bereit ist, eine interne Aufgabe an eine Masse von Freiwilligen (Solver oder Crowd genannt) auszulagern. Steht zwischen Solver und Seeker eine unabhängige
Webplattform, welche die Aufgabe eines Maklers übernimmt, entsteht ein sogenannter zweiseitiger Markt.


1. Meinungsvielfalt unter den Teilnehmern: Zum einen bedeutet dies, dass eine kritische Masse an Teilnehmern erreicht werden muss, um verschiedenartige Vorschläge überhaupt einbringen zu können. Zum zweiten ist hierunter zu verstehen, dass die Wahrscheinlichkeit, dass ein Teilnehmer eine passende Lösung findet, umso höher ist, je unterschiedlicher der Hintergrund der Beteiligten ist.

2. Unabhängigkeit der Teilnehmer: Beeinflussen sich die Teilnehmer gegenseitig, kann dies das Ergebnis verfälschen. Die Meinung des Einzelnen muss daher unabhängig von der letztlichen Gruppenmeinung sein, d. h. es darf keine Meinungsführerschaft geben.


Obwohl also Crowdsourcing-Kampagnen oder Ideenwettbewerbe per Definition nicht zwingend eine kollektive Intelligenz hervorbringen müssen (denn es geht wie beschrieben lediglich um den Prozess der Auslagerung), wird gerade eben dies von vielen Crowdsourcing-Anbietern versprochen. Oft wird die Crowd vom Plattformanbieter als Innovatoren-Community dargestellt, welche generell die Möglichkeit besitzt, in verschiedenartigsten Aufgabengebieten zu innovieren (man spricht oft von User-centric Innovation bzw. Customer Co-Creation). Im folgenden Kapitel werden diverse Crowdsourcing-Plattformen kategorisiert und näher betrachtet.
A.3 Kategorien von Crowdsourcing Plattformen

Abbildung 2 Crowdsourcing Plattform-Typen.

A.3.1 Crowdsourcing Makler
Crowdsourcing-Makler sind Intermediäre, welche als Bindeglied zwischen lösungs- oder Ideen suchenden Unternehmen auf der einen und einer ausführenden, designenden oder lösungsgenerierenden Crowd auf der anderen Seite fungieren. Nach Anforderungsniveau und gewünschtem Ausarbeitungsgrad von auf der Plattform gestellten Aufgaben können hier drei Sub-Typen von Maklern unterschieden werden:

1 Die in diesem Kapitel genannten Plattformen und Unternehmen werden nicht mit der kompletten URL oder als Literaturhinweis angeführt, sind aber über eine Websuche direkt auffindbar.

2 Vgl. u.a. die Angebote von help Ideenmanagement, hypeinnovation oder HYVE.
Forschungs- und Entwicklungsplattformen, Marketing-Design- und Ideenplattformen sowie Plattformen für Freelancer.


A.3.2 Direktes Crowdsourcing über die Unternehmenswebseite

**A.3.3 Verkaufsplattformen für Crowdsourcing Design**

**A.3.4 Unternehmensinternes Crowdsourcing**

In einigen Ländern sind innerbetriebliche Crowdsourcing-Plattformen aus Gründen der Mitbestimmung jedoch zwingend mit dem Betriebsrat abzustimmen. Es hat sich zudem ein breiter Markt an Technologieanbietern gebildet, die Großunternehmen bei Konzeption, Inbetriebnahme und im laufenden Betrieb innerbetrieblicher Crowdsourcing-Plattformen mit eigener Software unterstützen².

A.4 Analyse der Crowdsourcing Plattformen


² Vgl. u.a. die Angebote von hlp Ideenmanagement, hypeinnovation oder HYVE.
A.4.1 Stärken von Crowdsourcing Plattformen

A.4.2 Schwächen und Bedrohungen von Crowdsourcing Plattformen
Für Crowdsourcing Anbieter ist es heute deutlich schwerer, das Momentum der First Mover zu erreichen. Vier Fallstudien aus 2008 zeigen z. B., dass Solvers rein über den

A.4.3 Möglichkeiten und Chancen von Crowdsourcing Plattformen


A.5 Literatur


**Contribution B: Crowdsourcing as a Business Model: An Exploration of Emergent Textbooks Harnessing the Wisdom of Crowds**

<table>
<thead>
<tr>
<th>Title</th>
<th>Crowdsourcing as a Business Model: An Exploration of Emergent Textbooks Harnessing the Wisdom of Crowds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authors &amp; Affiliation</td>
<td>Thomas P. Walter, Andrea Back, Institute of Information Management, University of St. Gallen, Mueller-Friedberg-Strasse 8, 9000 St. Gallen, Switzerland, {thomas.walter, andrea.back}@unisg.ch</td>
</tr>
<tr>
<td>Publication Type</td>
<td>Conference Paper</td>
</tr>
<tr>
<td>Year</td>
<td>2010</td>
</tr>
<tr>
<td>Status</td>
<td>Published</td>
</tr>
<tr>
<td>Full Citation</td>
<td>Thomas P Walter, Andrea Back, Crowdsourcing as a Business Model: An Exploration of Emergent Textbooks Harnessing the Wisdom of Crowds, in 23rd BLED conference, Bled, Slovenia, July 2010, p. 555-568.</td>
</tr>
</tbody>
</table>

*Table 1 Bibliographical Information for Contribution B.*

**Abstract:**

The process of writing textbooks is still very traditional regarding the status of authorship and expert opinions. Recently we observe the emergence of authors who follow a different approach, tapping the wisdom of crowds as key resource of their own publications. In this paper, we explore business model innovation which leverages value propositions of textbooks by applying crowdsourcing. We use case study research methods to analyze four textbooks written collaboratively. Essential findings indicate occurrence of user-communities fulfilling peer-reviewing, editing or co-authoring despite a lack of monetary incentives. We further detect a tendency towards wiki software providing a community hub. This paper enters the field of partially crowdsourced textbooks and derives future questions of research.

**Keywords:**

crowdsourcing, business model innovation, collaborative writing, textbooks, case study research

**B.1 Introduction**

Within studies of Management or Information Systems textbooks are still state of the art. Approximately 30 % of US book sales can be traced back to textbooks, which include academic and management books. Even though the book-selling industry was hit by the recent crises, the textbook sector and the US college book in particular were least affected. 22.6 million single copies (Nielsen 2009), bringing in US$ 3.8 billion
net sales state best growing category of US resellers. (The Association of American Publishers 2009). Furthermore 2010 should be the year when book trade begins a recovery (Key Note 2009). However, production of textbooks still is mainly due to past restrictions regarding authorship and expert opinions. Commonly a group of self-constituted experts or single thought leaders write and publish textbooks. But more recently we find outstanding exceptions to this common practice. Some authors start breaking out of the common business model by harnessing the wisdom of crowds. Next to their own knowledge these authors utilize crowdsourcing as a key resource to their business model.

The crowdsourcing of textbooks has not been analyzed from an Information Systems perspective yet. During a literature search in titles or abstracts on Ebsco Host, no results were found including the terms crowdsourcing, textbook, management books, wisdom of crowds, business model innovation or any combination of these. Even though the idea of crowdsourcing is established in Information Systems and various forms of online collaboration also exist for years, the observed phenomenon of application is new, innovative and maybe disruptive. Hence, harnessing the wisdom of crowds as input of own publications states major changes to the industries common business model. Up to now professional crowdsourcing campaigns are more likely set up by companies and concerning technological aspects. Hence, established case studies discuss idea contests, open innovation (Chesbrough 2006), lead-user co-development of products (von Hippel 1986) or open source software products. Now pluralities of questions remain due to observed phenomena. How did the authors apply crowdsourcing? What triggered and spurred the crowds’ participation? Which roles did the crowds fulfill? Which old business models are affected and which new business models are occurring?

Within this article we attempt convergence of the phenomenon by using case study research methods. (Yin 2009) Analyzing authors who partly crowdsourced the composing of textbooks enters the field and shapes further questions of research as suggested in Eisenhardt (1989). We make use of business model theory as a constitutional perspective and focus on textbooks. Hence, we exclude collaboration projects in writing fiction or consumer books such as www.webook.com or www.bookbymany.com which are platforms of hobby authors. We also exclude collaborative online projects which were not published such as the www.wikibooks.com project. Last, we do not differentiate between channels of publishing, i.e. print versions, ebooks or audiobooks. Unlike Wikipedia had disruptive impact to printed encyclopedias we do not see evidence that ebooks will have medium-term consequences on the textbook industry. Chapter 2 provides basic literature of business models and the aspect of crowdsourcing. Furthermore, we deduce a grid to describe case studies. In chapter 3 we specify four descriptive case
studies and chapter 4 includes a brief cross-case analysis (Yin 2009) by adopting the grid of chapter 2. Finally chapter 5 provides a critical reflection and a short conclusion.

B.2 Crowdsourcing as Business Model Innovation

B.2.1 Business Model Definition as Framework

With their efforts of harnessing the wisdom of crowds as input to their publications authors have deviated from the default business model of book writing. Since business model is a frequently relevant (Magretta 2002) but strained term we initially have to give our understanding of it. Numerous business model definitions include the terms “value” or “revenue stream” (Al-Debi, El-Haddadeh and Avison 2008) and subject areas of business model research are mostly “e-business”, “strategy” or “information systems” (Pateli and Giaglis 2003, Pateli and Giaglis 2004). In regard to this papers topic we follow the business model definition of Rajala and Westerlund (2005) which includes the factor of collaboration:

“A business model describes ways of creating value for customers and the way business turns market opportunities into profit through sets of actors, activities and collaborations.”

Similar definitions which also include the aspect of collaboration can be found in Torbay, Osterwalder and Pigneur (2001), Camponovo and Pigneur (2003), Afuah and Tucci (2001) or Andersson et al. (2006). We see this paper as initial step to adjust further research. Therefore, we see business model theory as convenient because it is conform to broaden the topic in continuing research. However, within this paper we neither do want to evaluate if the observed business models fit in with defined standard e-business models as developed in Clemons (2009), Rappa (2005) or Timmers (1998) nor do we want to develop a new canvas to describe business models as Johnson, Christensen and Kagermann (2008) or Osterwalder and Pigneur (2009). Basically, we will apply the definition to set up a simple framework of cross case analysis and hence, make our cases comparable. Yin (2009) calls this to develop a case study protocol.

Value creation is often set as core of a business model. (Johnson, Christensen and Kagermann 2008, Magretta 2002, Timmers 1998, Afuah and Tucci 2001) We restrain our focus on the aspect of crowdsourcing as a way to create value for customers which in our cases are readers of textbooks. Hence, we will provide brief descriptions on how crowdsourcing is applied in the different cases by analyzing which part of value creation is taken by the crowd. The second part of the definition focuses on actors, activities and collaboration as factors of value creation. Hence, we will provide brief descriptions of how the crowd is assembled, which incentives of participation are offered, which hurdles exist, which tasks are fulfilled by the crowd and how collaboration is technically supported. Business model definitions also strongly focus on output and revenue streams. (Rappa 2005, Magretta 2002, Timmers 1998) With
crowdsourcing as major topic we concentrate on the production aspect of textbooks. A focus on the revenue streams would include the entire value chain from publishing houses to resellers, lectorship and customers and should be done in follow-up research. Table 1 summarizes our business model description framework of crowdsourced textbook writing.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Focus within case studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value creation by crowdsourcing</td>
<td>How does the crowd add value to the product (textbook)? Which tasks do crowds fulfill? Does action match with the definition of the term crowdsourcing?</td>
</tr>
<tr>
<td>Crowd description</td>
<td>What size is the crowd? How is the crowd assembled? Are there aspects of lead-users?</td>
</tr>
<tr>
<td>Incentives</td>
<td>Which incentives are set up by the main authors to spur participation? What types of incentives (monetary, acknowledgement, fame, learning, etc.) are set up?</td>
</tr>
<tr>
<td>Hurdles</td>
<td>What are hurdles of participation? How easy can the collaboration process be joined? Are there any strict limits to participation?</td>
</tr>
<tr>
<td>Technical Solution</td>
<td>How is the crowdsourcing process backed up technically? What web-solution to leverage collaboration is applied?</td>
</tr>
</tbody>
</table>

Table 2 Characteristics of business models using crowdsourcing as a key resource

B.2.2 Crowdsourcing as Key Factor of a Business Model


“Simply defined, crowdsourcing represents the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call. This can take the form of peer-production (when the job is performed collaboratively), but is also often undertaken by sole individuals. The crucial prerequisite is the use of the open call format and the large network of potential laborers.”

As we intend to evaluate the degree of value creation by crowdsourcing within the described case studies more classification of related topics is required. Consequently, crowdsourcing can be described as the process of harnessing or leveraging the wisdom of crowds. Wisdom of crowds state that the aggregation of information in groups results in decisions which are often better than made by any single member of the group. (Surowiecki 2004) Moreover, the term crowdsourcing reintroduces open innovation (Chesbrough 2006). Using crowdsourcing as key resource in a business model can also be found in various case studies on open innovation. (Brabham 2008, Howe 2008) Differentiating time and amount of crowd involvement, idea contests and co-development of products can be distinguished.
Idea contest state that a company (or a problem seeker) announces a monetary reward to motivate customers (or potential problem solvers) to contribute their proposal for a problems solution. (Chesbrough, Vanhaverbeke and West 2006) Among many, prevalent cases include the projects of Cisco iprize, IBM Jamming, or MyStarbucksidea. Also new business models containing and selling crowdsourcing as value proposition have developed out of the concept of idea contests. Companies such as InnoCentive or InnovationXchange function as intermediaries between problem seekers and potential crowds of problem solvers (Brabham 2008). Research also focuses on how to set up incentives, rules and technical platform of idea contests. (Leimeister et al. 2009, Ebner et al. 2008) The concept of idea contests also has to be circumscribed from intermediary platforms selling user generated contend such as istockphoto. Co-development of products was described as the lead-user concept by von Hippel (1986). Roughly the concept indicates to include best and technologically well proved customers into companies’ innovation process. Dependent on the intensity of the crowds’ involvement the lead-user concept can vary from co-development and advancement of existent products (as in the case of Lego factory) to the simple rating of internal ideas as market study (as in the case of threadless.com).

B.3 Case Studies
In this chapter we deliver brief descriptions of the applied case studies, following the characteristics of the simple framework given in Table 1. There has not been a pre-selection underneath a plurality of possible case studies. However, we see these case studies as representative first mover examples of crowdsourced textbooks. Additionally to adopting the framework, we provide brief descriptions of the textbooks content and which initial material was brought into the crowdsourcing process by the main authors. Finally we describe the various outcomes and to which degree the case study represents a business model innovation.

B.3.1 Charles Leadbeater – “We think”
Leadbeater (2008) explores the ways in which mass collaboration is dramatically reshaping our approach to work, play, and communicate. The author took the success of Wikipedia as a prime example and initiated a process of collaborative writing and editing of his book. Therefore he posted as a first draft of the book chapter by chapter to the web using a wikia wiki in 2006. (Leadbeater 2009a) During a time period of 12 month his initial version was downloaded thousands of time, edited and new information was added by unpaid and widely unknown collaborators (Leadbeater 2009b). By putting up an entire first draft of the book rather editing and peer-reviewing processes were crowdsourced than the actual writing of the book. However, regarding to the author at some parts there was not much left of the initial draft. There were no restrictions to participation and the wiki is still editable by everybody. (Leadbeater 2009b) Due to his “no barriers policy” Leadbeater was not able not track the real size of the crowd. But the authors proposition of “hundreds” of useful
comments but “thousands” of downloads tells its own tale. The definite amount of free-riders which only consumed but did not review or edit content was not measured either. To spur participation Leadbeater announced, that all editors and comment providers are acknowledged in the print version of the book, which are 237 contributors. There were no monetary incentives.

Consequently, this case study meets the requirements of the crowdsourcing definition, as Leadbeater invited everybody to contribute to the wiki. But the focal point has to be adjusted. Rather than crowdsourcing the content of the textbook, the value proposition arises from comments and feedback. Hence, Leadbeater sets a prime example how a first draft of a textbook can be peer-reviewed by applying crowdsourcing.

**B.3.2 Osterwalder and Pigneur – “Business Model Generation”**

Osterwalder and Pigneur (2009) present a business model framework, based on nine building blocks. The production of the book was based on the PhD-thesis of Osterwalder (2004), which had been accessible in the web for free. As the topic of the book is on business model generation the authors clearly pointed out that also during the process of creating a textbook they want to practice what they preach and hence, generate a new business model of textbook production. Based on www.ning.com they created a hub for a rising community of co-authors. His community grew to a final size of 470 co-authors. Especially strategy practitioners and business model experts responded to the authors open call for participation. Interested collaborators had to pay an initial participation fee of US$ 24. As the community grew, the authors stepwise raised fees until a final amount of US$ 250. (Osterwalder 2009) Main tasks of co-authors were to criticize chapters and contribute sample applications of business model innovation from practice. Additionally the authors organized a workshop, where the entire community of co-authors also met physically to discuss potential topics of the textbook. According to the authors, incentives to co-authors were to be the first to read and discuss new content on the topic of business model generation. Co-authors also paid to be part of the collaboration process during which they learned from each other. Finally all co-authors were mentioned within the textbook. As hub of the community the authors set up a social network service on the basis of www.ning.com. They unlocked nings’ premium services to enable their own URL www.businessmodelhub.com and remain free of advertisement.

Summed up, the case study meets most requirements of crowdsourcing a textbook. However, asking for community subscription fees and thus, creating an additional revenue streams do not go along with the idea of an open call. Furthermore, the authors’ demand of best practice insights can be characterized as co-development of the textbook content by the crowd. Therewith the value proposition shifted away from core content of the textbook. On the contrary the authors exceeded crowdsourcing requirements as they not only addressed a crowd but also built up a persistent expert-community.
B.3.3 Crumlish and Malone – “Designing Social Interfaces”

Crumlish and Malone (2009) present social web design principles and interaction patterns thus capturing user-experience best practices and emerging social web customs for web 2.0 practitioners. The authors set up a patterns wiki as a companion site to their book. They opened a major wiki category for each chapter of their book and various sub-categories equaling sub-chapters respectively. They shared content from the minute of production by providing it in the wiki based on MediaWiki software. Hereby they intended to collect community feedback and enhance discussions particularly during the stage of writing. (Crumlish and Malone 2010) Their intention was to strengthen the content with a variety of opinions and elevating reliability and representativeness. Hence, their open call addressed the community to make use of the wiki in a forum way, by adding opinions underneath an entry and not overwriting the same.

Next to the wiki the authors set up a photo stream on the flickr platform to provide illustrations which possibly could be included in the textbook. Contributions by the crowd concerning insights to business best practices were included as essays in separate boxes. Only contributors who added content by mentioning of their names were kept. Anonymous contributions or by nickname were deleted. 21 best practice insight essays found their way into the final textbook. This states half of all active contributors to the wiki and a tenth of all signed in users. There was no participation hurdle. Everybody was and still is allowed to enter information into the wiki. (Crumlish and Malone 2010) The authors did not announce any other kind of direct incentives as the aspect of learning from and discussing with each other. It was not announced that 21 “essayists” are thanked with reference to their specific contribution until near completion of the textbooks content. Summarized this case study provides a solid reference how crowdsourcing can be applied to harness best practice insights for a textbook. Once again, the value proposition is not producing key content of the textbook but rather peer-reviewing and giving insights to best practices.

B.3.4 Williams and Tapscott – “The Wikinomics Playbook”

Wikinomics written by Tapscott and Williams (2006) consists of 12 chapters of which only eleven are written. Within the twelfth chapter the authors invite readers to write it for themselves collaboratively: “Join us in peer producing the definitive guide to the twenty-first-century corporation.” As Wikinomics already deals with the power of mass collaboration transforming economy and society the authors intended crowds to collect references from practice regarding propositions made during the first 11 chapters of the textbook, called the Wikinomics Playbook (Tapscott and Williams 2008). Hence, they set up a wiki on the technical platform of socialtext as a community hub. As initial content the authors placed a reduced chapter outline to narrow down topics of interest. (Williams 2010) Thereby they provided a framework of potential questions and tasks of which they called for response. Over the course of
2007 a community of readers and experts formed a life of its own. As the authors did not anticipate the amount of participation they decided to transform the former twelfth chapter into a self-contained book. As wiki owners the authors initiated a role concept. They separated so called “researchers” which contributed most of the content from so called “writers” which mainly reformulated initial contend and “editors” which shaped, trimmed and reorganized the content. Furthermore and on voluntary basis, wiki contributors who strongly felt responsible for a specific article were asked to become “lead-authors”. This role additionally included linking articles to own user profiles and supervising changes within these specific articles. Since the decision to publish wiki entries as book the authors additionally set up an editorial board. Next to Williams the board consisted of three more members chosen from the community. Every contributor was asked to link to an own user page where real name, an institution and experiences were provided. Anonymous contributions were not allowed. The 20 most valuable contributors to the wiki are noticed on the cover of the textbook as authors, all contributors are acknowledged in the book. The authors measured top contributors by number of pages contributed to, whether the edit was the current (last) edit on a page and by significance of all contributions. Contributions from lead-authors where also published on a companion blog under a guest author column. (Tapscott 2010) The authors also announced that the opportunity to continue the dialogue is by no means over. The Wikinomics Playbook should be considered as version 1.0 and the wiki still should be filled. They offered the book in print, ebook and as audio version.

With their repeated open call to join the community and produce the last chapter of the Wikinomics collaboratively Williams and Tapscott provide a solid example of crowdsourcing content of a textbook. The value proposition of the crowd is set by contributing insights to industry adoptions. But here, the crowdsourcing approach is set as post-production-processes of textbooks to avoid outdated practice insights.

### B.4 Analysis and Shaping Hypothesis

Following Eisenhardt (1989), in a next step we analyze the data applying a brief cross-case pattern search. Initially, Table 2 provides a summary of observed characteristics within the cases. Following, we discuss each topic separately, intending to shape hypothesis and enter the field of further research.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Textbook-topic</td>
<td>mass intelligence and mass collaboration</td>
<td>business model generation</td>
<td>social media patterns</td>
<td>mass collaboration</td>
</tr>
<tr>
<td>Crowd-sourcing part</td>
<td>editing, commenting, re-writing an initial</td>
<td>discussing textbook-topics in a social</td>
<td>contributing and commenting to textbook content</td>
<td>writing, editing, reviewing and editing a</td>
</tr>
</tbody>
</table>

With their repeated open call to join the community and produce the last chapter of the Wikinomics collaboratively Williams and Tapscott provide a solid example of crowdsourcing content of a textbook. The value proposition of the crowd is set by contributing insights to industry adoptions. But here, the crowdsourcing approach is set as post-production-processes of textbooks to avoid outdated practice insights.
Within all cases authors practice what they preach. The textbook on business model generation states a new business model of book production, the textbook on mass intelligence leverages the wisdom of crowd as a reviewing resource and the book on designing social interfaces obtains practice examples of wiki adoption from users of a companion wiki to the textbook writing process. Justifiable questions remain: Can these business model innovations be transferred and deployed to write textbooks on more classical topics such as marketing, human resources or even controlling? Or is crowdsourcing of textbooks only possible if a textbook itself also deals with a related topic? Examples of collaborative written fiction found on www.webook.com or www.bookbymany.com prove that collaborative writing is possible. However, textbooks state a special issue. A major buying incentive of textbooks is to understand complex issues and achieve comparative advantage by comprehending expert knowledge. Hence, a preliminary answer derived from our case studies would be that crowdsourcing a textbook only is possible if the textbook concerns a related topic. At least our research does not show any counterexample, e.g. why one should buy a crowdsourced textbook concerning atom physics or electrical engineering?

### B.4.2 Crowdsourcing as value proposition

All case studies show that crowdsourcing approaches do not work from scratch. Correspondent authors provided initial material to spur the participation of crowds, i.e.
we found no example of crowds providing innovative content out of nowhere. Crowds more likely fulfilled reviewing, editing and commenting tasks but core messages still were provided by main authors. However, value propositions of crowds were to provide insights into daily business routine and to deliver use cases. Hence, we see this also as lead-user and customer co-development approaches. Another unmeasured value proposition is the degree to which authors leveraged network effects and hence, set off viral marketing campaigns by establishing communities. (Shapiro & Varian 1998, Rosen 2005) Two questions derive: Is it also possible to crowdsource an entire value creation, i.e. content production process, or is initial input by main authors required as incentive and basis of discussion? Here, the case of Tapscott and Williams (2008) can be seen as a first approach. A second question is, whether standard strategies for applying crowdsourcing throughout several steps of value creation can be derived? E.g. using a wiki to let the crowd peer-review a textbook, or setting up a community of experts to provide use cases and best practices?

B.4.3 Crowds
Crowds can vary significantly and hence, have to be measured whether they are adequate to the software which is used as community hub and the task which is expected to be fulfilled. Within all cases the active crowd consists of hundreds of people, i.e. 219 to 470 signed in users. Few statements of main authors evaluating the amount of active users show that approximately a tenth of the crowd accounts for nearly its entire value proposition. As educational and working experience backgrounds have not been measured by the main authors it remains difficult to profile crowds. According to Osterwalder and Pigneur (2010) mainly strategy practitioners and business model experts felt allured by their call for participation. But this also has to be considered as marketing statement to strengthen the value of the textbook. However, the case of Leadbeater (2009) showed that an open approach implicates a free-rider problem. How to deal with crowds implies issues of trust. Questions are about how open access to initial content is provided and what authors expect the crowd to share. Also the question what kind of contributors are necessary to provide value propositions remains. If authors primary focus on harnessing business insights then crowd members should at least have related job positions. Otherwise main authors run the risk of devaluing the content.

B.4.4 Participation Hurdles
Basically, two attitudes towards participation hurdles can be found in our case studies. Authors which used wikis as technical hub of the community set the participation hurdle to sign-in and contribute labor time at most, whereas Osterwalder and Pigneur (2010) set a monetary hurdle by demanding a participation fee. Clearly their approach of “asking crowds to pay if they want to do the authors’ work” appears to be contradictory. But according to the authors, the monetary hurdle was understood as a seal of quality for the community of business model innovators. From the authors’
perspective, participation fees involve the additional advantage of an early stage revenue stream. Last, with increasing crowd the amount of participation fee can be adjusted as the value of the network raises (Shapiro and Varian 1998). Hence, Osterwalder and Pigneur started with initial fee of US$ 24 per participant but were able to ask for US$ 250 at a crowd-size of 400+ due to classical network effects. From this single case perspective it seems like a prime example of setting up participation hurdles. Solely experts are attracted and accomplished plus additional revenue is generated which can be taken as funding of production costs. But the case also elevates the question of transferability. Once again, the question derives if open access to textbook content can be counterproductive because it is strongly associated with the “free credo” (Anderson 2008) and not noticed as a valuable textbook.

**B.4.5 Incentives**

Although we did not research for incentives specifically the cases show that crowds spur of participation does not compulsory have to be of monetary kind. Instead of attracting with potential dividends authors traded off crowd participation against the potential of getting noted as expert or being part of a community of experts. (Framke and Shah 2003) However, also free-riding and collecting expert-knowledge must be considered as incentive. The case of Osterwalder and Pigneur (2009) shows that people even are up to pay fees to participate in crowdsourcing campaigns if they want to be a part of something bigger, discuss a topic with experts or recognize chances of enriching individual networks. Still questionable is, if and how long “to be part of something big and have your name mentioned in a textbook” will remain as incentive? Here, recent case studies surely had a first mover advantage. Further research should focus on incentives separately. Are incentives dependant on the content of a textbook or correspondent to main authors or can they stand autonomous? What are the indicators to measure these incentives? Another task is to determine since when incentives cause negative effects, such as project free-riders and if hurdles are necessary to avoid these problems.

**B.4.6 Technical solutions**

The amount of only four cases does not allow general propositions on suitable technical solutions for crowdsourcing initiatives. However, we see a tendency, as within three cases authors used three different wiki solutions to create a hub for the community. This seems to be convenient as wikis support a collaborative but still sequential editing of text and can be filled with initial content by the main authors. The tendency towards wiki solutions diverges from the ways of wiki usage within the cases. Crumlish and Malone (2009) prohibited the overwriting of content and commentary. Technically, they converted the wiki into a forum. Williams and Tapscott (2009) provided a prime example how to apply wiki software and set up clear policies of usage. Leadbeater (2009) left behavior rules completely open and trusted crowds’ behavior to only overwrite parts they think they know better. From technical
perspective Osterwalder and Pigneur (2010) established the only real community underneath our case studies. Their premium community on ning.com is the only platform requiring rich user profiles from the crowd, and maintaining their own blog within the community. Furthermore Osterwalder and Pigneur (2009) as well as Tapscott (2010) and Leadbeater (2009) ran a companion blog, posting about the project and hence, handling marketing and public relations. It remains questionable if wikis can be taken as a standard tool to apply crowdsourcing approaches of textbook-production, or a text of common derivation respectively?

B.5 Conclusion
The studies illustrate that former business models and role allocations during a book writing process can be turned upside down by applying crowdsourcing approaches. People who were considered as readers (customers) in former business models can become reviewers (lead-users), editors or co-authors (co-developers). The wisdom of crowds can displace reviewers or paid lectorship or even add new content and offer valuable insights to best practices or business scenarios. Our analysis of business model characteristics of crowdsourced textbooks led us to questions of further research which are finally summarized in Table 3.

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Research-Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Book topic</td>
<td>Can prime examples of case studies be transferred and applied in books on topics not related to crowdsourcing or business models?</td>
</tr>
</tbody>
</table>
| Crowd-sourcing part | Is value proposition of the crowd limited to providing best practice insight and use cases?  
|                  | Are there standard strategies for applying crowdsourcing throughout several steps of value creation? |
| Crowd            | Are there optimal crowd sizes?  
|                  | How should projects deal with free-riders?                                      |
| Participation hurdle | Are monetary hurdles necessary to avoid free-riders and allocate an ambitioned crowd of experts?  
|                  | Are there standard entrance fees of participation and should they be raised as the network value growths? |
| Incentives       | Are incentives dependent on the topic or authors or autonomous?                
|                  | What are the indicators to measure these incentives?                              
|                  | Since when do incentives cause negative effects?                                |
| Technical solution | Can wikis be taken as standard tool to apply crowdsourcing approaches of textbook-production?  
|                  | Is a companion blog the new standard to provide marketing?                     |

*Table 4 Questions of further research derived from case study research*

Further research will originate findings regarding these questions. For the moment open issues remain. One main critical self-reflections, that authors did not measure data during the main production phase. Hence, it is not able to draw conclusions on impacts of certain incentives. Finally, we have to deal with the issue that all four case
studies state successful projects. Hence, we lack of information about failed approaches to apply crowdsourcing in textbook writing. A first instance could be Guy Kawasaki’s intention to “tap the wisdom of the crowd” for his next book and for which he opened a wiki. (Kawasaki 2006). Another example could be Kiruba Shankar’s (Shankar 2009) failed approach to write a book on crowdsourcing, including 140 opinions of users sent via the microblogging service twitter. Unfortunately these authors were not open for a case study.

B.6 References


Osterwalder (2004). The Business Model Ontology - a proposition in a design science approach. Dissertation from University of Lausanne, Switzerland.


Contribution C: Towards Measuring Crowdsourcing Success: An Empirical Study on Effects of External Factors in Online Idea Contests

Title
Towards Measuring Crowdsourcing Success: An Empirical Study on Effects of External Factors in Online Idea Contests

Authors & Affiliation
Thomas P. Walter, Andrea Back, Institute of Information Management, University of St. Gallen, Mueller-Friedberg-Strasse 8, 9000 St. Gallen, Switzerland, {thomas.walter, andrea.back}@unisg.ch

Publication Type
Conference Paper

Year
2011

Status
Published

Full Citation

Table 1 Bibliographical Information for Contribution C.

Abstract:

Crowdsourcing campaigns in general and online idea contests in particular have reached popularity in practice and research. Based on the open innovation paradigm, idea marketplace providers like InnoCentive broker ideas between a potentially large-scaled crowd of solvers and seeking firms by hosting idea contests in the web. So far, research has mainly focused on the setup of idea marketplaces, design patterns to define idea contests or motivation factors to spur solvers. Mostly, extrinsic incentives, especially rewards for best ideas, are used to guarantee high amounts and good quality of ideas. In contrast to the predominant qualitative and conceptional research on online idea contests, we attempt an empirical study. We derive external factors, like rewards, contest-duration or market-maturity, to construct a model depicting effects of these factors on the outcome of online idea contests. We use a dataset from Atizo, containing over 25’000 submitted ideas and over 83’000 ratings to evaluate validity of our model. Our findings show that current design elements of online contest may not contribute to the commonly expected effects. One interesting aspect is that using rewards to maximize quality and quantity of ideas creates a conflict of goals as rewards can have perverse effects on the outcome.

Keywords:
idea contests, crowdsourcing, open innovation, empirical study
C.1 Introduction

Online idea contest, also called innovation tournaments (Terwiesch and Ulrich, 2009) represent a relatively new form of electronic marketplace. Also often referred as crowdsourcing (Howe, 2006, Howe, 2008) approach, companies or institutions take an idea- or concept-seeking process, once performed by employees, and outsource it to an undefined (and generally large) group of people in the internet. In this occurrence of an electronic marketplace, ideas (potential innovations) are traded like information goods (Shapiro and Varian, 1998), which are produced by working consumers and demanded by solution-seeking companies or institutions (Singh and Wittman, 1988, Archak and Sundararajan, 2009). Like in other electronic marketplaces, operating the marketplace itself has become a promising business model. Companies such as InnoCentive, Presans or IdeaConnection run such idea marketplaces and offer crowdsourcing as value proposition. Through hosting idea contests online they act as intermediaries between idea- or innovation-demanding companies (called seekers) and a large-scaled crowd of potential working consumers (called solvers) in the internet. Idea marketplace providers gain revenues as seekers pay for the opportunity of tapping the wisdom of crowds (Surowiecki, 2004), particularly expressed through a large amount and high quality of submitted ideas. Thus, (Spradlin, 2010), CEO of InnoCentive, suggests thinking about their business model as “the ebay of innovation”. Following this description, idea marketplaces are causing changes to the supply side of innovation (Archak and Sundararajan, 2009). Through offering a large scaled base of potential solvers, including various customer experiences, skills, backgrounds and tastes, seeker’s potential benefits are based on reduced matching costs and increased diversity. In other words, seekers don’t have to search for suitable solvers, but pluralities of solvers can select tasks they feel attracted to and submit suggestions unconstrained.

Various success stories of idea contests (e.g. InnoCentive, Dell IdeaStorm, myStarbucksIdea, Cisco iPrize, etc.) also accelerated research. Mainly from a conceptual point of view, information systems (IS) research deals with design patterns of innovation platforms, economics develop models of prize structures or incentive schemes for such contests, whereas sociologists mostly search for intrinsic motivation factors among the solver crowd. Furthermore, the concept of idea contests is not an invention of the short termed past. E.g. research on free and open source software (FOSS) communities or electronic brainstorming tools deals with the parts of the concept for years. Despite this growing amount of research papers, basic rules and tenets of online idea contests have not yet been carved into stone. One major reason for this is, that less empirical studies have been undertaken, dealing with data from real online idea contest. Hence, even though there are various assumptions on influential factors, there is less clarification which factors can be applied to support high and valuable outcomes, i.e. a plurality of precious ideas. At a basic level, it becomes a crucial issue of idea marketplace providers (like InnoCentive) as well as innovation
seeking companies to understand, which external factors cause effects on the outcome. Once understood, in return marketplace operators may use these factors as a driving belt to increase the outcome and finally offer a stronger value proposition. This gave rise to our constitutional research question:

**RQ:** Which (measurable) external factors take effects on the outcome of online idea contests?

To answer this question we start with a literature review, focusing on IS research as well as, economic or sociologic work to provide a basis for hypothesis how idea contests should work from a theoretical point of view (Chapter 2). We restrict the analysis on extrinsic factors, which are either given externally by the nature of an online idea platform or regard definable structural and design elements of online idea contests. Thus, we exclude internal and behavioral factors which are directly dependent and only measurable by interacting with the solvers, e.g. measuring their individual efforts (Lakhani et al., 2006), their networks (Ye and Kishida, 2003, Huberman et al., 2009, Franke and Shah, 2003), their skills, backgrounds and intrinsic motivation level (Lakhani et al. 2005, Shah, 2006) or feedback processes during running idea contests (Nov et al., 2009, Yang et al. 2009). Based on the literature review we develop 4 models, predicting the effects of external factors on the outcome of online idea contests (Chapter 3). We evaluate these models by applying a dataset from practice, derived from Atizo (an idea marketplace provider like InnoCentive). We describe data collection and run descriptive analysis (Chapter 4) as well as OLS regression, using the sample of all closed contests since the first contest in 2008, containing over 25,000 submitted ideas (Chapter 5). Concluding we discuss the validity of our results and provide a glance of implications on idea contest design on idea marketplaces as questions of future research (Chapter 6).

### C.2 Theoretical Background

A plurality of partly overlapping or constitutive theories can be identified dealing with aspects of idea generation processes. Hence, our initial task is to define how crowdsourcing and idea contests relate to other existent theories (2.1) and subsequently to define tenets within these theories, which describe which external factors contribute towards the outcome of online idea contests (2.2).

#### C.2.1 Classification of relevant terms

Within the last decades a plurality of partly overlapping terms has been defined within the field of open innovation. Open innovation can be traced back to the definition of (Chesbrough, 2003) as the use of purposive inflows and outflows of knowledge to accelerate internal innovation, and expand the markets or external use of innovation, respectively. Hence, the open innovation paradigm can be understood as the antithesis of the traditional vertical integration model where internal R&D-activities lead to
internally developed products that are then distributed by the firm (Chesbrough et al., 2006, Laursen and Salter, 2006). Open innovation is set in the field of R&D and thus a part of new product development, which is defined the complete process of bringing a new product or service to market (Lilien et al. 2002, Ulrich and Eppinger, 2004). Furthermore, open innovation can include user innovation (von Hippel, 2005) or customer co-creation (Prahalad and Venkatram, 2000), which both refer to an innovation where users have performed a substantial part of the problem-solving process leading to a solution. A group of customers, which solve a company’s R&D problem, creates a user innovation community, which is defined as a distributed group of individuals focused on solving a general problem and/or developing a new solution, applying computer mediated communication (di Gango and Wasko, 2009). As the importance of customer inputs further increase with the market newness of a product, it becomes more critical to enlarge user innovation communities (Callahan and Lasry, 2004). FOSS Communities are yet another example of a user innovation community in which information, assistance, and innovations is freely shared (Franke and Shah, 2003).

Research from the area of open innovation often references the existence of a swarm- or collective-intelligence as defined by (Lèvy, 1997, Bonabeau, 2001, Bonabeau, 2009), and depicted as the wisdom of crowds by (Surowiecki, 2004). The wisdom of crowds states that the aggregation of information in groups results in decisions which are often better, than made by any single member of the group (Surowiecki, 2004). Various systems or tools which are used to tap the wisdom of crowds and support collaboration- or idea production-processes have been mentioned: (de Sanctis and Gallupe, 1987) defined the term group decision support systems, (Numaker et al., 1991) defined electronic brainstorming tools, (Berg et al., 1996) defined prediction markets, years later (Piller and Walcher, 2006) defined toolkits for idea competitions, and (Back and Wagner, 2008) defined group wisdom support systems. The commonness of such tools is, that they try to transform the phenomena of the wisdom of crowds (or collective intelligence respectively) into a structured process of value creation (Jones et al., 2009). This leads towards the definition of crowdsourcing. The basic concept behind crowdsourcing is that crowds of amateurs can fulfill tasks and gather contributions at least as valuable but less cost-intensive than internal R&D departments (Howe, 2006, Howe, 2008). Similar to open innovation, crowdsourcing has a strong focus on R&D processes, especially on lowering R&D costs or enhancing R&D quality (Howe, 2008, Leimeister et al. 2009). Crowdsourcing is increasingly seen as a strategic model to attract an interested, motivated crowd of individuals capable of providing solutions superior in quality and quantity to those experts can (Girotra et al., 2010, Brabham, 2008). Consequently, crowdsourcing can be described as an process of open innovation harnessing or leveraging the wisdom of crowds. During implementations of a crowdsourcing strategy, often idea contest (or innovation tournaments) are set up on company websites (e.g. Dell Idea Storm, Cisco iPrize, IBM
Innovation Jam, MyStarbucksIdea, Threadless) or on mentioned idea marketplaces (e.g. InnoCentive, PRESANS, TekScout, IdeaConnection, NineSigma, Innovation Exchange, 99 designs, Atizo, etc.). Using idea contests to leverage innovation processes is also not an invention of the short-term past, e.g. (Amabile, 1988) suggested idea contests to manage ideas within organizational structures. Generally, an idea contest is a game in which several agents spend resources in order to win one or more prizes (Moldovanu and Sela, 2001). (Ebner et al., 2008) define online idea contests as “the invitation of a private or public organizer to a general public group to submit contributions to a certain topic within a timeline. An idea-reviewers committee evaluates these contributions and selects the rewarded winner.”

C.2.2 Effects of External Factors on the Outcome of Idea Contests
As the previous chapter shows, idea contests do neither establish an entirely new, nor independent area of research. In fact, previous research on open innovation, electronic brainstorming, FOSS communities, but also research in the fields of economic contest design, behavioral economics or sociology also contribute predictions on the potential effects of external factors within idea contests. A basic tenet of all these research areas is that certain factors operate as incentives which potential solvers perceive and then activate. Hence, the right mixes of incentives are those that appeal to or match the solver’s motivation for participating (Leimeister et al., 2009, Riggs and von Hippel, 1994). Basically extrinsic and intrinsic incentives are distinguished. Accordingly, we say that a solver is intrinsically motivated if he engages in an activity to feel competent and self-determining in relation to the activity. There is no external reward, rewards are internal to the solver and take the form of feelings he has on himself (Deci, 1978, Amabile, 1997). Within our research we do not focus on intrinsic motivation factors. However, we will acknowledge the following:

Research showed, that attempts to measure intrinsic effects within online idea contests has to be based on qualitative research methodology, i.e. mostly interviews or surveys among a crowd of solvers. Significant intrinsic factors include the desire to acquire new skills, the desire to learn, the passion for problem solving and exploration (often at the boundary or outside fields of expertise), enjoyment or the interest in free-sharing of ideas (Lakhani et al. 2006, Lakhani and Wolf, 2005, Franke and Shah, 2003, Shah, 2006). For further qualitative attempts to measure intrinsic factors see (Antikainen and Vääätäjä, 2010, Lakhani et al. 2006, Stenmark, 2002). In the following we focus on our primary interest, the idea marketplace perspective. We will depict several extrinsic factors which may appear on marketplaces of idea contest. We structure the literature search for tenets on external factors by separating the research areas open innovation/ crowdsourcing from FOSS communities/ electronic brainstorming and economics/ sociology.
C.2.2.1 Open Innovation and Crowdsourcing Research

From an IS perspective external factors often are associated or identified as so called “design elements” (Nov et al., 2009). Similar to external factors, design elements are partly given by the circumstances of an online marketplace and can be set or adjusted by the marketplace operator. (Haller et al., 2009) deduce 24 design elements from analyzing 27 online idea contests. Among these design elements they find specificity of tasks, contest period and reward as relevant external and adjustable factors. 9 design elements focus on the idea evaluation process, which is not our focus and others are not adjustable, like the type of contest organizer. (Piller and Walcher, 2006) develop a matrix, by which toolkits for idea competitions can be defined through the task specificity and the required degree of elaboration. (Ebner et al., 2008) build up an online idea contests and define 12 characteristics. (Leimeister et al., 2009) further develop this project and narrow down the list to 6 essential characteristics, including the contest timeline, incentives, problem specification and elaboration. Additionally they define 8 types of incentives, including prizes, profit options and career options. They find that getting aligned with a seeking company (i.e. the seekers brand strength) has a stronger influence than monetary prizes. (Walter and Back, 2010) find similar results in a qualitative study of four crowdsourcing case studies. These findings go along with a plurality of case studies in open innovation, which all find that without direct monetary compensation, a vast number of resources are committed to open innovation (Chesbrough, 2006, Chesbrough and Appleyard, 2007, Nickerson et al. 2009, Brabham, 2008). (von Hippel and von Krogh, 2003) find reasons of this in relative low costs of submitting to online contests. Hence, even low amounts of monetary rewards can have sufficient effects on contributions. In their recent empirical study, (Yang et al., 2009) try to prove this and conduct research on factors which affect the outcome of online idea contests. They define the amount of attracted solvers as dependent variable and find that rewards attract more people as well as longer contest duration or fewer words used to define the contest question. (Huberman and Romero, 2009) define the possibility of free-riding as rationale of participating in crowdsourcing. They find that freely consuming the ideas of others satisfies most solvers. Additionally they find that paid attention by other solvers, e.g. via the existence of an idea rating system can additionally enhance participation. Research of (Antikainen and Väätäjä, 2010) supports this finding. According to their study 2/3 of all solvers contribute due to the pure existence of a reward system, which can contain a solver ranking, e.g. a top ten list, next to basic monetary rewards. (Morgan and Wang, 2010) develop a two step decision tree to design idea contests which is based on the demanded degree of disruptiveness of ideas and the solvers skill profile. Based on these factors they suggest different distribution styles and absolute values of monetary rewards. Their basic suggestion is to use higher and more divergent rewards, the more the skills vary among solvers.
C.2.2.2 Research on FOSS Communities and Electronic Brainstorming

Similar to idea contests, FOSS projects and electronic brainstorming groups are unlikely to be successful unless there is an accompanied community that provides the platform for developers and participants to collaborate (Ye and Kishida, 2003). FOSS Projects include external factors, as programmers often get paid by their employer to participate in order to improve their skills, which is also seen as investment in human capital (Lakhani and Wolf, 2005). (Lerner and Tirole, 2002) find the biggest external motivation factor of FOSS projects in a delayed reward, expressed by enhanced career opportunities of participants. (Nov et al., 2009) separate motivational, structural and tenure factors as causes of user contributions in online communities. Tenure factors especially focus on maturity of the group, measured by years of existence or experience. Within the research area of electronic brainstorming (Michimov and Primois, 2005) find that group productivity and creativity is most dependent on the factor of provided feedback. Additionally they find that social comparison, such as a ranking of suggestions also enhances productivity. In contrast, (Stenmark, 2002) finds that the introduction of rewards lead to poorer, i.e. uncreative, results in electronic brainstorming as a level of competition lowered the users’ ability to be creative. Additionally the limited amounts of time as well as the detailedness of task description are seen as positive external factors.

C.2.2.3 Economic and Sociologic Research

Economists are also interested in rewards as an external factor of idea generation. Absolute values as well as reward structure are central research issues in tournament- and contest-design. In contrast, contest duration and contest declaration aspects are subordinate as tournament models commonly use two participants and one or two rounds of action (Archak and Sundararajan, 2009, Singh and Wittman, 1988, Yang et al., 2009). Even though economists do no longer base their entire argumentation on the assumption of rational behavior, rewards are still used as main external factor and thus, mostly used to model human behavior on marketplaces. But the effects of monetary compensation on performance are no longer seen as monotonic, which means that offering money does not always produce an improvement (Gneezy and Rusticini, 2000). Economic research recently states some law of diminishing utility of rewards on participation in innovation tournaments. The main statement is that relatively high monetary incentives can have perverse effects on performance, which means adding additional incentives will decrease performance from some point (Ariely et al. 2009a). (Ariely et al. 2009b) further show that tasks, which involve quantitative effort only, e.g. pressing a button as often as possible within a definite time, are more likely to benefit from increased rewards, whereas tasks, which require a cognitive component, include a level beyond which further increases of incentives induce negative effects on performance. Sciologist and behavioral economics call this overjustification-effect, depicting paying too much can induce a crowding-out of so
called prosocial behavior and finally lead to poorer quality of results (Frey and Oberholzer-Gee, 1997, Bènabou and Tirole, 2006). This is based on tenets like intrinsic rewards as superior to extrinsic in causing action of individuals (Amabile, 1997), rewards significantly undermining the free-choice of intrinsic motivation (Deci et al., 1999), extrinsic and intrinsic rewards not being compatible in incentive design (Deci, 1978) and finally rewards undermining real interests, ignoring reasons and discouraging risk-taking (Kohn, 1993). Economic research also states that investments of solvers tend to be sunk (Singh and Wittman, 1988). Hence, with a large-scaled group of solvers, each will have relatively small chance of winning, so the winner's investment and hence, the quality of the winning suggestion will tend to be low (Che and Gale, 2003). Furthermore huge amounts of ideas cause higher costs as the real value proposition offered by marketplace operators comes in locating, filtering and communicating what is useful to the seeker (Varian et al., 2004). In other words, large amount of ideas also causes higher costs of extracting valuable from useless ideas. Based on this (Terwiesch and Xi, 2008) state that inefficiency of idea contests is resulting from the solvers' underinvestment and this can be reduced by changing the rewards structure from a fixed-price to a performance-contingent type.

C.3 Hypothesis and Prediction Model

During chapter 2 diverse external factors (signaled by italic presentation) have been deduced from various areas of research. Chapter 3.1. states hypothesis, derived from this literature review, 3.2 explains how variables will be measured and 3.3. finalizes the prediction models.

C.3.1 Hypothesis

- Hypothesis 1: Higher rewards will lead to a higher outcome of online idea contests.
- Hypothesis 2: Longer contest duration will lead to a higher outcome in online idea contests.
- Hypothesis 3: A shorter description of tasks will lead to a higher outcome of online idea contests.
- Hypothesis 4: Highly specific tasks will have a lower outcome of idea contests.
- Hypothesis 5: The suggested type of answering will lead towards different outcome levels of idea contests.
- Hypothesis 6: The maturity of the marketplace will have positive influence on the outcome of online idea contests.
• Hypothesis 7: The brand-strength of seekers will have positive influence on the outcome of online idea contests.

• Hypothesis 8: Idea contest where brand strength of a seeker and rewards are relatively high will generate a high outcome.

• Hypothesis 9: The specificity of tasks will response to the set reward.

• Hypothesis 10: The requested type of answers wills response to the duration of online idea contests.

All derived hypothesis are collected and shown in Figure 1, an illustration of the potential external factors on the outcome of online idea contest.

![Diagram of Hypothesis and Variables](image.png)

**Figure 1 Hypothesis of potential external factors on the outcome of online idea contests**

**C.3.2 Variables Measurement**

We are restricted to measurable outcomes of idea contests. Literature suggests measuring effectiveness of idea contests in 1) number of submitted ideas, 2) the quality of ideas and 3) the rarity of ideas (Conolly et al., 1990). Other suggestions are to measure the amount of attracted solvers (Yang et al., 2009) or solely the quality of the best (winning) idea (Girotra et al., 2010). We decided to exclude rarity of ideas as well as the quality of the best idea as these are mainly subjective factors. Attracted amounts of solvers are also blurry as the stages of participation vary from submitting ideas, towards commenting the ideas of others until only following contests and reading ideas. An additional bias would be that users are allowed to submit various ideas to one contest. Hence, we set submitted ideas as first dependent variable, meaning we count the total amount of submissions to an idea contest. Solvers are allowed to submit several ideas to one idea contest and seekers can already mark submitted ideas as “interesting” while a contest is still running and therewith maybe influence the solvers behavior. Seekers intervened idea contests by marking ideas as interesting 170 times (which states 0.6% of all ideas within our dataset). In a second regression we define
the average quality of submitted ideas to an idea contest as dependent variable. Therefore we use the given rating system within the Atizo community, where each idea can be rated on a 5-point scale from 1, “almost useless idea”, to 5 “idea with options to win the contest”. The rating of ideas is also done by the solver community. Table 1 provides an overview of the dependent variables (amount and quality), the seven independent variables, derived during hypothesis building in chapter 3 and all corresponding methods of measurement.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Method of Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount (dependent)</td>
<td>Total submitted ideas of an idea contest.</td>
</tr>
<tr>
<td>Quality (dependent)</td>
<td>Average quality of all ranked ideas of an idea contest.</td>
</tr>
<tr>
<td>Reward</td>
<td>The amount of money in Swiss Chf, accounted for one or many winning ideas of an idea contest.</td>
</tr>
<tr>
<td>Duration</td>
<td>Days from contest start till contest end of an idea contest.</td>
</tr>
<tr>
<td>Description length</td>
<td>Amount of words used to describe the task of an idea contest.</td>
</tr>
<tr>
<td>Specificity</td>
<td>Subjective measure of required skill level, necessary to submit an idea to an idea contest (scale of 1-5).</td>
</tr>
<tr>
<td>Task Type</td>
<td>Suggested answer types categories: Naming, Designing, Engineering or Business Solution</td>
</tr>
<tr>
<td>Market Maturity</td>
<td>Month counted from September 2008 till launch of an idea contest.</td>
</tr>
<tr>
<td>Brand Strength</td>
<td>Subjective measure of seekers brand-strength (scale 1-5).</td>
</tr>
</tbody>
</table>

Table 2 Dependent and independent variables and methods of measurement

All variables refer to the entity of an online idea contest. The task type is defined by the seekers description of task, especially by mostly given rating criteria, in which certain submission styles are suggested or announced. Brand-strength and specificity of tasks are purely subjective factors. In order to measure the effects of these we defined values of seekers brand-strength and task specificity on the scale of 1 (lowest, worse) to 5 (highest, best). We tried to eliminate potential biases due to subjective determinations of values by applying inter-rater reliability technique, i.e. multiple classifications by 3 research assistants. Additional determination of values was undertaken by Atizo in order to affirm no local specifics have been ignored. Fleiss’ Kappas of 0.94 (brand-strength) and 0.81 (specificity) shows high concordance among all raters’ valuations (Fleiss, 1971).
C.3.3 The Model

We use ordinary least squares (OLS) for multiple regression analysis to determine size and strength of relationships between dependent and independent variables. OLS is commonly used during structural equation models like this. It is also commonly used in cited studies on online idea contests (e.g. see Ariely et al., 2009b, Giora et al. 2010, Lakhani and Wolf, 2005, Yang et al. 2009) what makes results partly comparable. The dataset shows no runaway values, so we do not see the requirement of a logarithmic description (e.g. as in (Yang et al., 2009)) and we found no bias through multicollinearity underneath independent variables. A test of inter-item reliability for independent variables revealed acceptable Cronbach alpha of 0.72 (Aiken and West, 1991). Finally this leads us towards four prediction models. The first two predict effects of basic external factors on quantity and quality of idea contests as the following:

$$\text{Outcome} (\text{Amount, Quality}) = \beta_0 + \beta_1(\text{Reward}) + \beta_2(\text{Duration}) + \beta_3(\text{Description Length}) + \beta_4(\text{Specificity}) + \beta_5(\text{Task Type}) + \beta_6(\text{Market Maturity}) + \beta_7(\text{Brand Strength}) + \xi$$

The second two models enhance the predictions of the basic model by including interaction terms between variables:

$$\text{Outcome} (\text{Amount, Quality}) = \beta_0 + \beta_3(\text{Description Length}) + \beta_6(\text{Market Maturity}) + \beta_8(\text{Reward}^*\text{Specificity}) + \beta_9(\text{Reward}^*\text{Brand-Strength}) + \beta_{10}(\text{Task-Type}^*\text{Duration}) + \xi$$

In both cases the outcome is separated into amount and average quality of ideas. The term $\xi$ covers a random error. Only the extended models include hypothesis 8, 9 and 10, which all state that variables intercommunicate. Hence, we cut out variables in their detached form (as in the basic model) and only use the combinations. The data for the regression was obtained via access to the Atizo database, which included all closed idea contests, including the corresponding rewards, amounts of submissions, average quality of submissions towards an idea contest, start- and ending dates of contests and name of seeking organization. R (Version 2.11.0) was used to calculate the regression models.

C.4 Data Collection and Descriptive Analysis

C.4.1 The Atizo Platform

Atizo was founded June 2008 in Bern, Switzerland. The company defines its business model as the following:

Atizo administers a growing web-community of creative thinkers, who are characterized by their user, consumer and special knowledge. For the mobilization of
this community and yet other innovator teams, Atizo continually develops innovation management tools, which are applied in innovation projects of companies and organizations of all sizes and sectors (Atizo, 2010).

Atizos’ standard idea generation process is the following. Seekers formulate a task description including suggested or required formats of submissions. Additionally seekers decide on rewards, whereas Atizo consults, which reward seems adequate to the specified problem. In the following, the task is posted on the platform, free for all signed-in solvers to submit ideas, comment on other ideas and rates other ideas within a public online brainstorming phase. In a next step seekers select one or more best ideas (which does not have to be in accordance with the average rating from the crowd of solvers) and divide the rewards. Hence, before or during a running contest solvers are not aware, whether there will be one fix price winner or a reward split into the best five ideas, but only the total size of distributed rewards.

C.4.2 Descriptive Statistics of Online Idea Contests on the Atizo Platform
The first idea contest was opened and hosted by Atizo in September of 2008. Since then 74 idea contests have been closed. Contests lasted between 4 and 135 days, but 57.3 days at the average (Std Dev 20.7). 7013 solver accounts have been opened yet, from which 1545 should be counted as active, requiring having posted an idea at least one time every six month. 156.500 Chf of rewards have been distributed yet. This states an average reward of 2200 Chf (Std Dev 1175.7) per idea contest (which is way below the prizes, InnoCentive offers to their solvers). Atizo suggests seekers to use rewards between 500 and 10’000 Chf. Six contests were set up by non-profit-organizations or seekers which requested to “donate an idea”, stating no rewards (0,- Chf). In general, idea contests at Atizo are mainly of a general type. (Hallerstede et al., 2010) analyzed ten idea marketplaces and found that contest on Atizo are rather focused on brainstorming and hence, less expert knowledge and detail is required. However, specificity of tasks also varies among the Atizo platform and of course it is not forbidden for solvers to submit highly sophisticated ideas. Seekers announced their idea contest tasks using 140 words in average (Std Dev 24.8). Thereby the most frequent suggested answer type (and hence, the biggest answer category) was the form of describing a product or strategy with text (35%). 21 Contests (28%) demanded engineering solutions (e.g. a technical proof), 17 (23%) only searched for product or service names and therewith could be answered by one line and 10 Contests (13%) requested ideas on design which includes to submit a drawing in the appendix.

Within all contests 25’730 ideas have been posted, which turns to an average of 347 ideas per contest (Std Dev 164.8). 925, and therewith most ideas underneath all contests, were submitted to a question of suggesting a new name of a student platform, offering 1000 Chf only. Only one contest includes less than 100 submissions, stating 94 ideas as a minimum. Considering different durations of idea contest, between 1.4
and 46.8 ideas were submitted per day and contest, concluding an average of 7.5 (Std Dev 12.1) submitted ideas every day regarding all contests.

In total 83.245 ratings of ideas have been submitted. Not each submitted idea was rated, but only 10’856 (42% of 25’730 totally submitted). The average value of these rated ideas is 1.68 on a scale of 1-5 (Std Dev 0.49). Hence, the dataset for modeling the quality was narrowed down to these 10’856 ideas (still including all 74 idea contests). This states the average value of a rated idea is based on 7.7 submitted ratings (Std Dev 5.3), and the value of average quality of ideas within an idea contest is based on 1’124.9 submitted ratings (Std Dev 432.3).

C.5 Results and Analysis

Using the dataset, described in chapter 4.2, we test 4 prediction models, derived during chapters 2 and 3, Table 3 shows the results for all 4 models. In each case, the basic model applies the basic variables whereas the extended model also includes interactions between variables, which are requested by hypothesis 8, 9 and 10.

Measuring effects on the first dependent variable, the amount of ideas, the basic model shows a significant effect of rewards ($\beta_1$). This supports hypothesis 1. A second, strong effect is given by the market maturity ($\beta_7$), which is significant in each of the cases, the basic and the extended model. Further significant, but less strong, factors are set by duration of contests ($\beta_2$) and the seekers brand-strength ($\beta_6$). This supports hypothesis 2 and 7, which suggest that longer contest periods and strong brands will have positive effects on the amount of ideas. ($\beta_4$), which represents the effects of task specificity, is negative in the basic model, but positive in combination with rewards within the extended model. This suggests that idea contests tend to have less submitted ideas as task complexity rises, but high rewards are able to reverse this effect. Hence we are able to support hypothesis 4 and 9. The results also show that description-length ($\beta_3$) and type of required submission ($\beta_5$) style have no significant effect on the amount of ideas. In other words, no matter how seekers describe a task or what kind of submission style they suggest, solvers will send ideas. This also counts in cases of the extended model ($\beta_{10}$), where the assumption is that answer types interact with contests duration. Hence, we reject hypothesis 3, 5 and 10. Regarding their prediction quality, the basic model is slightly stronger, explaining about 60% of variance on the amount of submitted ideas in online idea contests (compared to 37% in the extended model).

Our second, dependent variable is average quality of rated ideas per contest. A central statement is that in this case different external factors have significant effects than in modeling the amount of submitted ideas. As Table 2 shows, within the basic model rewards ($\beta_1$), duration ($\beta_2$), description-length ($\beta_3$), specificity ($\beta_4$) and brand-strength ($\beta_6$) cause no significant effect on the average quality of rated ideas within a contest. This suggests rejecting hypothesis 1,2,3,4 and 6 in this case.
<table>
<thead>
<tr>
<th>((\beta_0)) Variables</th>
<th>Effects on the average quality of rated ideas per contest</th>
<th>Effects on the amounts of submitted ideas per contest</th>
</tr>
</thead>
<tbody>
<tr>
<td>((\beta_0)) Intercept</td>
<td>-29.31 (15.45)*</td>
<td>-12938 (4770)**</td>
</tr>
<tr>
<td>((\beta_1)) Reward</td>
<td>0.000</td>
<td>0.052 (0.017)**</td>
</tr>
<tr>
<td>((\beta_2)) Duration</td>
<td>0.004 (0.003)</td>
<td>1.712 (1.020)*</td>
</tr>
<tr>
<td>((\beta_3)) Description-Length</td>
<td>-0.001 (0.002)</td>
<td>0.196 (0.758)</td>
</tr>
<tr>
<td>((\beta_4)) Specificity</td>
<td>-0.0566 (0.057)</td>
<td>-29.732 (17.513)*</td>
</tr>
<tr>
<td>((\beta_5)) Answer-Type (Naming)</td>
<td>0.045 (0.342)*</td>
<td>-</td>
</tr>
<tr>
<td>((\beta_6)) Brand-Strength</td>
<td>0.035 (0.038)</td>
<td>20.810 (11.991)*</td>
</tr>
<tr>
<td>((\beta_7)) Market-Maturity</td>
<td>0.009 (0.005)**</td>
<td>3.929 (1.419)**</td>
</tr>
<tr>
<td>((\beta_8)) Reward*Brand-Strength</td>
<td>-</td>
<td>0.000</td>
</tr>
<tr>
<td>((\beta_9)) Reward*Specificity</td>
<td>-</td>
<td>0.342 (0.118)**</td>
</tr>
<tr>
<td>((\beta_{10})) Answer-Type*Duration (Naming)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>((\beta_{11})) Answer-Type*Duration (Designing)</td>
<td>-0.115 (0.011)**</td>
<td>-1.334 (2.606)</td>
</tr>
<tr>
<td>((\beta_{12})) Answer-Type*Duration (Technical Sol.)</td>
<td>-0.014 (0.015)</td>
<td>-2.199 (5.373)</td>
</tr>
<tr>
<td>((\beta_{13})) Answer-Type*Duration (Business Sol.)</td>
<td>-0.02 (0.008)**</td>
<td>0.442 (2.803)</td>
</tr>
<tr>
<td>((\beta_{14})) Answer-Type*Duration (Business Sol.)</td>
<td>-0.018 (0.007)**</td>
<td>-1.334 (2.606)</td>
</tr>
</tbody>
</table>

| N | 74 |
| R² | 0.1712 | 0.6164 | 0.5972 | 0.3695 |

Level of significance: ***p < 0.001; **p < 0.05; * p < 0.1

Table 3 OLS Regression Analysis for Modelling the Amount and Quality in Online Idea Contests

Also the validity of predictions using the basic model is weak, as it only explains about 17% of the variance in idea quality per contest. However, the prediction quality of the extended model is unequivocally better (R² = 0.62). Within the extended model, especially the variables describing the interaction between external factors (β9 and β10) show significant effects. These results suggest that higher rewards only lead towards better quality of ideas, if they are offered in highly specific contests. As effects of answer-types were slightly positive in the basic model, this effect is partially turned around
through the extended model. The regression suggests that in all answer-types (except designing) a longer duration leads towards a lower average quality of ideas. Therewith we see support for hypothesis 10. Finally, again the market-maturity ($\beta7$) has a strong positive effect on the outcome.

**C.6 Conclusions and Limitations**

Even though IS research has been concerned with open innovation (e.g. Chesbrough, 2003, Laursen and Salter, 2006, Thomke and von Hippel, 2008) or community design (e.g. di Gangi and Wasko, 2009, Ebner et al., 2008, Nov et al., 2009, Stenmark, 2002) for years, the rules for online idea marketplaces have not yet been carved into stone. Platforms (e.g. Hallerstede et al., 2010), design elements (e.g. Haller et al. 2009, Leimeister et al., 2009, Piller and Walcher, 2006) as well as seeker- (e.g. Gassmann and Enkel, 2004, Lilien et al., 2002) or solver-behavior (e.g. Franke and Shah, 2003, Lakhani et al., 2006, Lakhani and Wolf, 2005) has been studied, but yet little is known about the actual effects such factors might cause. Of cause, this is also partly due to the fact that such marketplaces are young phenomena and thus research lacks of a plurality of datasets. To draw near the rules of these marketplaces we searched for effects of external factors on the outcome of online idea contests. Figure 2 depicts our significant findings, separated by effects on quality and amount of submitted ideas in contests.

**Figure 2 Effects of external Factors on the average quality of rated ideas in online idea contests**

Limitations in space do not allow discussing each result in detail, so we concentrate on major findings. Given the circumstances of an open platform, with less information asymmetries and the asynchronous suggestion process, theory from various research areas is not sufficient to explain effects on online idea contests. Monetary incentives are widely used, but (concerning our dataset) only have an effect on the amount of submitted ideas, not on their quality. Idea marketplace providers as well as seekers should be careful with setting high rewards, as they may cause large amounts of solver submissions, often of lower quality. This could lead toward higher transaction costs (especially for unraveling valuable from useless ideas) and hence, lower the ROI of crowdsourcing. Furthermore the effects of some basic design elements from IS
research show no effects on the outcome at all. This includes the task description or partly also the contest duration.

In contrast Figure 3 shows the existence of external factors, which can be equal driving forces of the market, but cannot directly be influenced. E.g. when it comes to pure participation, i.e. submitting ideas, solvers are also attracted by the pure existence of a seekers’ corresponding brand. Hence, seekers with strong brands may get informed by idea marketplace providers, that it does not need a high reward. In general we find that amount and quality of ideas are affected by different kinds of external factors. For idea marketplace providers this is important to understand. Seekers, trying to incent higher quality of ideas by higher rewards may even create a conflict of goals.

However, there are also limitations to this paper. One is given by a strong finding. We detect a kind of inflationary effect on amount and quality of ideas by the external factor of market-maturity. This simply shows that the (perceived) outcome also rises due to the age (i.e. growth) of the idea marketplace. This represents the fact, that more and more potential solvers signed in over time. Future modeling may consider using market growth as a underlying effect. Of course one can also claim that our sample size (74 contests) is rather small. In contrast, variables have been chosen wisely and there is no visible tendency towards multi-collinearity among variables. Another concern is that we were rather selective in choosing the external factors. As chapter 2 suggests, there is a plurality of further factors that may cause effects on the outcome of idea contests. As mentioned, our initial step was to exclude all factors which are directly dependent on solver behavior. Hence, we see a next step in focusing on these factors, especially on effects of the crowd’s structure. Social network analysis methodology might be an approach here. To finalize towards a system-theoretical approach research also have to be conducted from a seeker based view. For now, we cannot claim that these external factors can or should be used as driving belts to
leverage the outcome, but rather to help all participants better understanding market conditions and being aware of the possible external effects.

C.7 References


Ebner, W., Leimeister, J.M., Bretschneider, U. & Krcmar, H. “Leveraging the Wisdom of Crowds: Designing an IT-supported Ideas Competition for an ERP Software


**Contribution D: Controlling von Kundenmeinungen durch Text Mining: Entwicklung einer Lösung zur Analyse von In-App-Feedback**

<table>
<thead>
<tr>
<th>Title</th>
<th>Controlling von Kundenmeinungen durch Text Mining: Entwicklung einer Lösung zur Analyse von In-App-Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authors &amp; Affiliation</td>
<td>Thomas P. Walter, Institute of Information Management, University of St. Gallen, Mueller-Friedberg-Strasse 8, 9000 St. Gallen, Switzerland, <a href="mailto:thomas.walter@unisg.ch">thomas.walter@unisg.ch</a></td>
</tr>
<tr>
<td>Publication Type</td>
<td>Journal Paper</td>
</tr>
<tr>
<td>Year</td>
<td>2012</td>
</tr>
<tr>
<td>Status</td>
<td>Published</td>
</tr>
</tbody>
</table>

**Abstract:***


**Abstract (English):***

Today enterprises offer their customers various channels to express their opinion, to provide ideas or concrete suggestions for improvement as well as to ask for help or information. One of such channels is the frequent invitation to provide feedback within so called enterprise-apps. As customers use this opportunity, enterprises face the problem of a large number of customer opinions, foremost given as short and unstructured text data. This paper deals with in-app feedback functionality that is implemented within four apps of an world leading airline. We develop a solution which supports customer value controlling. In particular, we apply methods of text
mining to automatically structure customer opinions, determine the sentiment of feedback and support the initiation of countermeasures towards negative opinion making.

**Keywords:**

Apps, Customer Feedback, Customer Value Controlling, Text Mining

**Danksagung:**

Das eingesetzte Excel Tool wurde in Zusammenarbeit mit der Masterarbeit von Herrn Marcel Hungerbühler entwickelt. Die Masterarbeit hatte den rein technischen Fokus des Toolaufbaus. Der Autor dankt Herrn Hungerbühler für die Bereitstellung der Codes für die R-Software sowie der Codes für Excel-VBA.

**D.1 Kundenfeedback in Apps**


**D.1.1 Problemstellung: Kundenverhalten bei App-Feedback-Funktionalität**

Der vorliegende Beitrag befasst sich mit der Feedback-Funktionalität in vier Apps einer weltweit führenden Fluggesellschaft. Diese ist als first mover bereits seit 2009 mit einer App für das Planen, Einsehen und Buchen von Flügen, sowie mit einer

D.1.2 Anforderungen an das Controlling
besteht seitens des Controllings hier Bedarf an technischer Unterstützung des Feedback-Controlling-Prozesses.


D.2 Grundlagen des Text Mining

In der wissenschaftlichen Literatur hat sich für den Begriff Text Mining noch keine allgemeingültige Definition durchgesetzt. Dennoch existieren Gemeinsamkeiten zwischen einzelnen Definitionen welche in der folgenden Auflistung dargestellt sind. Im Folgenden vgl. (Feldman et al., 2007, S. 1; Franke et al., 2003, S. 1; Heyer et al., 2008, S. 3; oder Miller, 2005, S. 104).

- Die Datenbasis liegt in Form von unstrukturierten Textdaten vor.
- Ziel ist die Erkennung und Identifizierung von Mustern in Dokumentenkollektionen.
- Es finden verschiedene Text Mining Analysetechniken Anwendung.

Text Mining gliedert sich für gewöhnlich in die Phasen der Text Vorverarbeitung (das sogenannte Pre-Processing), der Anwendung eigentlicher Tex Mining Analysetechniken und der Aufbereitung und Darstellung der Ergebnisse (Visualisierung). Abbildung 1 veranschaulicht diese drei Phasen und nennt Beispiele der darin zum Einsatz kommenden Standard-Techniken. Im Anschluss werden diese drei Phasen detaillierter erläutert.
Abbildung 1 Phasen des Text Mining und beispielhafte Techniken, eigene Darstellung auf Basis von (Feldmann und Sanger 2007, S. 1 sowie Miller 2005, S. 104)

D.2.1 Pre-Processing


Die „TDM“ stellt den Inhalt eines Textes in numerischen Werten dar und bildet somit den Ausgangspunkt für Text Mining Analysetechniken. Jede Zeile einer TDM entspricht einem Term (oder Wort) und jede Spalte einem Dokument aus einer Dokumentenkollektion. Die Gewichtung der einzelnen Terme in der TDM können
binär (darin enthalten oder nicht), per normalisierter Termfrequenz (tf) oder per
inverser Dokumentenfrequenz (tf-idf, um sowohl die häufigsten als auch seltene
Begriffe aus einer Dokumentenkollektion hervorzuheben) vorgenommen werden

D.2.2 Text Mining Analysetechniken

Text Mining Analysetechniken sind Verfahren, welche zur Gewinnung von neuen
Erkenntnissen sowie von interessanten Mustern aus einer Dokumentenkollektion
dienen. Grundlegende Text Mining Analysetechniken sind die Häufigkeitsanalyse,
Textkategorisierung, Spracherkennung, Clustering und die Sentimentanalyse.

Die „Häufigkeitsanalyse“ ist eine der einfachsten Analysetechniken des Text Mining,
bei der die identischen Begriffe in der Dokumentenkollektion auf der Basis der TDM
summiert werden. Die Häufigkeitsanalyse kann auf der Basis der verschiedenen
Gewichtungsarten der TDM durchgeführt werden. Eine der meist verbreiteten
Analysetechniken des Text Mining ist die „Kategorisierung“ von Texten. Die
Hauptaufgabe besteht in der Zuweisung von vordefinierten Kategorien zu neuen
Texten (Dokumenten). Der Prozess verläuft hierbei nicht vollständig automatisch,
sondern beinhaltet zunächst einen überwachten Lernprozess, bei dem manuell eine
bestimmte Anzahl von Dokumenten vordefinierten Kategorien zugeordnet wird. Aus
diesen Trainingskorpus erlernt der verwendete Algorithmus spezifische Muster für
die Kategorien und weist im Folgenden jeden neuen Text nach erlerntem
Schema der vordefinierten Kategorien zu (Weiss et al., S.39). Im Unterschied werden
beim „Clustering“ die Cluster aus der inhaltlichen Struktur der zugrunde gelegten
Textkollektion automatisch abgeleitet. Im Clustering wird vorwiegend zwischen
hierarchischem und nicht-hierarchischem (flachem) Clustering unterschieden. Meist
wird nicht-hierarchisches Clustering, vor allem der k-means Algorithmus angewandt
(Heyer et al., 2008, S. 199). Die „Spracherkennung“ dient zur automatischen
Klassifizierung der Sprache eines Textes. Text Mining Software beinhaltet hierfür in
der Regel eine auf der N-Gramm-Klassifizierung basierende Bibliothek. Ein N-Gramm
ist die Aufteilung eines einzelnen Wortes in verschiedene lange Buchstabenfolgen,
welche jedoch kontinuierlich sein müssen. Aus dem Wort „Haus“ resultiert
beispielsweise folgendes Bigramm (N=2): _H, HA, AU, US, S_. Durch die Häufigkeit
einzelner N-Gramme in einem Text und der Berechnung der Distanz zwischen den
einzelnen Klassierungen der N-Gramme kann sehr verlässlich das jeweilige
Sprachprofil definiert werden (Cavnar und Trenkle, 1994, S. 165). Die „Sentiment
Analyse“ ist das konzeptive Untersuchen ausgedrückter Meinungen, Sentiments und
Emotionen in Textkollektionen (Feldmann und Sanger 2007, S.82). Es werden die
Forschungsschwerpunkte der Sentiment Kategorisierung (positiv, negativ, neutral) der
eigenschaftsbasierte Sentiment Analyse (positiv, negativ oder neutral bezüglich einer
Eigenschaft) und der vergleichsbasierten Sentiment Analyse (positiv, negativ oder
neutral bezüglich eines Objektes, z.B. einem Konkurrenzunternehmen).
D.2.3 Visualisierung

D.3 Entwicklung einer Lösung für das Feedback-Controlling
Mehrere hundert Kundenmeinungen und –Wünsche, sowie direkte Anfragen und Forderungen pro Woche stellen Unternehmen, welche eine eigene App mit Feedbackoption herausgeben, vor Kapazitätsprobleme. Eine manuelle Auszählung und Auswertung des gesamten Feedbacks ist aus Zeit- und Kostengründen nicht möglich. Die Hauptanforderungen an die technische Unterstützung für das Controlling, welche aus Interviews mit der Fluglinie abgeleitet wurden, sind daher:

- Übersicht über den Status Quo des gesamten Feedbacks.
- Mehrsprachigkeit.
- (Teil)automatisierte Identifikation von Problembereichen.
- Filtern relevanter Feedbacks für Fachbereiche.
- (Teil)automatisierte Weiterleitung an die Fachbereiche.
- Bearbeitung der Feedbacks direkt in der Lösung.
- Einfache Nutzeroberfläche in bekannter Arbeitsumgebung.


Abbildung 3 Screenshot der Dashboard Funktion Sentiment Analyse

Abbildung 4 Screenshot der Funktion Feedback bearbeiten
Der Benutzer navigiert mit den Interaktionselementen (Pfeile) durch die selektierten Feedbacks und kann somit relevante Feedbacks direkt bearbeiten, bzw. einzeln oder gebündelt an die relevanten Fachabteilungen weiterleiten. Im abgebildeten Fall wird das Feedback mittels Sentimentanalyse vorselektiert. Aussagen wie „nicht korrekt“, „verspätete“ oder „längst vorbei“ klassifizieren das Feedback als negativ. Der Mitarbeiter kann nun über die Option diese Information an die betroffene Abteilung (in Verantwortlichen über die Informationen zum Flugstatus) weiterleiten und dem Kunden eine dankende Antwort oder eine Erklärung zukommen lassen. Die Signale werden von Eingangs beschriebenem Rauschen getrennt womit die Entscheidung unterstützt wird, auf welche Kundenmeinungen direkt reagiert werden sollte.

D.4 Ergebnisse

D.4.1 Evaluation der Genauigkeit der Lösung

<table>
<thead>
<tr>
<th>Analysemethode</th>
<th>Cohen Kappa</th>
<th>Feedback-Länge</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt; 20 Zeichen</td>
<td>20 – 200 Zeichen</td>
</tr>
<tr>
<td>Spracherkennung</td>
<td>1.0</td>
<td>47,6%</td>
</tr>
<tr>
<td>Sentimentanalyse</td>
<td>0.94</td>
<td>63,9%</td>
</tr>
<tr>
<td>Kategorisierung</td>
<td>0.77</td>
<td>37,4%</td>
</tr>
</tbody>
</table>

N = 1269 Feedbacks

Die automatische Spracherkennung erweist sich bei langem Feedback (> 200 Zeichen) zu 96,3% als richtig, jedoch nimmt die Genauigkeit bei kürzeren Feedbacks deutlich ab. So hat die Spracherkennung für Feedbacks mit geringer Zeichenlänge (< 20 Zeichen, z.B. „Top App, weiter so!“) nur noch eine Genauigkeit von 47,6%. Die Sentimentanalyse weist ebenfalls bei den ausführlicheren Feedbacks die höchste Genauigkeit von 68.2% auf. Die Ungenauigkeit resultiert vor allem aus

D.4.2 Bedeutung für das Controlling
D.5 Fazit
Gezieltes Monitoring des In-App Feedbacks erhöht den Servicelevel des ganzen Unternehmens, denn sämtliche Feedbacks können ressourcenschonend ohne weiteren Personaleinsatz verarbeitet werden. Unter den Funktionen innerhalb der Anwendung wird die Vorselektion nach bestimmten Kriterien von den Mitarbeitern als am wertvollsten eingestuft. Somit gelingt nach dem Überblick über mögliche Problembereiche (Selektion negatives Sentiment) die direkte Bearbeitung aller Feedbacks zu einer konkreten Kategorie (z.B. alle Nutzer des Blackberry Betriebssystems). Heute noch nicht möglich sind vor allem die vollautomatisierte Analyse des Sentiments, bzw. das Clustering, was den grössten Anreiz für zukünftige Forschung im Bereich darstellt.


D.6 Literatur


Contribution E: A Text Mining Approach to Evaluate Submissions to Crowdsourcing Contests

<table>
<thead>
<tr>
<th>Title</th>
<th>A Text Mining Approach to Evaluate Submissions to Crowdsourcing Contests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authors &amp; Affiliation</td>
<td>Thomas P. Walter, Andrea Back, Institute of Information Management, University of St. Gallen, Mueller-Friedberg-Strasse 8, 9000 St. Gallen, Switzerland, {thomas.walter, andrea.back}@unisg.ch</td>
</tr>
<tr>
<td>Publication Type</td>
<td>Conference Paper</td>
</tr>
<tr>
<td>Year</td>
<td>2013</td>
</tr>
<tr>
<td>Status</td>
<td>Published</td>
</tr>
<tr>
<td>Full Citation</td>
<td>Thomas P Walter, Andrea Back, A Text Mining Approach to Evaluate Submissions to Crowdsourcing Contests, in 46th Hawaii International Conference on System Sciences (HICSS), Wailea, USA, January 2013.</td>
</tr>
</tbody>
</table>

Abstract:

This survey deals with the problem of evaluating the submissions to crowdsourcing websites on which data is increasing rapidly in both volume and complexity. Usually expert committees are installed to rate submissions, select winners and adjust monetary rewards. Thus, with an increasing number of submissions, this process is getting more complex, time-consuming and hence expensive. In this paper we suggest following text mining methodology, foremost similarity measurements and clustering algorithms, to evaluate the quality of submissions to crowdsourcing contests semi-automatically. We evaluate our approach by comparing text mining based measurement of more than 40’000 submissions with the real-world decisions made by expert committees using Precision and Recall together with F1-score.

Keywords: Crowdsourcing, Text Mining, Clustering, Decision Support Systems,

E.1 Introduction

In 2006, Thomas Davenport argued in an article in Harvard Business Review that the latest strategic weapon for companies is analytical decision making, providing examples of companies that have used analytics to better understand their customers and optimize extended supply chains to maximize their return on investment while providing the best customer service [1]. A large component of this understanding comes from analyzing the vast amount of data that a company collects. The cost
storing and processing data has decreased dramatically in the past, and, as a result, the amount of data stored in electronic form has grown at an explosive rate [2].

As mentioned in the call for paper of this minitrack, social media, encompassing a range of web sites such as blogs, microblogs, wikis, forums or social networks generate tremendous volumes of numerical and textual data that can be mined and analyzed for both research and commercial purposes.

This paper deals with the problem of analyzing data from crowdsourcing websites, a form of social media, on which data is increasing rapidly in both volume and complexity. Crowdsourcing websites, which often claim to tap a collective intelligence [3], [4] or the so-called wisdom of crowds [5], have attracted worldwide attention of both, practice and the scientific community. In 2006, Jeff Howe defined crowdsourcing as the new pool of cheap labor: Everyday people using their spare cycles to create content, solve problems, even do corporate R&D, mostly by using crowdsourcing websites [6]. Whereas the basic idea behind the concept of crowdsourcing is rather clear, so far, neither crowdsourcing platforms nor research succeeded in submitting evidence which methods of measurement can or should be applied to analyze and evaluate submissions towards crowdsourcing websites. On the other hand, the return on investment of crowdsourcing is questioned by firms. For instance, firms, in quest of innovative product solutions via crowdsourcing websites, often obtain up to 1000 submissions, an amount that can be similar to 1000 pages of plain text data. Usually expert committees are installed to rate submissions, select winners and adjust monetary rewards. However, they often are unable to cope with this sheer quantity and complexity of data. As a consequence, firms are running the risk of missing the benefits of crowdsourcing.

In this paper we suggest to follow a text mining approach to address the given problem. Text mining is the semi-automated process of extracting patterns (useful information and knowledge) from large amounts of unstructured data sources [2]. Text mining works by transposing words and phrases in unstructured data, such as submissions to crowdsourcing websites, into numerical values which can then be linked with structured data in a database and analyzed with data mining techniques [7], [8]. Our goal is to provide decision support to the expert committees’ process of analyzing and evaluating submissions to crowdsourcing websites. As it is a longstanding dream of the community to have algorithms that are capable of automatically reading and obtaining knowledge from text our initial research question is stated as following:

*RQ1: How can text mining methodology be applied to support the submission evaluation process on crowdsourcing websites by suggesting most innovative solutions?*
For this purpose we provide theoretical background on crowdsourcing websites in general and current methods of crowdsourcing evaluation in particular during chapter 2. Furthermore we conduct a brief literature review on papers which deal with text mining approaches in crowdsourcing evaluation. Chapter 3 focuses on the development of the text mining approach itself. This includes the description of the applied text mining methods as well as the dataset we use to test our approach. We exploit platform data from a real crowdsourcing website. This data includes over 100 finished crowdsourcing contests together with all raw text data given by over 40’000 submissions. Furthermore we make use of real-world expert committees’ decisions about those submissions, foremost which are most valid to seeking companies and hence, are rewarded. This enables us to state our second research question as following:

\[ RQ2: \text{To what extent can a text mining based evaluation of submissions to crowdsourcing websites reproduce the results of expert committees in regards to selecting most innovative submissions?} \]

We present our answer to RQ1 during Chapter 3 and results to RQ2 in Chapter 4. Analysis of the data was performed using accuracy measures from the field of Information Retrieval, Precision, Recall and F1-score. This enables us to compare results from the manual expert committees’ decisions with the semi-automated, text mining based evaluation approaches. Chapter 5 aims on drawing conclusions from this survey, including managerial and theoretical impact as well as current limitations and an outlook to further studies.

### E.2 Background

The increasing popularity of open innovation approaches [9] in practice has led to the rise of various literature streams within the area of crowdsourcing. Following, we will focus on three central aspects: how crowdsourcing success currently is defined, how submissions from the crowd are evaluated on crowdsourcing websites and to what extent the evaluation is already supported by text mining approaches.

#### E.2.1 Success Patterns of Crowdsourcing

Defining success patterns of crowdsourcing opens a two-sided discussion. On the one hand, various studies find positive effects of monetary rewards on quantity aspects of crowdsourcing, foremost the amount of attracted solvers or the amount of submissions [10–15]. In general firms can benefit from larger crowds because they obtain a more diverse set of solutions [3], [5], [16], [17], which mitigates and sometimes outweighs the effect of the crowds’ underinvestment [18], [19]. Accordingly research states that it requires a large amount and variety of submissions to achieve a high quality best idea [19–22].
On the other hand, economists state that with a large-scaled crowd, each member will have relatively small chance of winning, so the winner's investment and hence, the quality of the winning submission will tend to be low [23], [24]. Furthermore large amounts of submissions slow down the evaluation process due to the necessity of filtering signals from noise. Finally, to the best of our knowledge, there is only little empirical evidence on what drives the quality of crowdsourcing. [25] find that highly connected crowds tend to produce lower quality, [21] find quality to be dependent on adequate crowd coordination techniques, [26] finds individual quality to be positively related to current effort, but negatively related to past success within crowdsourcing, [27] find that crowd performance rises after they recognize being above average, [28] find that, in comparison to experts, on average crowd submissions score higher in novelty and customer benefit, but lower in feasibility and [29] find that whether a task was framed as meaningful does not induce greater or higher quality output.

E.2.2 Measuring the Quality of Crowdsourcing
Next to these findings on general success patterns, research is unclear about how to measure and define the quality of crowdsourcing. As in [15] the size of the attracted crowd is taken as indirect measurement of quality, [14] use the scale of every submission on a five-star rating, [29] take the information whether a task was completed as a measurement and [26] uses the information whether a submission was eventually implemented as primary dependent measure of quality. Also qualitative approaches can be found. [25] use data from external experts to measure quality and [28] take the evaluation from independent executives to compare crowd and expert submissions.

In contrast, [30] find that all simple rating mechanisms, such as thumbs up/down or 5-star ratings are not sufficient to measure the quality of submissions and suggest a multi-attribute scaling including ratings from both, independent expert committees and crowds. In the context of crowd-generated product ideas, [31] aggregate literature and define that idea quality consists of four distinct dimensions: novelty, feasibility, strategic relevance and elaboration. Novelty typically is defined as something being unique, rare or not been expressed before [32]. Another attribute of novelty is the relatedness among submissions [33–35]. This refers to a revolutionary submissions character of being radical and not related to others. Closely related to novelty is originality. Originality of submissions can be defined by their ability to surprise, imaginairiness or degree of unexpectedness [36]. Hence, following this Schumpeterian definition of innovation, many researchers see novelty and originality as the most important facet of creativity [30], [34], [36] and hence, as most suitable measurement of crowdsourcing quality.

E.2.3 Using Text Mining to Measure the Quality of Crowdsourcing
Literature applying text mining methodology is manifold and spread over diverse research fields. For instance, text mining has become an appreciated research
methodology different research areas, from patent analysis [37] towards biology [38]. However, although crowdsourcing websites are generating tremendous volumes of numerical and textual data, a brief, but specific literature review of applied text mining methodology on crowdsourcing or the related area of collective intelligence does not provide a plurality of papers. Therefore, we scan an IS-specific database (The Association of Information Systems electronic Library, AISel) using the search terms “text mining” and major topics crowdsourcing and collective intelligence combined by a logical AND. Findings are diffuse and the coverage of crowdsourcing websites can be described as rather vague. [39] apply four commonly used text classification algorithms and propose a text classification framework for finding helpful user-generated contents in online knowledge-sharing communities. [40] present and evaluate different manual, semi-automatic, and automatic text analysis methods for summarizing transcripts transforming tacit knowledge into explicit form and to substantially reduce the time required to perform this transformation. [41] run text mining methodology on user opinions, expressed via twitter to analyze the appearance of a collective intelligence. [42] develop a taxonomy for combining text and data mining. [43] use text mining to analyze different genres of spam and [44] apply text mining to depict crime networks. Table 1 summarizes the provided background literature.

The final row within Table 1 also represents the research gap we address with our survey. Text mining methodology is used to analyze various aspects of online communities, but to the best of our knowledge not yet to evaluate submissions to crowdsourcing websites.

<table>
<thead>
<tr>
<th>Research Stream</th>
<th>Representative Literature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Defining success patterns of crowdsourcing</td>
<td>[3], [5], [9–16], [18]</td>
</tr>
<tr>
<td>Measuring quality aspects of crowdsourcing</td>
<td>[14], [15], [21], [25], [26], [28]</td>
</tr>
<tr>
<td>Analyzing and defining metrics to measure the</td>
<td>[30–36]</td>
</tr>
<tr>
<td>quality of submissions to crowdsourcing websites</td>
<td></td>
</tr>
<tr>
<td>Applying text mining on open online communities</td>
<td>[39–44]</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2 Summary of background literature.

**E.3 Methodology**

This study exploits text mining to analyze a real-world data sample from an international crowdsourcing website. For the semi-automated step of evaluating the quality of submissions by text mining, clustering is implemented and compared to real-world results, that is to say expert committee decisions. The background literature sets the focus on two central points which describe a current research gap.

- A submission that can be described as representative or average will seldom be able to convince a problem seeking firm as it typically is often not the
richest in information. Under the given circumstances of crowdsourcing websites, it is more useful to select submissions that offer an interesting, unusual or particularly revealing set of circumstances, submissions which are outstanding.

- Yet, text mining is not applied to fulfill a semi-automated selection of submissions to crowdsourcing websites.

During this chapter we will describe our approach to fill this gap. For text mining procedures foremost Provalis Researchs QDA Miner, including the extension package Wordstat is used and for the statistical analysis R is used.

**E.3.1 Dataset and descriptive Statistics**

To apply a text mining based approach of analyzing crowd submissions we make use of crowdsourcing website data. The website was launched in 2008, currently has over 7,000 active members, called solvers. Since its launch, 112 crowdsourcing contests have been closed. In average a contest is open for two month. Firms (called seekers) use the website to state a problem or task and crowds participate by logging in and submitting ideas or concepts to contests. External incentives of participation are monetary rewards and a community ranking of solvers, which is also based on earned total rewards. The quartile of rewards is US$ 92 to 350 per selected (winning, rewarded) submission, but the reward structure is defined by expert committees after a contest has been closed. Hence, the quartile of contest reward-budgets is set by US$ 2’532 to 4’211 which is used to price selected winning submissions per contest. On average 379 ideas or concepts are submitted per contest. As the average length of a submission is 25 words of plain text (which equals 7 sentences or 115 characters), this makes our raw data 42’448 submissions (or 3.65 Mio words). Table 2 summarizes the metrics of the website data.

<table>
<thead>
<tr>
<th>Unit</th>
<th>Total</th>
<th>Avg.</th>
<th>Std.dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crowd Submissions</td>
<td>42’448</td>
<td>379.0</td>
<td>87.41</td>
</tr>
<tr>
<td>- by sentences</td>
<td>154’110</td>
<td>2’653.25</td>
<td>559.12</td>
</tr>
<tr>
<td>- by words</td>
<td>1’078’771</td>
<td>9’631.89</td>
<td>2’108.4</td>
</tr>
</tbody>
</table>

Table 3 Metrics of submissions to the crowdsourcing website.

To make this numbers more feasible, one could say that per average crowdsourcing contest text, twice the length of this conference paper, including 379 more or less outstanding ideas or concepts is submitted, and has to be evaluated and rewarded by expert committees. Contests concern different areas such as product ideas, marketing concepts or technical solutions and are demanded by firms operating in various industries. The following text may serve as an example of a representative crowdsourcing contest. It is an excerpt from a task, provided by a global player in the sports apparel industry, offering a total of US$ 6’000 for the most innovative submissions:
“How can the clothing of the future better regulate the sportsman's body temperature? Athletes at the Olympic Games efforts were challenged by the high temperatures. Athletes produce heat but only around 20% of it is utilized as energy, around 80% literally becomes "hot air", if the body cannot get rid of this heat, and this can lead to cramps and even heat strokes. So that our body does not overheat, it uses four preventive measures: sweating (evaporation), fanning oneself (convection, ventilation), channeling extra heat (conduction), radiates heat. Other examples exist, albeit lavish ones: ice-vests to cool down before competition, integrated ventilation systems in clothing (Air Force pilots). How can we develop a simple piece of clothing, which utilizes the four mentioned mechanisms to prevent athletes from overheating? The Evaluation Criteria are a) A Product that has not yet been developed (originality) and b) Quantifiable temperature reduction (effectiveness).”

E.3.2 Pre-Processing Crowdsourcing website data

The data was processed following a standard text mining procedure, e.g. as in [7]. The first major step of text mining is pre-processing. After extracting all raw text data (the so called corpus) from the website using simple MySQL statements, the raw data is imported to the QDA Miner software. Figure 1 illustrates this by using the given example of the sports apparel contest.

![Figure 1 Typical answer to a crowdsourcing contest, shown as raw text data within the QDA Miner software.](image)

This particular contest has one of the lowest amounts of submissions (98, depicted as cases in Figure 1). Nevertheless, the crowd submitted manifold types of answers to
this contest, some from a technical focus, some from a rather simple minded focus. To be able to analyze this data with traditional data mining techniques, text mining works by transposing words and phrases into numerical values which can then be linked with structured data in a database [7], [8].

Hence as a next step we run pre-processing steps, to be exact stemming, stop-word cleaning and tokenization. Stemming is the process for reducing inflected words to their stem, base or root form. For instance, stemming algorithms like [46] are used to delete suffixes. As a result a stemming algorithm would reduce words like computer, computing, and compute to their stem, which is “comput”. Stop-word cleaning usually is partly a manual process. An algorithm searches text by a predefined list of so called stop-words and deletes them from the text. Most common, short function words, such as “the, is, at, but, which and on” are set onto stop-word lists [7]. Tokenization is the process of breaking a stream of text up into words, phrases or symbols and Part-of-speech tagging is the process of marking up a word in a text as corresponding to a particular part of speech (to syntax) based on both its definition, as well as its context.

The term document matrix (TDM) is the final result of pre-processing. A TDM describes the frequency of terms which occur in a collection of text. In a TDM, rows correspond to documents (D) in the collection and columns correspond to terms (T). In our case documents are represented by submissions (called cases within Figure 1) to a specific crowdsourcing contest and terms are represented by words used within those submissions. Weighting of terms can be calculated binary (e.g. a certain expression is included in a collection), normalized (term frequency, tf) or by inverse term frequencies (tf-idf), which means overweighting less used terms within an collection purposely.

Following the background literature we use both, normalized (tf) and inverse term frequencies (tf-idf), to be able to compare results afterwards. Hence, we calculate two TDMs for each of the 112 crowdsourcing contests. In the enlightened case, the sports apparel contest, the TDM contains 194 x 89 data fields, stating 194 different words used at least in one of 98 submitted cases. However the largest TDM within our dataset is given by a 242 x 957 matrix, stating 242 different words within 957 submissions to one single contest.

E.3.3 Clustering Submissions to Crowdsourcing Contests
Text clustering is the application of certain algorithms to automatically detect patterns within a TDM. Clustering is used to explore the similarity between documents. Often so called non-hierarchical (or centroid-based) clustering is applied, foremost the k-means algorithm [47–49]. In this survey k-means aims to partition text-documents (submissions) into k clusters in which each observation belongs to the cluster with the nearest mean. At the bottom k-means is based on principal component analysis or minimalizing least squares [50]. However, determining the k number of clusters in a
data set is a frequent problem in data clustering, and is a distinct issue from the process of actually solving the clustering problem. In text mining, a frequently used method to determine the number of clusters can be estimated by the following formula \((D \times T) / t\) where \(t\) is defined as the amount of non-zero entries in the entire TDM [51].

Two broad types of clustering can be applied: first- and second-order clustering. First order clustering will group together words appearing in the same document and second order clustering will consider that two words are close to each other, not necessarily because they co-occur in the same document, but because they both occur in similar environments. One of the benefits of this clustering method is its ability to group words, and submissions therewith that are synonyms or alternate forms of the same word. For example, while TUMOR and TUMOUR will seldom or never occur together in the same document, second order clustering may find them to be related because they both co-occur with words like BRAIN or CANCER [52]. As a consequence we apply second order clustering.

![Figure 2 Excerpt of similarity index matrix showing cosine coefficients between submissions to a crowdsourcing contest.](image)

Ultimately, clustering legitimizes a statement about the distance between all submissions within a contest. When the clustering is set to be performed on documents (submissions), a distance matrix used for clustering and multidimensional scaling consists of cosine coefficients computed on the relative term frequencies (in tf or tf-
idf) of the various words within documents. The more similar two submissions will be in terms of the distribution of words, the higher will be this coefficient [7], [53]. Figure 2 illustrates the similarities within the given example of the sport apparel contest.

Hence, as a final step we have to select, which submissions should be selected within the text mining approach? Following the background literature (c.f. Table 1), an average, or typical submission is often not the richest in terms of novelty or originality. In other words, a selection that is based on representativeness will seldom be able to produce highly valuable insights for seeking firms. In clarifying lines of history and causation it is more useful to select submissions that offer an interesting, unusual or particularly revealing set of words. Following the literature on clustering, these kinds of submissions will stand out by their very unique set of used words. Hence, their cosine coefficients will be low towards most other submissions. Figure 3 illustrates an excerpt from the clustering in form of a dendrogram.

Figure 3 Dendrogram of clustered submissions within a crowdsourcing contest.

Clusters are visible by color, their aggregation is defined by their cosine coefficients, and the amount of clusters is calculated following the formula from [51]. Hence, there are different amount of clusters, including different amounts of submissions for each contest. As mentioned, following theory most innovative submissions should stand out, which means, at best they are not even part of a cluster at all (stating a so called single-item cluster). Therefore, we take two sets of submissions into the analysis, that is all submissions which appear as single document cluster, and second, all submissions which are part of clusters with up to three documents (submissions).
E.3.4 A Text Mining Based Evaluation of Submissions to Crowdsourcing Contests

We made use of theory to define what separates high quality submission from average submissions and we applied text mining methodology for semi-automated detection of submissions which supposed to be of high quality. Table 3 summarizes the process in analogy to the standard text mining process depicted in [7] and gives an answer to our first research question: How can text mining methodology be applied to support the submission evaluation process on crowdsourcing websites by suggesting most innovative solutions?

<table>
<thead>
<tr>
<th>Process-Step</th>
<th>Method</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Data extraction</td>
<td>SQL-statements on website database, Raw-text data: 112 contest with 42’448 submissions</td>
</tr>
<tr>
<td>2</td>
<td>Pre-Processing</td>
<td>Stemming, stop-word cleaning and tokenization, Cleared data set</td>
</tr>
<tr>
<td>3</td>
<td>Term Document Matrix (TDM)</td>
<td>Two weighting algorithms: tf and tf-idf, Term frequencies in all contest, calculated by submissions</td>
</tr>
<tr>
<td>4</td>
<td>Text Mining</td>
<td>Calculating similarity (cosine coefficients) and clustering (k-means), Clustered submissions per contest. (similar submissions as cluster)</td>
</tr>
<tr>
<td>5</td>
<td>Submission selection</td>
<td>Defining single-case- and cluster containing two or three submissions as outstanding, Semi-automated selection of best submissions</td>
</tr>
</tbody>
</table>

Table 4 Process of a text mining based evaluation of crowdsourcing contests.

After raw text data is extracted from the crowdsourcing website, pre-processing is used to clear the data from meaningless terms and preparing it for text mining procedures. For each crowdsourcing contest, the TDM is calculated as words (terms, T) by submissions (documents, D). Overweighting less used terms by using the tf-idf format may already highlight outstanding ideas. Calculating the similarity of submissions within one contest by using their cosine coefficients opens the possibility to aggregate submissions into cluster. Following literature, we define cluster which include only one, or a maximum of three submissions, to contain ideas of outstanding quality and hence, the submissions which are selected to be rewarded.

E.4 Results

To evaluate the text mining approach, described during chapter 3, we measure its overall accuracy. Therefore we compare the two given kinds of selection processes, the real-world decisions by expert committees against the semi-automated submission selection process applying the described text mining approach. The intention is to
answer our second research question: To what extend can a text mining based analysis of submissions to crowdsourcing websites reproduce the results of expert committees in regards to selecting most innovative submissions?

Hence, the simple overall model to test is whether the text mining based approach is capable of reproducing the expert results. This makes text mining based selections our independent variable and the real world expert committee decision our dependent variable. Still, as we use different methods of measurement during the text mining approach, four different models have to be evaluated.

- Model A uses a TDM of type tf and clusters with only one submission to define which submissions are selected.
- Model B uses TDM of type tf-idf and clusters with only one submission.
- Model C uses TDM of type tf and clusters with up to three submissions.
- Model D uses TDM of type tf-idf and clusters with up to three submissions.

Following, all 42’448 submissions are used to evaluate those models. Evaluation follows standardized measurements from the field of Information Retrieval [54]. A descriptive analysis of all four models in terms of the selection task is shown in Table 4. The four quadrants of the so called confusion matrix [55] exhibit the absolute values of classifications made by the text mining approach, i.e. true positive results at top left, false positive at top right, false negative at bottom left ad true negative at bottom right. For instance, using model A a total of 635 submissions are selected. 522 of them are true positive, i.e. selected by both, the text mining approach and expert committees. Therefore, these are so called “hits”.

<table>
<thead>
<tr>
<th>model</th>
<th>Selected by expert committees</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>A</td>
<td>522</td>
<td>113</td>
</tr>
<tr>
<td>B</td>
<td>367</td>
<td>69</td>
</tr>
<tr>
<td>C</td>
<td>1’222</td>
<td>351</td>
</tr>
<tr>
<td>D</td>
<td>949</td>
<td>237</td>
</tr>
</tbody>
</table>

Table 5 Selection of submissions made by the text mining approach compared to expert decisions.

However, at the same time model A produces 1’798 false negatives, which are called “misses”. These submissions are only selected by the expert committees and not found by the text mining approach. On the other hand, 113 false positive submissions are only selected by the text mining approach, but neglected by experts. Finally, the vast
amounts of submissions are true negatives, stating not being selected in any of the two ways. In a next step the absolute values from confusion matrix are used to calculate common metrics which measure the accuracy of the text mining approach, that is to say Precision, Recall and F1-score. Those metrics are calculated as following [54]:

- Precision = true positives / (true positives + false positives)
- Recall = true positives / (true positives + false negatives)
- F1-score = 2 * Precision * Recall / (Precision + Recall)

Table 5 summarizes those metrics for all four models. The results show that all models score rather high in precision and rather low in recall. This means all models tend to have a low amount of false positives, but unfortunately also a rather high amount of false negatives. In other words, selected submissions are mostly included in the experts picks, but experts mostly pick further submissions on top. That also explains higher Recall and F1-score values for models C and D. As those models include cluster including up to three submissions, they simply select more submissions, which comes closer to the behavior of the experts.

<table>
<thead>
<tr>
<th>model</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>82.2 %</td>
<td>23.2 %</td>
<td>0.362</td>
</tr>
<tr>
<td>B</td>
<td>84.1 %</td>
<td>16.3 %</td>
<td>0.273</td>
</tr>
<tr>
<td>C</td>
<td>77.9 %</td>
<td>54.3 %</td>
<td>0.639</td>
</tr>
<tr>
<td>D</td>
<td>80.0 %</td>
<td>42.2 %</td>
<td>0.552</td>
</tr>
</tbody>
</table>

Table 6 Precision, Recall and F1-Scores for all models.

By not overloading rare terms, that is using the tf instead of the tf-idf type of a TDM, and not limiting selections to single submission clusters only, Model C is most valid in reproducing the decisions from various expert committees (F1-score of 0.639), mostly because is scores highest in Recall. The inverse relation between Precision and Recall can be described as rather typical. For instance, one can often increase Recall by simply retrieving more documents [54]. In our case this would be possible by expanding the applied cluster sizes within the models. Ultimately, those results are a direct consequence to our initial definition of quality. The results show that in terms of used words, expert committees are also rewarding standard or average submissions. This does not mean that outstanding ideas get lost. In fact, high Precision in all models shows that using unique sets of words correlates with the chance of a submission of being selected. But the results also show that this aspect does not do the entire trick of evaluating submissions.

**E.5 Discussion and Conclusion**

As stated, it is a longstanding dream of the community to have algorithms that are capable of automatically reading and obtaining knowledge from text, that are capable
to understand human language. Despite great achievements in the field of text mining and natural language processing [7], [49], [53], [56], we will not have such possibilities in the near future. As stated in [58], many researchers think it will require a full simulation of how the mind works before we can write programs that read and understand the way people do.

So what can we learn from our study? We used long existing text mining algorithms and applied them on the modern research field of crowdsourcing contests. Our intention was to detect outstanding, innovative ideas, submitted by crowds, due to their likelihood of using unique sets of words and hence, separating them from a mass of so called noise. The empirical results are based on over 40'000 submissions. They show that text mining can serve as an approach to detect outstanding ideas. However, our approach has shortcomings and hence, should rather be seen as an initial step.

Overall all four models can be described as rather conservative selectors [54], meaning very few documents get selected in general and this is causing rather low Recall scores. This is due to the fact that following literature on crowdsourcing quality, we intended to focus on uniqueness. In contrast, expert committees seem to rather give plenty of lower rewards than following a “winner takes it all” strategy, which is a slightly different definition of quality. Hence, when it comes to selection of winning submissions, a text mining based approach must also be aligned with the reward structure of a crowdsourcing platform. Additionally, we treated all contests the same in regards to the semi-automated selection process. However, the 112 different contests addressed different topics, had different expert committees who applied individual reward structures and had differentiating opinion about relevant measuring criteria, especially the weighting of a submissions’ uniqueness. Though it is our belief that taking those criteria into account would improve the results, it also goes beyond the scope of this paper, i.e. taking an initial step. A final shortcoming to mention is that we clustered submissions by their common use of unique words and therewith neglected the possibility of using more sophisticated clustering, e.g. by n-gramms or phrases. Future research should also elaborate on the question, which clustering method works best for comparing submissions to crowdsourcing platforms.

To sum it up, we do not suggest that a complete evaluation process should be based on text mining. Text mining could be used as decision support of expert committees as it provides fast and direct entrance to unique ideas. Concerning the rising problem of an increasing number of ideas, concepts or solutions being submitted by the crowd, text mining could facilitate the current situation of which expert committees commonly are unable to cope with. A final example should give further evidence to this result. Within our used example of the sports apparel contest the following idea (in raw text) received the highest reward from the expert committee:
“You could embed super-absorbent polymers (SAPs) into your clothing. The garment is now able to produce several types of cold: Total cool down (SAPs with cold water), cool down (SAPs with water at body temperature, cooling by evaporation), warming (SAPs with warm water). I don’t know SAPs by hard, eventually evaporation of water from SAPs is too slow in your case and [...]”

As in this case the expert committee had to read through 98 submissions to detect this submission, applying the text mining approach it becomes visible on first sight, mainly because of the uniqueness of the term “SAPs” within this contest. Hence, within the text mining approach, the same submission had the lowest average cosine coefficient, which means it had fewest in common with the other ideas, and therefore was selected by all four models. Still it can and should not be concluded from a study of a single data sample of the applied text mining approaches, which combination or combinations of these should be implemented in any particular situation. However, the identification in this study may guide future efforts to determine ideal combination of text mining algorithms during the evaluation of crowdsourcing contests.

E.6 References


[38] A. B. Clegg, Computational-Linguistic Approaches to Biological Text Mining, School of Crystallography Birkbeck, University of London, 2008.


**Contribution F: Data Mining and Social Network Analysis of Ideation Contests: A Repeated Measures Design**

<table>
<thead>
<tr>
<th>Title</th>
<th>Data Mining and Social Network Analysis of Ideation Contests: A Repeated Measures Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authors &amp; Affiliation</td>
<td>Thomas P. Walter, Andrea Back, Institute of Information Management, University of St. Gallen, Mueller-Friedberg-Strasse 8, 9000 St. Gallen, Switzerland, {thomas.walter, andrea.back}@unisg.ch</td>
</tr>
<tr>
<td>Publication Type</td>
<td>Journal Paper</td>
</tr>
<tr>
<td>Year</td>
<td>2013</td>
</tr>
<tr>
<td>Status</td>
<td>Under Review</td>
</tr>
<tr>
<td>Full Citation</td>
<td>Thomas P Walter, Andrea Back, Data Mining and Social Network Analysis of Ideation Contests: A Repeated Measures Design, HSG, Working paper, January 2013.</td>
</tr>
</tbody>
</table>

*Table 1 Bibliographical information for Contribution F.*

**F.1 Introduction**

“Innovation is not the product of logical thought, although the result is tied to logical structure.” Albert Einstein once said. Given the right conditions, ideation contests, which are web-based competitions of users who use their skills, experiences and creativity to provide a solution for a particular contest challenge defined by an organizer, can achieve outstanding results. For instance, recent solutions derived from ideation contests at InnoCentive, an ideation contest organizer that focusses on a broad range of contest domains such as engineering, computer science, math, chemistry, life sciences, physical sciences and business, include a low-cost rainwater storage system which now is in use across Africa, an off grid illumination device that now is used in areas of developing countries that are not connected to an electricity network or the concept of a school system which is based on cellular phone technology that distributes educational material by causing no-costs [1]. In a single company case study, market analyst Forrester Research illustrates the financial impact of using InnoCentive ideation contests in the research and development department of a large consumer products organization. The report arrives at the conclusion that ideation contests at InnoCentive achieve a return on investment (ROI) of 74%, with a payback period of less than three months [2].

Notwithstanding these success stories, it is not possible to plan innovation provoking ideation contests from scratch. Foremost, the initially mentioned logical structure of ideation contests itself, that is to say the incremental procedure which leads towards such impressive solutions (the so called ideation function), has not yet been fully unraveled. While we know lots of anecdotical examples where ideation contests have led to remarkable results, about the benefits of certain technical features supporting the
ideation process or about the relevance of various influencing factors like monetary rewards or feedback mechanisms, yet little is known about the incremental process steps within ideation contests. Research still struggles with questions like “How many ideas does it require to achieve high ideation quality?”, “Does time play a significant role when it comes to ideation quality?” or “How does idea aggregation work?”. As a result, detailed findings on the processes within ideation contests are rare. Foremost, ideation contests are threatened as a black box, which transform some adjustable independent variables into a plurality of highly valuable ideas. Figure 1 illustrates this research gap.

![Figure 1 Analogy of an ideation contest as black box](image)

Since [3] introduced his theory of brainstorming in 1957, researchers have been focusing on such input-output correlation. Likewise, basic tenets of ideation remained stable, albeit Information Systems (IS) like Electronic Brainstorming Systems (EBS) [4] or recent Crowdsourcing Platforms have rolled up the entire process of ideation. Today a variety of modern online ideation contests frequently provide a variety of corporate tasks to the anonymous crowd of internet users. The common understanding follows [5], who stated that as long as certain amounts of everyday people elaborate on these tasks, their aggregated results can excel the results a single expert achieves. MIT’s Center for Collective Intelligence defines such situation as the appearance of collective intelligence (CI): Groups of individuals doing things collectively (connected by computers) that seem intelligent [6], [7]. Next to the laws of large numbers, frequently cited requirements of CI (or a wise crowd) are diversity, decentrality and independence in opinion and a subsequent process of aggregation [8–10]. But despite theory references the aggregation mechanism as highly relevant (for ideation quality or a CI), the aggregation within ideation (the ideation function itself) is a key question largely left behind by IS-research. Hence, this study’s key research question is if and how the ideation function can technically be extracted from an ideation contest and if so, what we can learn on the aggregation process within the contests. The guidance of this question leads us to phrasing two proper research questions. In consideration of the appropriate procedural approach to extract and analyze the ideation function of ideation contests we ask:

**RQ 1:** How can the ideation function technically be measured and analyzed so that we can draw implications on the quality of ideation?
As RQ 1 has a technical focus, our second research question is stated with regards to the research field of crowdsourcing, ideation contests and CI. Therefore we ask:

**RQ 2: Which cognitive abilities of a solver crowd can be detected within ideation contests and how are they visible in intermediate and final results of an ideation contest?**

Thus, the second question aims on converting techniques from RQ1 into practice. Subsequent questions are how the final result of an ideation contests emerge from the crowd’s submissions, whether ideas or concepts are built up on each other and to what extent winning ideas are a proof of an outstanding creativity or a result of CI? To elaborate these questions our survey will proceed as following. Chapter 2 will provide the theoretical background. We conduct a literature analysis on ideation theory, empirical studies on crowdsourcing and ideation contests and the methods of measurements applied during those studies. Chapter 3 focusses RQ 1 and develops a procedure model to measure and analyze the ideation function in online ideation contest. We suggest a repeated measures design, including data mining techniques and social network analysis to describe the ideation function. Chapter 4 puts the procedure model into practice. Applying repeated measures after every idea, we exploit an exemplary ideation contests called “The motorbike of the future”, including 725 idea submissions. We discuss our findings and aim on drawing conclusions from this survey in chapter 5.

### F.2 Theoretical Background

Ideation is defined as the process of generating or conceiving of ideas and concepts that may be useful for attaining some desired state or outcome [11]. Research has come up with a variety of techniques designed to increase the number or quality of ideas produced during ideation. Next to brainstorming [3], the Delphi method [12], or the five W’s, ideation contests are nothing else than yet another technique [13]. Nevertheless, the ultimate purpose of every ideation techniques is to produce good, or outstanding ideas [14]. For the theoretical background of our survey, we intend to unravel the state of the art of the question how ideation contests produce good ideas. The following three questions serve us as guidance for the following chapters: How does ideation work theoretically?, How do ideation contests work empirically? and Which research methods are applied to analyze ideation contests? The first question is used to understand ideation as processes from a theoretical perspective. The second question hence is, for which of those theoretical assumptions empirical evidence has been found and the third question focusses the methods and techniques applied during those empirical studies. Herein, the aim is to detect the current usage of data mining and SNA techniques within the literature.
F.2.1 The underlying theory of ideation contests

[3] proposed an ideation protocol, coined brainstorming, for improving ideation. Following his theory, ideation should foremost aim for high quantities and defer judgment. The theory does not instruct clear process steps but rather provides guidelines such as to form groups of twelve people, to address one specific question at a time, to welcome unusual ideas or to combine and improve ideas. He suggests that ideas with better quality would be generated when people were hold back from criticizing one another’s ideas, were open to wild or unusual ideas, focused on generating a large quantity of ideas, and sought to build and expand on the ideas of others [15]. Remarkably, [3] suggests that the first ideas that are mentioned are unlikely to be among best ideas. Subsequent, he suggests to focus on the second half of the ideation process, as more good ideas would be mentioned there. This theory indirectly induces that, all else being equal, the more ideas submitted, the more likely it is that good ideas are included. Until this very day, this argumentation is used as foundation of most studies on ideation contests. However, critics of Osborne’s brainstorming approach argue that evaluation apprehension, production blocking, social matching and freeriding may obstruct high quality ideation [16]. Modern information systems (IS), mainly electronic brainstorming systems (EBS, [17]) provided opportunities to leapfrog those pitfalls by facilitating the ideation process electronically.

Ideation contests are a modern form of electronic brainstorming. Following the ideas of [3], [4], [17], [18], in theory ideation contests lead to a plurality of good ideas. In terms of ideation contests, good ideas (high ideation quality) are ideas that contain novel information, that are feasible to implement, that would attain the goal, and that would not create new unacceptable conditions [15], [19–21]. Furthermore, theory often suggests, that in best case, ideation contest can invoke a CI [6], [9], [22], [23], a so called wise crowd [8]. But similar to ideation, the theory does not imply a strictly required process instruction, but rather provides guidelines for successful implementation. [8] sums up diversity, decentrality and independence in opinions as well as the existence of an aggregating mechanism as requirements. However, he does not provide specific information how aggregation does work. [9] distinguishes between outreach, additive aggregation and self-organization to impose CI. Again, outreach and additive aggregation are seen as a required process steps. The value of outreach is seen in a larger number of opinions and additive aggregation is required to get the optimum out of such large quantities. The additive process suggests collecting a large number, but also variety of opinions and building the mean. However, [24] argues by using the law of large numbers and estimation games, but neglecting ideation.

Regarding the aggregation process within ideation, contrary conjectures on the ideation function are made by theories of groupthink [25], [26], the tipping point [27]
and Bounded Ideation Theory [15]. Groupthink occurs, if the crowd’s desire for harmony overrides a realistic appraisal of alternatives. Participants then try to minimize conflicts and reach a consensus decision without critical evaluation of alternative ideas or viewpoints [28], [29]. In the end groupthink can lead to the loss of creativity, uniqueness, and independent thinking, which in turn are required criteria for collective intelligence. Possible causes for groupthink are defined in high group cohesiveness, structural faults like the lack of norm-requiring methodological procedures and the situational context like recent failures or moral dilemmas. Whether groupthink occurs in a situation is largely a subjective perception. However, groupthink theory does not define how the aggregation of opinions can leads to groupthink in a step-by-step manner. Such assumption is taken by the theory of a tipping point [27]. The tipping point states that ideas, products, messages and behaviors spread just like viruses do. Hence, the tipping point reconditions the idea of network effects and critical masses [30] in ideation. The assumption is that a former linear or steady process is swapped by a marginal idea, an idea that has major impacts, a signal that stands out from the noise and hence, changes the direction of the entire ideation process. However, applied to ideation contest, the theory does not conjecture whether the effect should be of positive or negative nature. A positive effect could be explained by an eye-opening, game-changing idea that enters a new field and inspires others to be more creative. A negative effect could be explained by the situation of one idea representing the marginal idea. After the submission of this marginal idea, all follow up ideas do not add value to the ideation contest, neither in novelty nor feasibility. To compare those theoretical assumptions we make use of their assumptions on ideation functions, which is the relationship between the total ideation values (quality of ideation) produced during an ideation session and the total number of ideas (a time factor) contributed. Figure 1 sorts theories by their time of publication and illustrates their ideation function.

Figure 2 Ideation functions implied by different theories

[15] sums up [3], [4], [18] and [16], [25] and presents the arguments of a so called Bounded Ideation Theory, a new causal model of the ideation function. Their curve argues, that the ratio of good ideas to total ideas may be smaller early within the ideation process on due to limited understanding of the task, and then larger as understanding of the task increases, and then smaller again due to cognitive overload and physical exhaustion.
F.2.2 Empirical studies on ideation contests

Even theory remains inconclusive, the analysis of ideation is not merely a recent trend, but has long tradition, especially in IS-research. IS-researchers discuss the question of computer supported ideation processes since the early days of the web, from EBS [31], creativity software [32], over the development of wiki software [33], [34] until web-based ideation platforms [35], [36] or Group Wisdom Support Systems [37]. Moreover, with an increasing amount of web-platforms similar to InnoCentive (e.g. see NineSigma or IdeaConnection) and organizations which follow open innovation approaches [38] and start crowdsourcing idea generation processes on their own (repeatedly cited examples include MyStarbucksIdea, Dell IdeaStorm or IBM’s jams), also the body of IS-literature dealing with the topic is growing. To get an overview see [39] or [40] for crowdsourcing taxonomies, [41–43] for literature reviews or [20] for a meta-analysis of 90 studies that deal with ideation quality.

A number of recent empirical studies (for example [44–49]), which aim to detect success factors of ideation contests, have revealed that design patterns (rewards, feedback and rating mechanisms, task description and many more) have significant impact on the outcomes of ideation contests. Such factors can be seen as initial “settings” to an ideation contest. They all can be set and changed by an organizer prior to the start of a contest. A second set of independent variables frequently used in empirical studies (for example see [6], [50–53]) is given by variables that describe the characteristic of the solver crowd (age, gender, profession, origin, income, hobbies, educational level, attitudes and believes). Such factors are of demographic nature and describe what kind of crowd is involved in the ideation contest. The third set of independent variables deepens the description of the crowd by additional structural variables. For instance, researchers use the number of contacts a solver keeps within the solver network [53], the past experience of solvers within online communities [54], the attitudes of solvers to a seekers brand [48], attitudes towards sharing and collaborating within a group of solvers [21], [55], [56] or the structural position within the solver network [57], [58].

When it comes to measuring the impacts of these factors, it’s merely their direct impact on the outcome, foremost the quantity and quality of ideas. In most studies the output, and in consequence the success, of ideation contests is measured indirectly. For instance, [45] use the amount of attracted solvers as indirect measurement of quality, [48] use the average rating of every idea on a five-star scale within a solver network, [59] measure quality by the boolean information whether a task was completed and [54] uses the information whether a submission was eventually implemented as primary dependent measure of quality. Also qualitative approaches can be found. For instance, [57] use data from external experts to measure quality and [60] take the evaluation from independent executives to compare crowd and expert submissions. As mentioned, only a few studies point to the characteristics of the ideation process when
defining neither independent nor dependent variables. For instance, [61] show a
temporal strategic analysis of solvers based on their decision when to enter a contest
and when to submit an idea, [36] define a mixture of interactive methods to be applied
during an ideation process. To the best of our knowledge, the closest ideation function
based measurements is presented by [62], who develops a practical, web-based
asynchronous ideation contest, which allows the implementation and test of various
incentive schemes. In this survey the amount of ideas that refer to an initial idea as
their root of inspiration are taken as quality measurement of this initial idea.

F.2.3 Research methods to analyze ideation contests
As the previous chapter shows, a tremendous volume of literature on ideation contests
has been accumulated in IS-research. At the same time, a brief but specific literature
review within this literature shows an unequal distribution of applied research
methodologies. Ideation describing theory mostly are results of experiments in
laboratory-like environments [3], [26], or are based on few remarkable incidents that
are documented within exploratory single case studies [8], [27]. Applying those
theories, explanatory research also runs experiments [18], [59], [63] or deepens
insights by interviewing solvers [57], [64].

With an increasing amount of crowdsourcing platforms, a noteworthy amount of case
study research approaches are realized. The goal is to describe the phenomena, as well
as to develop ideation design artifacts such as taxonomies, models or patterns [6], [47],
[49], [65–72]. Such research is often enhanced by descriptive statistics [36], [40], [51],
[57], [73–75]. Evaluating these design artifacts and analyzing aforementioned
independent variables, a large set of empirical studies are conducted. Using actual user
data (taken from platform data) or indirect user data (from surveys) literature applies
regression models such as OLS, PLS, ANOVA to analyze the effects of the
aforementioned independent variables on the outcome of ideation contests [14], [33],
[44–46], [48], [53], [55], [76], [77]. In a similar vein, research exploits logit models
[50], [54], [78], cluster analysis [79] or various group comparison techniques like
Cronbach’s alpha, U-test or t-test [52], [58], [80]. Additionally we find game
theoretical and other econometrical approaches to develop the structure of an ideation
contest [35], [55], [62], [81], [82].

To this point, data mining as well as SNA techniques are remarkably neglected
research methods. A brief, but precise literature review verifies this proposition.
Therefore we scan The Association of Information Systems electronic Library, AISel
using “data mining”, “text mining”, “clustering”, and “social network analysis” as
search terms to define our intended methodology and “ideation”, “idea contest”,
“crowdsourcing” and “collective intelligence” to describe our primary field of
interests. Thus, the simple question is how those methods are represented in the field?
We pairwise combine all search terms by a logical AND and intend to find a
combination of those terms in titles or abstracts of papers. The significant finding
depicted in Table 1 is, that the 32 queries (title and abstracts) leads to only three hits. There are hardly any studies available that attempt to make a contribution to the field by using data mining or SNA as a research method, even though these methods are well-rehearsed and reputable within IS-research. Expanding the literature search by a forward and backward search on the findings reveals scattered coverage of the methods within the field.

<table>
<thead>
<tr>
<th>Search Term Combination</th>
<th>Ideation</th>
<th>Idea contests</th>
<th>Crowdsourcing</th>
<th>Collective intelligence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data mining</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Text mining</td>
<td>0</td>
<td>0</td>
<td>[83]</td>
<td>[84]</td>
</tr>
<tr>
<td>Clustering</td>
<td>[85]</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Social network analysis</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

*Table 2: Results from a literature search for applied research methods within the field of ideation*

For example, [86] apply text mining to depict crime networks, [87] apply four commonly used text classification algorithms and propose a text classification framework for finding helpful user-generated contents in online knowledge-sharing communities, [88] proposes a software tool that uses the concepts of swarm intelligence and text mining to analyze free/open source software development communities and [84] run text mining methodology on user opinions expressed via twitter to analyze the appearance of a collective intelligence. In an 2006 paper [85] suggest to use (but don’t apply) hierarchical clustering and multidimensional scaling (MDS) techniques in the design of group support systems and in an 2013 paper [83] apply text mining to select best ideas from crowdsourcing campaigns semi-automatically.

<table>
<thead>
<tr>
<th>Research methodology</th>
<th>Focus on input/output relations in ideation</th>
<th>Focus on ideation function (process)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Conceptual studies</strong> (exploring the phenomena by case studies, artifact design, taxonomies, and literature reviews)</td>
<td>[6], [20], [47], [49], [66], [69], [72], [89], [11]</td>
<td>[40]</td>
</tr>
<tr>
<td><strong>Empirical studies</strong> (analyzing the impact of ideation contest design features and solver attributes by surveys among solvers and ideation contest data.)</td>
<td>[14], [44], [46], [48], [50], [52], [54], [55], [62], [76]</td>
<td>[33], [70]</td>
</tr>
<tr>
<td><strong>Structural studies</strong> (analyzing ideation by applying data mining, text mining, clustering or social network analysis)</td>
<td>[53], [57], [58], [83–85]</td>
<td>Research Gap</td>
</tr>
</tbody>
</table>

*Table 3: Summary of background literature by representative findings and identified research gap*
Table 2 summarizes the provided background literature and therewith represents the research gap we address with our survey. This gap can be split into two parts, namely the research objective and the research methodology. First, we have shown that research is rather focused on analyzing the effects of input factors on the outcome of ideation contests. The ideation function itself, that is to say the process how ideas or concepts are aggregated and enhanced, represents a somewhat mistreated research objective. Second, we detect a unilateral commitment of research methodologies, foremost the use of descriptive and multivariate statistics. Structural research methodologies, such as text or data mining techniques or SNA are seldom used to analyze ideation contests. To the best of our knowledge, our study is the first that combines those two gaps.

F.3 A procedure model to analyze ideation contests
This study exploits a procedure model to analyze the ideation function within ideation contests. Aiming on answering our first research question, we consider method engineering literature [90] to build a procedure model as systematic approach to analyze ideation functions. Therefore, we will explain all steps of the procedure model including relevant activities, techniques, as well as suggested tools and the expected outcome documents. In that way we intend the procedure model to be generic and repeatable within the domain of online ideation contests. During the development we adopt the blueprint of a standardized data mining and text mining procedure as described by [91–93]. To be exact, we will follow their common three step approach of pre-processing, processing or actual data mining and the final visualization. The following chapters explain those steps in detail.

F.3.1 Preprocessing
Once an ideation contests is chosen for analysis, all raw text data (the so called corpus) should be extracted from the website’s database using a query language such as MySQL. The central unit of analysis is “ideas”. To be able to analyze the ideation function, all extracted ideas must include attributes such as timestamp, rating, allotted reward or corresponding solver. The corpus then is imported to software for qualitative data analysis (QDA) together with its describing attributes. As central element of preprocessing, we suggest a manual coding of categories or concepts which are transported by the idea. Recoding existing data is widely seen as valuable within qualitative data analysis. An observed key mechanism in idea generation for product development is the association of one category of idea with another category [80]. For example, within our ideation contest example (c.f. the next chapter) the product “motorcycle” could be coded with a desired benefit category (eventually “fuel saving”) and a product feature (eventually “a lightweight frame”) to satisfy that benefit. Hence, we suggest that multiple codes per idea are possible. In that manner, also the combination [3], aggregation [8] or additive outreach [24] of separately mentioned concept categories is covered.
While coding ideas, existing taxonomies (for example, a taxonomy of motorbikes could list all parts of the bike) or a manual coding can be used to define a codebook. In any case, literature suggest to test the reliability of the coding by computing Krippendorf’s alpha [94]. Therefore, at least two coders independently have to assign codes to ideas using a commonly provided codebook. Hereafter, inter-coders agreement is used to compare the consistency of coding between coders and can be useful to uncover differences in interpretation, clarify equivocal rules, identify ambiguity in the text, and ultimately quantify the final level of agreement obtained by coders. Acceptable Krippendorf alphas can be defined according to the importance of the conclusions to be drawn from the survey. [95] suggests to rely on data that achieves alpha-values of $\geq 0.8$, consider data with $0.8 > \alpha \geq 0.667$ only to draw tentative conclusions, and discard data whose agreement measures $\alpha < 0.667$.

As final result of preprocessing, a code document matrix (CDM) can be computed. In a similar manner to the term document matrix (TDM) in text mining [93] the CDM is of $n$ by $m$ type, where $n$ represents rows of ideas and $m$ represents columns of allotted codes, that is to say concept categories. Hence, contrary to a TDM (which can be computed binary, by term frequency or inverse term frequency) the CDM is binary by definition, with “1” representing the fact that a code exists in an idea and “0” representing that it does not.

**F.3.2 Ideation Process Mining: A repeated measures design**

After pre-processing is conducted, a set of data mining techniques represent the core of our procedure model. The basic idea is to apply repeated measurements of the distance between concept categories (or ideas that are coded with these concepts) to be able to explore the changes of their relationships over time. The repeated measures design is a key constraint to our approach. Repeated measures design means to use the same research object (which are coded ideas) with every condition of the analysis [96]. Similar to a longitudinal study, we want to assess the change of concepts used throughout an ideation contest. Figure 3 illustrates our approach.

![Procedure model](image)

**Figure 3 Procedure model**

The relationships among concepts as well as idea similarity can be identified using distance measures, clustering, MDS and SNA. But those metrics will change as the process of ideation continue and it is those changes that we find particularly interesting and focus on. Hence, we suggest running all following steps of measurement...
(distances, clustering, MDS) in a cascading manner from the first to the last idea of the ideation contest.

Still, the procedure model might leave us with a cluttered outcome, especially as soon as the repeated measures design approaches larger N. This is a typical problem in reading the results from MDS or SNA [97], [98]. The final result (the overall situation) then might appear obscure and is of low explanatory character. A common approach in such situations is limiting the amount of total items and running analysis on a defined subset. In our context, subsets can easily be defined by idea’s attributes. In that manner we suggest to use subsets of ideas for deeper analysis. In terms of a noised ideation function, that might result when too many concepts are mentioned, we suggest to limit ideas by a preselected amount concept. In terms of the ideation quality, we suggest to limit the analysis to rewarded ideas. Later the results can be compared with the overall contests. Both suggestions will be applied below.

F.3.2.1 Ideation Distance Measurement

Based on the CDM, distance matrices between concepts can be computed. The matrices are of $n$ by $n$ (one-mode) type, which means that concepts represented both, rows and columns and the values represent the distance between them. By default, distance measures allow statements on the co-occurrence of concepts within ideas. Co-occurrences are said to happen every time two concepts appear in the same idea. Hereafter distance matrices between concepts can be computed using Ochiai coefficients. Contrary to the usage of Cosine coefficients to measure distances based on a TDM, the coding of ideas by concepts results in the binary CDM. Hence, we suggest to use Ochiai's coefficient as the binary form of the cosine coefficient, to measure distances between codes: The distance $d_{ij}$ between two codes $i$ and $j$ then is defined by $d_{ij} = \frac{\sqrt{a^2}}{\sqrt{(a+b)(a+c)}}$, where $a$ represents ideas where both codes occur, and $b$ and $c$ represent cases where one code is present but not the other one. For example, two codes occurring together in various ideas could be represented by a variety of ideas taking up on the aforementioned idea of “saving fuel” (say code $i$) by building “a lightweight frame” (say code $j$). The similarity of two ideas will range from “0” to “1. This construes values of “1” meaning ideas are using exactly the same set of concepts to “0”, usually indicating a total independence of ideas, and in-between values indicating intermediate similarity or dissimilarity of ideas. The more similar two documents are in term of the distribution of concepts, the higher the coefficient between them [91], [93].

F.3.2.2 Ideation Clustering

Clustering is the application of certain algorithms to automatically detect patterns within the CDM. Therefore, clustering applies the computed distance matrices to explore the similarities between various groups of concept categories. Often so called non-hierarchical (or centroid-based) clustering is applied, foremost the k-means
We suggest to use k-means to partition the concept categories into \( k \) cluster in which each observation belongs to the cluster with the nearest mean. At the bottom, k-means is based on principal component analysis or minimalizing least squares [101]. In other words, clustering will minimize the average distance of a group of co-occurring codes to other groups of co-occurring codes. However, determining the exact \( k \) number of clusters in a data set is a frequent problem in data clustering, and is a distinct issue from the process of actually solving the clustering problem. In text mining, a frequently used method to determine the number of cluster is the elbow criterion [102], which suggests to choose the \( k \) number of cluster by the maximum \( R^2 \), so that neither dropping nor adding a cluster does rise the percentage of variance explained by the cluster [103].

**F.3.2.3 Ideation Cluster Visualization**

As a final step of the three-step data mining process [91] we are following, we suggest to read-out the results from clustering. We suggest cluster visualization based on MDS. MDS is often used in information visualization for exploring similarities or dissimilarities in data. Also known as principal coordinates analysis, MDS takes the distance matrices and outputs a coordinate matrix whose configuration minimizes average distances. We suggest using concept maps, as in our case a network of concept categories evolves through the repeated measures. For sufficiently small \( N \), the resulting locations of the coordinate matrix can be displayed in a network-graph. This step can be supported by a plurality of SNA software tools, for example UCInet, Pajek, R or Network Workbench. Within the evolving concept map (network) of co-occurring concepts in ideas, nodes will represent a specific concept and weighted edges the fact how often two concepts co-occurred in ideas. The radius of a node will represent the relative frequency of the concept within the overall network. Cluster of concept are visualized short distances and unique color. For larger \( N \), idea networks might not benefit as much from MDS as some distortion may result. As a consequence, some concepts which are computed as close to each other or which are parts of the same cluster may still be plotted far from each other due to the optimization of mean maximal distances within the entire network. In such cases, the suggested approach of limiting the amount of ideas to a subset and performing MDS on a fewer number of items usually produces a less cluttered map which can be interpreted again.

**F.3.3 Ideation function and ideation network key characteristics**

The repeated measures design generates a lot of new data. To be precise, for each idea that is added, we suggest computing the overall situation (network). This approach allows us to define new metrics that are closely related to the ideation function. To plot the ideation process we can look on the sums of concepts, ideas, words, participating solvers, rewards and ratings over time. These metrics already allow a first comparison of a specific ideation contest to the ideation functions implied by different theories.
(c.f. Figure 2). Furthermore, we suggest adding a growth-perspective, which is defined by the marginal input of every idea. Subsequent questions are whether an idea adds one or more new concepts (new nodes) or focuses on recombining already existing concepts (additional edges within the ideation network). Furthermore marginal ideas can also simply strengthen an edge (adding weight), which over time will lead towards two concepts being part of the same cluster. Additionally to these ideation functions describing characteristics we suggest to include concept network characteristics, especially betweenness centrality and network density, into the analysis.

F.3.3.1 Betweenness centrality and Concept Network Density

In [8] a central suggestion to achieve a collective intelligence is keeping loose connections. Referring to the idea that often it does not require very strong connections to spread an idea, this reintroduces [104] theory on the strength of weak ties. In terms of a concept network, a tie (edge) represents the fact that two concepts (nodes) co-occurred in one idea. Hence, following SNA literature, the betweenness centrality of a concept should be the most relevant measure to analyze such characteristic [105]. The betweenness centrality \( g(k) \) of a node \( k \) is defined as the share of times that a node \( i \) needs the node \( k \) in order to reach a node \( j \) via the shortest path [106]. Specifically, if \( g_{ij} \) is the number of shortest paths from \( i \) to \( j \), and \( g_{ikj} \) is the number of these paths that pass through node \( k \), then the betweenness centrality of node \( k \) is given by:

\[
g(k) = \sum_{i \neq j \neq k} g_{ikj} / \left( g_{ij} \right)
\]

In our context the idea is that a concept which is in-between serves as transmission belt, like a bridge to a yet unseen concept combination. Such network position is defined as powerful, as it implies the ability to broker (or block the combination) of two separately mentioned concepts. Hence, ideas which introduce concepts of high betweenness centrality could be boundary spanning, or covering structural holes in the ideation contest. If later ideas use identical concept combinations, the betweenness centrality of such ideas will tend to be lower due to the existence of multiple shortest paths. Therefore we suggest comparing betweenness centrality of concepts for the subset of rewarded ideas (see above) at the final stage plus during intermediate stages (for example whenever a winning idea is submitted) of the ideation contests.

F.3.3.2 Concept Network Density

Finally, we suggest two different density measures for the concept categories network. Density describes the ratio of existing edges over possible edges (defined by the number of nodes) [107]. But as edges of the concept network are weighted, the density can also be defined as the average strength of edges across all possible (not all actual) edges [98]. The latter will take the network’s cohesion, a growing strength between a distinct subset of nodes, into account, whereas the first measure uses the binary form of a connection between two nodes. We suggest to compute the network density \( D_t \) at time \( t \) using

\[
D_t = \frac{2E}{N(N-1)},
\]

where \( E \) is the sum of weighted existing edges and \( N \) is the amount of Nodes of a network [107]. To sum up and close the development of our
procedure model, Table 2 reassembles all suggested steps, including techniques, tools and pursued results.

<table>
<thead>
<tr>
<th>Step</th>
<th>Activity</th>
<th>Technique</th>
<th>Tools</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Data extraction</td>
<td>Download the platform’s database-dump. Use a query language to extract the required information (ideas, timestamp, reward and/ or rating, user, etc.) in tabular format.</td>
<td>Query language like MySQL</td>
<td>Plain ideation contest data (the corpus)</td>
</tr>
<tr>
<td>2</td>
<td>Preprocessing</td>
<td>Code all ideas by a concept category. Check inter-coder reliability by Krippendorff’s alpha. Compute a concept document matrix (CDM).</td>
<td>Software for qualitative data analysis (NVivo9, Provalis QDA miner, Netminer, etc.)</td>
<td>Coded Ideas Reliability of coding CDM</td>
</tr>
</tbody>
</table>

**Repeat steps 3 to 5 for idea 1 to N.**

**Repeat steps 3 to 5 for defined sub-sets of ideas.**

<table>
<thead>
<tr>
<th>Step</th>
<th>Activity</th>
<th>Technique</th>
<th>Tools</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Ideation distance measurement</td>
<td>Compute distances between concepts by Ochiai's coefficient based on co-occurrence in ideas.</td>
<td>Software for QDA or Statistics (R, Stata, SPSS, Excel, etc.)</td>
<td>Code distance matrix</td>
</tr>
<tr>
<td>4</td>
<td>Ideation clustering</td>
<td>Use k-means algorithm for clustering concepts. Determine the amount of cluster using the elbow criterion (maximizing R² values).</td>
<td>Software for QDA or Statistics</td>
<td>Category cluster</td>
</tr>
<tr>
<td>5</td>
<td>Ideation clustering visualization</td>
<td>Use multidimensional scaling (MDS) to layout the concept maps (ideation networks).</td>
<td>Software for SNA (UCInet, Netdraw, Pajek, NWB, etc.)</td>
<td>Concept cluster in network layout (concept maps) or as dendrogram</td>
</tr>
<tr>
<td>6</td>
<td>Ideation function and ideation network characteristics</td>
<td>Compute the sums for ideas, words, solvers, concepts, rewards, ratings, edges over time. Compute the growth rates of ideas, concepts and edges over time. Compute network measures for ideas and concepts (betweenness centrality of ideas and overall network density measures (binary and weighted) over time.</td>
<td>Software for QDA, Statistics and SNA</td>
<td>Ideation function Process oriented ideation contest analysis Ideation network analysis Winning ideas in context</td>
</tr>
</tbody>
</table>
Compare rewarded and unrewarded ideas.

Table 4 Procedure model steps to analyze ideation contest with data mining and SNA techniques

F.4 Results
To evaluate the suggested procedure model, we apply it to analyse a real-world ideation contests. As mentioned, this approach can be seen as single case study approach and therewith is also subject to the limits of case study research [108]. Our goal is not to generalize on ideation theory, but rather to explore how research could benefit from applying methods of data mining and SNA to the field. In the following we first will describe the example dataset and thereafter the result of applying the suggested procedure model.

F.4.1 Dataset and Coding
We use data from a real world crowdsourcing website. The website was launched in 2007 and since then, 128 different ideation contests, including nearly 50’000 ideas, 20’000 ratings and 500’000US$ of rewards, have been completed. The website has 7.512 solver accounts. The average age of solvers is 41.8 years. 71.2% of solvers are male and 52.6% keep a university degree. The formal crowdsourcing procedure on the website runs as following: Various ideation contests are announced on the website simultaneously. Solvers can sign-up to the website (create a solver profile) for free, but in order to receive potential rewards they have to provide their bank account. During an ideation contests, solvers can submit various ideas, but also see, comment and rate the ideas of other solvers. The rating mechanism is similar to facebook’s “like” button, or google’s “+1” button, allowing solvers simply to express that they like another solvers idea, but not to differentiate the extent. However, the rating is solely a solver community feature with no consequence on winning a reward. Seekers pay an initial fee to the platform to get the contest online. Prior to the ideation contest seekers decide on the reward (total amount and structure), the duration of the contest, deliver a contest (task) description, announce their relevant criteria of rewarding (e.g. “most radical solution” or “feasible concepts with business impact”) and already suggest the type of expected answer (e.g. “We would like to receive plain text.”). After an ideation contest is closed, all submitted ideas are rated by a seeker’s expert committee and best ideas are rewarded. Table 3 summarizes the overall website’s and our chosen ideation contest’s descriptive statistics.

As mentioned, we chose one particular ideation contests as an example to test our procedure model. This contest was called “The motorbike of the future” and its seeker was a global manufacturer of motorbikes. The contests lasted for 104 days, during which 725 ideas were submitted by 402 solvers (1.82 ideas per solver; 6.98 ideas per day). 27 solvers submitted at least five ideas and 84 at least three ideas. The seeker rewarded twelve different ideas, aggregating a total reward of 5’000 US$. The task
description of the seeker was rather broad: “Motorcycle, scooter, moped riders and enthusiasts: We want your ideas! What will be the perfect motor bike of the future – what features should it have, how much would it cost, who would buy it? What will be the next exciting trend in motor bikes that will fascinate the youth of tomorrow? We are looking for unique, breakthrough products and services for the future of motorcycles, scooters, mopeds or the like vehicles. The solutions should embody innovative technology, passion and modern lifestyle.”

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Crowdsourcing website</th>
<th>Ideation contest</th>
<th>Ideation example</th>
</tr>
</thead>
<tbody>
<tr>
<td>∑ ideation contests</td>
<td>128</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>∑ ideas (avg. per contest, std. dev.)</td>
<td>49’284 (385.0, 89.1)</td>
<td></td>
<td>725</td>
</tr>
<tr>
<td>∑ words</td>
<td>2’089’472</td>
<td>45’759</td>
<td></td>
</tr>
<tr>
<td>• Avg. per contest (std. dev.)</td>
<td>16’324.3 (6222.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Avg. per idea (std. dev.)</td>
<td>42.4 (17.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>∑ solvers (avg. per contest, std. dev.)</td>
<td>8’512 (305.3, 181.9)</td>
<td></td>
<td>402</td>
</tr>
<tr>
<td>Avg. solver age (std. dev.) in years</td>
<td>38.12 (10.3)</td>
<td>41.8 (7.7)</td>
<td></td>
</tr>
<tr>
<td>Solver’s gender (male, female) in %</td>
<td>66.2/ 33.8</td>
<td>71.2/ 28.8</td>
<td></td>
</tr>
<tr>
<td>∑ rewards (avg. per contest, std. dev.) in US$</td>
<td>524’800 (4100.0, 522.3)</td>
<td></td>
<td>5000</td>
</tr>
<tr>
<td>∑ ratings total</td>
<td>211’921</td>
<td>2175</td>
<td></td>
</tr>
<tr>
<td>• Avg. per contest (std. dev.)</td>
<td>1’655.6 (310.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Avg. per idea (std. dev.)</td>
<td>4.30 (0.82)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>∑ Contest duration (avg. per contest, std. dev.) in days</td>
<td>11’712 (91.5, 14.2)</td>
<td>104</td>
<td></td>
</tr>
</tbody>
</table>

Table 5 Descriptive statistics of the dataset

We followed the procedure model (c.f. Figure 3, Table 2) to analyze all ideas that were submitted to this task. While extracting the raw ideation contest information (the ideas in plain text) from the crowdsourcing website's database using MySQL, we also included various types of arrays of every idea. This included the idea’s timestamp, solver ID, reward and rating. All this information was imported to the Provalis Research QDA-Miner software. Within the software, all ideas were coded by responding categories. A first coder coded all ideas by a category following basic rules of coding [109]. The goal of categorizing ideas was to describe how (which part or by which act or process) motorbikes should be innovated. The depth of coding was ceded to the coder. For example, various ideas addressed environmental issues. However, as they suggested various aspects, resulting codes included E-motor, Hybrid, Fuel-saving, CO₂-emission, Solar power, heat-reusage, etc. Likewise, ideas that addressed a business issue included codes such as Vendor Service, Brand- refreshment, Versioning/ Customization, Salesprice, etc. The first coding resulted in a codebook
containing 70 concept categories. This codebook was provided to two other coders, who also coded all ideas, being allowed to only using codes from the codebook. Interrater reliability was assessed by calculating Krippendorff’s alpha for all 70 codes. The reliability for most of the codes is above .67, the commonly canonical value. Perfect agreement was achieved for concept categories that described parts or types of a motorbike (for example Wheels, Seats, Storage, Frame, Handlebars, Sidecar, Quad, Trike, Jetski or Snowmobile all resulted in values of 1.0). Lowest agreement was achieved in rather technical categories or categories that described the benefit of an action. For example, the agreement values of fuel cell, heat re-usage, on-board electronics, novel vendor service, brand-refreshment or visibility are .62, .66, .66, .65, .66, .67, respectively. Remaining codes received values above 0.67. Given the difficulty of the specific task (and the fact that none of the coders is a motorbike experts), those results seem to be satisfactory [95], [109], [110].

**F.4.2 Repeated measures to analyze concept cluster development**

Applying steps 3 to 5 from the procedure model enables us to analyze and visualize how the solver’s ideas are built up and relate to each other while creating an ideation network. We apply the repeated measures design by using the ideas’ timestamps and calculating a new CDM after every idea that is added during the ideation contest. Code distances are used to cluster ideas, which are visualized by applying MDS to draw concept maps.

![Figure 4 Concept maps based on CDM, clustering and MDS after ten, 20, 30 and 40 of 725 ideas (top left to bottom right).](image)

In other words, following the procedure model we are able to describe the ideation process by using 725 concept maps and the included network characteristics. Figure 4 illustrates four instances, that is to say the concept maps after ten, 20, 30 and 40 of the 725 ideas.
Stringently, early ideas contain high levels of novelty. To be precise, in terms of words or categories, the very first idea that is submitted to an ideation contest will have endless novelty by definition. In our case the first idea was the following: "My idea is a networked motorcycle. The bike of the future should include lots of networked features. I think, that based on GPS and your online profiles it should be possible that your bike suggests locations for you to stop, find the way to appointments that you have in Outlook or is recognized by your home whenever you approach your garage by 100m or so. A networked bike would also include a feature that allows you to stream your music from your home computer or a feature that allows your mechanic to get remote access to your bike’s system."

This rather comprehensive idea already mashes up a technical idea (new sensors) with new business offers (remote access by a mechanic, streaming). Appurtenant concepts are Sensors, On-board-electronics and Vendor service, and they are represented within the blue cluster, visible in the top left image of Figure 4. As ideas are visible to other solvers, subsequent ideas try to circumvent initial ideas, which leads to other concepts also visible in Figure 4. Consequently, after 10 ideas the following situation emerges: Twelve different concepts are mentioned within these ten ideas. The elbow criterion suggests two cluster (R² = 0.8655). Apparent by larger node-radius, four concepts (Driver Suit, Airbags, Sensors, Vendor Service) already are mentioned at least twice and as the edges signalize concepts also briefly co-occur in ideas (as within the first one). With further ideas being submitted, nodes are added (due to new categories), node sizes change (due to concept frequencies), edges are added (due to co-occurrence of categories), the weight of edges changes (due to multiple co-occurrence of categories), the amount of cluster changes (due to the elbow criterion) and a concepts membership to a cluster changes (due to the distance measures). Figure 5 illustrates this by opposing the concept networks at halftime (N=362) and finish (N=725) of the ideation contest.

Figure 5 Concept maps of an ideation contests after N=362 (left) and 725 (right) ideas

Figure 5 illustrates that by the nature of ideation, the likelihood of novel concepts decreases over time. Table 4 supports this hypothesis, by presenting equivalent network characteristics. By the end of the first half, the majority of concepts (60 of 70 or 85%) are already submitted. Conversely, the second half of ideas that are submitted to the contests significantly increase the amount and weighting of edges. This is given
by the concept network’s density of 11.8% (208 edges of 1770 possible edges between 60 concepts) at N=362 against 20.1% (486 of 2415 for 70 concepts) at N=700 ideas. In other words, whereas the first half of ideas is rather focused on brainstorming and rapid mentioning of new concepts, the second half of ideas is rather focused on recombining and aggregating concepts that already have been mentioned. Considering the weighting of this effect becomes even stronger. At N=362, 298 edges are created (making 90 or 30.2% duplicates) while at N=700, 820 edges are created (making 334 of them or 40.7%) duplicates. Thus, not only the creating of new edges accelerates during the second half, but also the weighting. Using a negative expression, this means that during the second half 99.0% of the mentioned concepts are copies, and also nearly every second combination (46.7%) of concepts has already been mentioned.

<table>
<thead>
<tr>
<th>Network characteristic</th>
<th>N=1 to N=362</th>
<th>N=363 to N=725</th>
<th>N= 725</th>
</tr>
</thead>
<tbody>
<tr>
<td>∑ Nodes (binary)</td>
<td>60</td>
<td>10</td>
<td>70</td>
</tr>
<tr>
<td>∑ Nodes (frequency)</td>
<td>621</td>
<td>954</td>
<td>1575</td>
</tr>
<tr>
<td>∑ Node duplicates frequency (%)</td>
<td>561 (90.3)</td>
<td>944 (99.0)</td>
<td>1505 (95.6)</td>
</tr>
<tr>
<td>∑ Edges (binary)</td>
<td>208</td>
<td>278</td>
<td>486</td>
</tr>
<tr>
<td>∑ Edges (frequency)</td>
<td>298</td>
<td>522</td>
<td>820</td>
</tr>
<tr>
<td>∑ Edge duplicates frequency (%)</td>
<td>90 (30.2)</td>
<td>244 (46.7)</td>
<td>334 (40.7)</td>
</tr>
<tr>
<td>Network density (%)</td>
<td>11.8</td>
<td>29.5</td>
<td>20.1</td>
</tr>
</tbody>
</table>

Table 6 Ideation concept network characteristics, split by two halves of an ideation contest

In that manner, clustering leads to the emergence of visible, dominant concepts as well as so called single item cluster (a single concept defining a cluster of its own). For example, as by the time of N=362, ideas which included safety concepts are caught within the major cluster (colored red in the top-left network of Figure 5), but increasing co-occurrence of a subset of safety concepts leads to the pink colored cluster the left image. This represents the fact that between N=362 and N=725 ideas, safety concepts like Airbags, Protectors and Safetybelts remarkably co-occur, but also together with concepts like Comfort, new kinds of Seats and Driver Suits. In other words, as the contests proceeds, solvers adopt the safety issue and increasingly recombine it with an issue of higher comfort. The repeated measures design shows, that ideation is by no means a linear process. It becomes visible, that there is a brainstorming phase in the beginning of the contest and an elaborating- or framing-phase in the second half of the ideation contest. While novelty decreases in terms of new concepts, further ideas aggregate and hence, help to carve out more specific or detailed concepts. Figure 6 shows this by presenting the ideation contests metrics and characteristics as suggested in step 6 of our procedure model.
During the first 20% of the contest’s allotted time already 80% of the concepts are mentioned, but the vast amount of submitted ideas during the remaining 80% of the time do not add new concepts, but just recombine or worse) repeat existing ideas. The ideation function shows some solvers as productive early mover. The first 20% of all participating solvers already mention over 50% of all concepts in the first 18% of ideas and within less than 5% of the contest time. By half time close to 90% of the concepts are mentioned, but yet only 35% of the contest participating solvers have joined the contest. Hence, approximately 60% of the users seem to be laggards (maybe freeriders), as they join the contest at a time already 90% of the concepts are mentioned.

Like [3] conjectures, the second half of the ideation contest is different. But in our case it is not the high quality that causes separation, but rather it’s merely changing network characteristics. 60% of the solvers join the contest within this phase creating 50% of the final raw material (in terms of ideas or words). On first sight, these ideas are of lower novelty, as they do not add new concepts. On the other hand, those ideas appear to be equally valuable as they receive 50% of all ratings and also more than half of the total rewards. Consequently, we will deepen the analysis of idea and concept aggregation within the next step.

F.4.3 How a specific concept develops within an ideation contest

As aforementioned, performing a MDS on a large number of items usually produces a cluttered map that is hard to interpret. This is the case if we consider the concept maps for higher N (c.f. the network at N=725 in Figure 5). To avoid the noise, created in those maps and provide more detailed view on the ideation process we will limit the amount of ideas within this step. As a result, we will again analyze the co-occurrence
of concepts, but only in regards to a centralized concept. We use the concept of Versioning/Customization, which is the concept with the highest betweenness centrality within the finished ideation network (at N=725). The concept is used by 75 ideas which suggest that the motorbike of the future should be built modular, so that a vendor or the customer himself can reassemble various parts and hence, customize the motorbike towards differentiating customer needs. Below, we will present the results of applying the procedure model on this limited subset of N=75 ideas, neglecting all the other (650) ideas.

The first co-occurrence of the concept happens with the submission of the 23rd overall idea. As the 2nd idea overall already suggested variety in frame designs and storage capacity to create customizable motorbikes (coded with Frame/Storage and Versioning/Customization), the 23rd idea reintroduced the concept, of customization, but from a business model perspective, suggesting focusing on new kinds of customer groups (coded with Brand/Customer Segment and Versioning/Customization). With the ideation contests continuing, more ideas adopt the concept of Versioning/Customization. Thus, due to repeated measures and MDS Versioning/Customization represents a centered concept whose frequency will be plus one for every idea added. This is visible by increasing node size of the centered concept in Figure 7.

Figure 7 Cluster development of concepts co-occurring with the centered concept of Versioning/Customization at N=2, 5, 10 and 15 ideas using the centered concept

Figure 7 shows, that concepts are rarely copied during this stage of the contest, but rather enhanced. For instance, at N=5 (by the 34th overall idea) the first concept (from N=1) is enhanced. The solver takes the initial idea (to vary frames and storage) and combines it with the idea to build it as a trike (c.f. Figure 7, top right, blue cluster). At N=15, four cluster are built. For example, the pink cluster collects ideas that suggest reinventing the sidecar. However, this idea is enhanced by the ideas of an electrical engine and an idea which suggests building bike and sidecar as modules of a quad (which is a 4 wheel motorbike). But the idea of an electrical quad itself has not yet
been mentioned at this time of the contest as we can recognize by the missing edge between those concepts.

During the ideation contest, discrete “sub-sub-concepts” such as customizable motorbikes with sidecars, are gleaned, reintroduced and enhanced with other concepts, but also copied, neglected or forgotten. Staying with the mentioned example of a sidecar illustrates further characteristics of the ideation function. Ideas suggest special protection solutions for the passenger sitting in the sidecar (at N=18), multimedia to entertain the passenger (at N=42), or highly comfortable seats for the sidecar (at N=44). However, at the same time the relative frequency of concepts within this sub-network is decreasing. The reason for this is that other ideas (for instance the multimedia solution) get adopted by different, faster developing concepts. This leads to gaining density of the red colored cluster, which is created around the centered concept. Hence, at N=36, the concept of an electrical engine switches from the cluster around the idea of a sidecar towards this centered (red colored) cluster. Repeated co-occurrence of codes within the centered cluster (for example a scooter-type of bike that has roll bars and hence, can be seen as a convertible) leads to higher density and cohesion within this cluster. Figure 8 illustrates these changes of the code network structure, foremost the situation of the central cluster becoming more powerful.

![Figure 8 Cluster development of concepts co-occurring with the centered concept of Versioning/Customization at N=25, 35, 45, 55, 65 and 75 (from top left to bottom right)](image)

Despite this development, it remains hard to determine a certain tipping point [27] nor opinion building or a groupthink situation [25] in this ideation function. Likewise the
situation does not reflect ideation as it is suggested by [3] and even provides an inverse ideation function to the one that is suggested by bounded ideation theory [15]. Like the ideation network of the overall ideation contest, also the sub-network of the developing concept is built in phases. Again, the second half (from N=38 until N=75) is not less than productive than the first half. Although it does not stand for many novel concepts (nodes), 70% of the new combinations or aggregation (edges) is set here. To sum up the repeated measures steps from our procedure model, this single case study shows a three-phased ideation function, depicted for the single cluster development in Figure 9.

From the ideation contest’s start, early movers and the early majority starts brainstorming. During this phase a large variety of novel concepts is introduced. Some of those ideas will be solid ground for other solvers to adopt on and some of those early ideas will be neglected by the crowd. During the middle part of the ideation function solvers already focus on aggregation. This results in strong cluster further being strengthened while some “outstanding” ideas (representing single-item cluster) still are not further pursued by the crowd. The final phase of ideation focusses on strengthening already existing co-occurrences of various concepts (weighting of existing edges). The negative description of this phase is “copying”. A clear sign of saturation is the network’s density (c.f. Figure 9) or the ratio of new edges per new idea.

![Figure 9 Ideation function of a centered concept within an ideation contest](image)

To deepen the aspect, how ideation characteristics are reflected by the final results of an ideation contests we will provide a final comparative analysis of rewarded and unrewarded ideas in the next chapter.
**F.4.4 Where good ideas come from**

As literature review and theory have shown, the issue of ideation quality opens a multi-sided discussion between researchers. To this point, research only achieved uncertainty what factors have influence on ideation quality and how ideation quality should be measured. In addition to factors which measure the impact of contests design features, extrinsic or intrinsic motivation on quality, we are able to determine if ideation network characteristics (introduced in 3.3) may are aligned with the quality. Using our existing data provides two opportunities of indirect quality measurements. First, we can make use of the solvers ratings of their own ideas and second, we can use rewards given by the seeker. As ratings might be biased by individual connections among solvers as well as their different perception and hence, practice of the rating mechanism [44], we will make use of the rewards that are given by a seeker’s expert committee. Therefore, we act on the assumptions, that the expert committees follow a fair rewarding process, are able to cope with the quantity of ideas (which is a problem according to [83] and also stick to their announced measurement criteria.

In the selected case the seeker’s task description calls novelty (*We are looking for unique, breakthrough products and services.*) and feasibility (*The solutions should embody innovative technology, passion and modern lifestyle.*) as relevant criteria of quality and reward twelve out of the 725 ideas. Therefore, our next step is to turn to the “best versus the rest” of the ideas in terms of their network characteristics. As noted above, in terms of an evolving ideation network, an idea is novel if it creates a new node (introduces a new concept) or a new edge (is the first that combines two nodes which have been mentioned before, but separately). A third nuance of novelty can be defined in the fact, whether the combination of nodes in an idea already has been used at the time of submission. For instance, an idea might use four different concepts (nodes) and hence, six edges, but none of the edges represents novelty as they might have been used pairwise by various seekers already. However, the situation of combining exactly those four concepts might still be novel. Table 6 applies those metrics to compare twelve rewarded and the 713 unrewarded ideas from our example.

As Table 4 shows, 70 different concepts are used within 725 ideas. Enhancing existing concepts leads to a total concept frequency of 1575, meaning on average, every idea includes 2.17 concepts. Remarkably, none of the rewarded ideas introduces a new concept, although on average they combine more concepts than unrewarded ideas (3.0 to 2.21). In other words, all rewarded ideas use concepts, which, by the time the rewarded ideas are submitted, already exist within the contest. Hence, the assumption is that rewarded ideas do well in enhancing these concepts. Looking at the edges stretches evidence to this hypothesis.

Within the 725 ideas the 70 concepts are tied by 468 edges. Additionally, 334 times an edge is strengthened by repetition. As a result, on average an idea coins 1.09 edges, of which 0.68 represent a new connection and 0.41 a repetition. In contrast, the twelve
rewarded ideas create 3.5 edges on average of which 1.33 are new combinations and 2.17 repetitions. Additionally, by the time they are submitted, 9 of the 12 rewarded ideas establish a new, and hence so far unique, combination of concepts.

<table>
<thead>
<tr>
<th>Concept network metrics</th>
<th>Node frequency</th>
<th>New nodes created</th>
<th>Node weight</th>
<th>Edge frequency</th>
<th>New edges created</th>
<th>Edges weighted</th>
<th>Novel node combination at time of submission</th>
<th>Betweenness centrality</th>
</tr>
</thead>
<tbody>
<tr>
<td>12 rewarded ideas</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• mean</td>
<td>3.0 (1.33)</td>
<td>0</td>
<td>1.33</td>
<td>2.17</td>
<td></td>
<td></td>
<td>75.0%</td>
<td>17.88 (12.2)</td>
</tr>
<tr>
<td>• (SD)</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td>22.35</td>
<td></td>
</tr>
<tr>
<td>• median</td>
<td>3.0 (1.33)</td>
<td>0</td>
<td>1.33</td>
<td>2.17</td>
<td></td>
<td></td>
<td>75.0%</td>
<td>17.88 (12.2)</td>
</tr>
<tr>
<td>713 unrewarded ideas</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• mean</td>
<td>2.21 (0.69)</td>
<td>.10 (0.08)</td>
<td>1.09</td>
<td>.68</td>
<td></td>
<td></td>
<td>37.2%</td>
<td>9.15 (15.6)</td>
</tr>
<tr>
<td>• (SD)</td>
<td>2</td>
<td>0</td>
<td>0.92</td>
<td>0.51</td>
<td></td>
<td></td>
<td>9.15 (15.6)</td>
<td></td>
</tr>
<tr>
<td>• median</td>
<td>2.21 (0.69)</td>
<td>.10 (0.08)</td>
<td>1.09</td>
<td>.68</td>
<td></td>
<td></td>
<td>37.2%</td>
<td>9.15 (15.6)</td>
</tr>
<tr>
<td>Mann-Whitney U-test U-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Z-value</td>
<td>-1.355 (.045)</td>
<td>.887 (.023)</td>
<td>-1.82 (.067)</td>
<td>-2.015 (.052)</td>
<td>-1.956 (.50)</td>
<td>-1.744 (.38)</td>
<td>-2.26 (.016)</td>
<td></td>
</tr>
<tr>
<td>• (p-value)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7 Comparison of concept network metrics for rewarded and unrewarded ideas

We can see in Table 6, that the rewarded ideas have higher betweenness centrality than unrewarded ideas (17.88 to 9.15). Also there is a lot of variation in betweenness centrality for the overall network (for example SD = 15.62 relative to a mean betweenness of 9.15 for all unrewarded ideas). This makes sense, because we know that as edges can be created by the use of one out of of 70 concepts, most connections between ideas can be made in this network without the aid of any intermediary. Hence, combining the network numbers and the fact that there cannot be a lot of betweenness, the difference between rewarded and unrewarded ideas is a significant hint that I this case, rewarded ideas are somehow bridging structural holes within the ideation network. Figure 10 illustrates the network of all ideas. Rewarded ideas are colored yellow and the edges towards their ten closest neighbors (with shortest distances in the distance matrix) colored red. We can see that some of the rewarded ideas (for example the three depicted in the middle-top position of the network) are among closest neighbors to each other, stating they are very similar according to coded concept categories. Even though the network of 725 ideas is a bit cluttered and per definition produces some distortion, we are able to recognize rather central positions of rewarded ideas.
Figure 10 Idea network, based on coding. Rewarded ideas in yellow, edges to 10 nearest neighbors in red

Summing up this analysis, we can say that in this particular example, rewarded ideas are not the ones which introduce new concepts, but rather those which combine them in a clever fashion. The network characteristics of rewarded ideas suggest that they are rich in information (given the fact that they include 3 concepts on average), even though particular concepts have been applied before. Rewarded ideas are able to introduce a novel way of combining those concepts (given the fact that on average they produce 1.33 new combinations). This means that even though their content richness includes some repetitive nature, the ideas over all represent novel combinations. As a result they are positioned central, in-between all other ideas in the ideation network. A hypothesis, taken from these results could be, that high quality ideas might be boundary spanning. Their main strength might be situated in reaching across different concept borders in order to build relationships, interconnections and interdependencies that finally solve a complex problem.
F.5 Discussions

We recombine long existing methods, which are to say coding, distance measures, clustering mechanisms, MDS and SNA to develop a procedure model which aims to analyze the process of ideation within online ideation contests. Our intention is to develop an approach that might help to understand how innovative ideas are formed and embedded within ideation contests. We focus on the process of ideation (often called the ideation function), not on input factors or the results. We put the developed procedure model to practice by applying it to analyze an ideation contests, taken from a crowdsourcing platform. The empirical results are based on 725 ideas, which were submitted to an ideation contest searching for innovative concepts to create the motorbike of the future. The analysis of repeated measures for the overall contest and the development of a single concept as well as the comparative analysis of rewarded and unrewarded ideas allows us to discuss some theoretical as well as managerial contributions as well as analogous limitations.

Literature on ideation is often vague, and does not proof the assumed ideation function by empirical studies. To the best of our knowledge, our study is the first one that uses a repeated measures design to analyze an ideation contests in a step-by-step manner. Furthermore we make use of methodologies which, in our opinion, are underrepresented in the rapidly growing research field of online communities. The results from applying our procedure model to a real world dataset suggest that ideation theory might not be sufficient in explaining the process (ideation functions) of modern online ideation contests.

Contrary to [3] conjecture that the value of ideas in ideation will increase over time, we find that increasing time itself changes the context of ideation. Ideas which are submitted during the second half of an ideation contests find utterly different starting conditions and hence, are hardly comparable with ideas that are submitted in the early stage. In terms of novelty we find a degressive collinear ideation curve within our example. The reason for this is that an early brainstorming phase is highly productive and forecloses the chance of later ideas to find a concept that has not yet been mentioned. Albeit this seems to be frustrating in terms of ideation novelty, late ideas still give the impression of being highly useful in terms of feasibility. As the nature of ideation changes within the second half, ideas base on existing concepts and enhance them, mostly by recombining and drawing analogies. Hence, in our case we are also not able to provide evidence to the theories of groupthink [25] or the tipping point [27]. Additionally, our ideation function is converse to the one conjectured by bounded ideation theory [15]. The difference we suggest is constituted by the fact that in modern online ideation contests, solvers mostly are competitors, not collaborators. In that manner, taking a solver’s perspective, theories like tipping points, bounded ideation theory, collective intelligence or groupthink do not provide helpful strategies to get an idea rewarded.
As a result from our survey, we argue that the often-quoted quality vs. quantity discussion [6], [14], [15], [20], [21], [24], [48], [57], [60], which essentially implies that it requires high amounts of diverse and autonomous ideas to include highly valuable good ideas or even breed a collective intelligence, might not be the best way of understanding the problem. Of course, we can find evidence for such assumption, as rewarded ideas are also a product of previous ideas. But at the same time such quality vs. quantity conjecture misses the point of ideation. Our study suggests that albeit a seeker’s task description, solvers address different quality criteria throughout a single ideation contest. Within hundreds of solvers participating, chances to score high in quality criteria such as novelty or feasibility are not uniformly distributed over time. As a result solvers that participate early in an ideation contest might be more likely to briefly search what kind of concepts are already mentioned and simply add a different concept to be the first that has mentioned it. Dealing with saturation in terms of concepts, solvers that enter during later stages have to scrutinize previous ideas to be able to differentiate themselves. Therewith their task is different. Eventually their only chance to stand out is a particularly clever combination of concepts. Hence, as a plurality of follow this strategy, the amount of ideas, words as well as the network density rises. The result of this process is that researchers as well as seekers perceive ideation contests as a whole like noise from which it is hard to detect the signals [83].

Consequently, our study draws three final conclusions. The first is, that in order to understand ideation, research has to analyze ideation. This addresses the mentioned research gap of treating ideation contests as a black box. The consequence is that surveys among solvers, or their perception of contests characteristics, are less significant than the real behavior and action happening inside an ideation contest. As we’ve shown, there is a plurality of yet only occasionally used methods, which can be applied to measure ideation processes. Still, like theories, such methods must undergo rigorous empirical testing to see if propositions made are useful. Our second conclusion is, that the discussion about where good ideas comes from might has to take a step back and concentrate on the quality defining criteria separately. Researcher’s common quality criteria, which often are a measure of aggregated characteristics like novelty, feasibility, relevance or elaboration, might be to manifold and of no avail for ideation contests. Instead of searching for magic bullets, crowdsourcing platforms or research could split ideation contests into different phases and aligning them with different quality criteria.

Therefore our third conclusion is that crowdsourcing platforms might try to offer ideation contests in a bit by bit fashion. As structuring process steps is a common approach (see [12], [94]), the ideation function from our study reasons an ideation contest to run as following: An ideation contests could start by a broad task description by the seeker and a few brainstorming days. The assumption is, that as there will be an early mover crowd, this phase will already collect obvious ideas, analogies, and curios
concepts. Maybe limiting the amount instead of the timespan might be a helpful adjustment. Then the seeker ends this phase, preselects concepts and rewards the bests based on novelty or the ideas ability to surprise, etc. Subsequently, the seeker introduces a new ideation contest using the feedstock from brainstorming as input (maybe in the task description or as compulsory reading to enter the contest). The second ideation contests may intend to achieve higher feasibility, to elaborate on the existing ideas in detail or strengthening their relevance. Some contest parameters such as a required minimum length of an idea or required tagging of ideas by using the concepts from brainstorming might help as guidance and to avoid a cluttered overall result. Our hypothesis is, that such a split might attract different types of solvers. Creative thinker (maybe without a high educational background) will be more efficient within the first part, while gifted tinkerers (say puzzlers) and inventors will enhance the second phase.

Still it can and should not be concluded from a study of a single data sample of the applied procedure model, which combination or combinations of these techniques should be implemented in any particular situation. However, the identification of helpful techniques in this study may guide future efforts to determine ideal combination of data mining and SNA during the analysis of ideation.

F.6 References


CV

Family name: WALTER
First names: Thomas Pierre
Address: Müller Friedberg Str. 8
9000 St. Gallen/ Switzerland

Mobile: +41 76 251 1548
Office: +41 71 224 2791
Email: thomas.walter@unisg.ch

Work Experience

04/2013–ongoing  Senior Consultant @ Namics AG, St. Gallen, Switzerland

08/2012–01/2013  Research Fellow and Lecturer @ City University of Hong Kong, School of Creative Media, Hong Kong.
• Invited to the “CityU International Transition Team” Scheme as Graduate Teaching Assistant (ITT-GTA).
• Research on Crowdsourcing and Collective Intelligence.
• Teaching in “Visual Literacy (English)”, for Bachelor, fall term 2012.

07/2009–12/2012  Research Assistant and Lecturer @ University of St. Gallen, Institute of Information Management, St. Gallen, Switzerland.
• Research on Crowdsourcing and Mobile Business.
• Project-leader in several research and consulting projects, e.g. for T-Systems, Lufthansa or ABB.
• Author of market reports and organizer of international business conferences.

07/2008–06/2009  IT-Consultant @ Lufthansa Systems Business Solutions GmbH, Basel, Switzerland.
• Project Manager for the development of “AirWarder”, software for airfreight capacity management.
• Responsible for release management, collaboration between software development units and conceptual design.

04/2008–07/2008  Student Trainee @ Deutsche Lufthansa AG, Central IT-Strategy Department, Frankfurt, Germany.
• IT architecture design.
• Thesis: “A Framework to Analyze Capabilities of Business Process Automatization”.

07/2006–11/2006  Student Trainee @ LSG Lufthansa Service Holding AG, IT-Application Management, Neu Isenburg, Germany.
• IT support in product development processes.

03/2006–07/2006 Intern @ LSG Lufthansa Service Holding AG, Infrastructure Management, Neu Isenburg, Germany.
• Project collaborator in a change management project.
• Swap of the mobile phone provider for 10,000+ employees.

Education

07/2009–current Ph.D.-Student in Business Innovation @ University of St. Gallen, St. Gallen, Switzerland.
• Doctoral Studies (Ph.D. in Management) at the Institute of Information Management (IWI-HSG) with specialisation in “Business Innovation”.

05/2004–10/2008 Diplom-Student in Mechanical Engineering and Economics @ Technical University of Darmstadt, Germany.
• Specialisation in “Virtual Product Development”, “Business Process Management” and “Information System Engineering”.

08/2005–02/2006 Exchange Student @ Department of Technology Management, Technical University of Eindhoven, Netherlands.
• Specialisation in “Innovation Management”.

10/2002–04/2005 Undergraduate Student in Mechanical Engineering and Economics @ Technical University of Darmstadt, Germany.

Further Information

Scholarships
• e-fellows.net scholarship (2005-2008).
• Erasmus exchange scholarship (2005-2006).
• smART, Lufthansa talent relationship management program for high profile interns (2006-2008).

Interests
• Coffee (certified SCAE Barista, level 2).
• Engineering (lightweight constructions).
• Amateur cycling races (palmares include 7th place in Maratona dles Dolomites, 23rd in Tour Transalp, 22nd in Highlander Radmarathon and 54th in Arlberg Giro).
• Competitive swimming (until 2004, international level, 1st German division from 2000-2004).
List of Publications

Peer reviewed articles


Interviews (given)


Mobile Geschäftsprozesse brauchen die richtige Strategie (by Uwe Kerinnes) in Computerwoche 07/2012, p.1-3.


Firmen haben Apps, aber keine Strategie (by Christoph Hugenschmidt) in inside-it 05/2012, p. 1-3.


**Book Contributions**


**Other pertinent publications**


Thomas P Walter, Keynote: Enterprise Apps, in *KnowTech Konferenz*, Bad Homburg, Germany, September 2010.

2009