2015 11th International Conference on Innovations in Information Technology (IIT)

An Investigation into the Implementation Factors Affecting the Success of Big Data Systems

Patrick Cato
University of Erlangen-Nuremberg
Nuremberg, Germany
patrick.cato@fau.de

Philipp Gölzer
University of Erlangen-Nuremberg
Nuremberg, Germany
philipp.goelzer@fau.de

Walter Demmelhuber
University of Erlangen-Nuremberg
Nuremberg, Germany
walter.demmelhuber@fau.de

***Abstract*—Big Data systems have significantly changed the possibilities and ways of data processing and analysis. Although many practitioners and scholars have written about the benefits of Big Data systems, little research has been conducted on how to succeed with the use of Big Data. The implementation of Big Data systems is a complex undertaking that requires a new technological and organizational approach. This paper seeks to identify the factors that may impact the success of implementing a Big Data system. The factors are identified based on the content review and analysis of industry and academic publications. We identified 21 implementation factors grouped into eight main categories.**

Keywords-Big Data, Information System Success, Implementation Success

I. INTRODUCTION

Big Data is one of the dominant IT topics of the last two years and many software vendors and consultancy firms have published various use cases for the application of Big Data technologies. The positive impact of Big Data on value creation is widely discussed by both practitioners and academics [1]–[3]. Surveys among CIOs in various industries show that investments in Big Data technologies have top priority [4], [5]. Big Data systems fundamentally differ from other technology projects and the implementation of these systems require new technical and organizational approaches [6]. There is considerable knowledge and wisdom among practitioners with regard to implementation of Big Data systems, however, very little academic research has been conducted on the implementation factors affecting Big Data system success [7]. Building on these thoughts, we propose the following research question: What factors impact the implementation success of Big Data systems? To achieve this, we collected publications from both practitioners (such as experience reports and whitepapers) and scholars based on a systematic literature search. We then conducted a content analysis and identified the relevant factors. This paper is structured as follows: Section 2 focuses on the theoretical background, Section 3 introduces the data and methods being used and Section 4 presents the results. Section 5 discusses the results and limitations.

II. THEORETICAL BACKGROUND

A. Big Data

Big Data is one the most discussed IT trends and publications regarding Big Data have grown exponentially since the mid-2000. However, the first publications date back to the 1970s where the term Big Data was mainly used to describe the problem that data is too big to be efficiently processed by conventional data base systems [8].

There is no consensus among practitioners and scholars on the definition of Big Data. Some authors have proposed a problem-centric definition and define Big Data as "datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze" [2, p. 1]. Other authors, such as the research firm Gartner, define Big Data based on characteristics of data. They define Big Data as "high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making" [9]. This definition (also known as the 3V model) is one of the most commonly quoted definitions in both academic and academic publications, however, it is also criticized for a lack of clear distinction between high and low values for the three dimensions [1]. Moreover, several authors have criticized Gartner's 3V model due to incompleteness and proposed further dimensions, such as veracity or value [10]. Veracity refers to the reliability and uncertainty of data which is partly due to issues with data quality. The value dimension of Big Data refers to the information asset, implying that Big Data lends value to the organization [1]. Some authors view Big Data from a business perspective and regard Big Data as a resource that can be used in the value creation process of the organization. Dapp and Heine [11, p. 6] for instance states that "Big Data can become a factor in production and competitiveness that will open up new possibilities for value creation".

Traditional data base systems, based on the relational data model, have reached their performance limit for the processing of Big Data. New database technologies (Big Data systems), such as NoSQL databases, enable the efficient processing of Big Data; the underlying principle of these databases is to divide data into small subsets and to process them concurrently

978-1-4673-8511-4/15/$31.00 ©2015 IEEE

[12]. Furthermore, these database systems are optimized based on alternative data structures (e.g. column-oriented data structure) and alternative hardware architectures (e.g. in-memory storage).

B. Information System Success

The identification and investigation of factors affecting and determining information system success (IS success) has been studied thoroughly in the last three decades and many measures, such as IS usage [13], user information satisfaction [14], decision accuracy [15], have been proposed for the evaluation of IS success. DeLone and McLean [16] have organized the diverse research and proposed a taxonomy with six major dimensions (system quality, information quality, use, user satisfaction, individual impact, organizational impact) based on a review of the literature. Their D&M success model (updated in 2003) has become the most cited framework to measure IS success [17]. As emphasized by DeLone and McLean [18], each IS success model needs to be adapted to the context of the empirical investigation. Consequently, different IS success model were developed for different types of Information Systems, such as Data Warehouse systems [19], Knowledge Management systems [20] etc. Since Big Data systems share notable commonalities with Data Warehouse and Business Intelligence systems [21], the subsequent paragraphs elaborate upon IS success in that application context.

Wixom and Watson [19] and Dinter et al. [22] have developed an IS success model for Data Warehouse and Business Intelligence systems respectively. They distinguish between three layers: Implementation factors, implementation success and system success (Fig. 1). Implementation factors can be regarded as the first layer in the IS success model impacting the implementation success and consequently the overall system success. The second layer refers to different facets of implementation success; for instance success with organizational issues, success with project issues and success with technical issues have been identified as implementation success factors in the Data Warehouse context [19]. The third layer refers to the overall system success and may contain various dimensions of the D&M IS success model, depending on the context of investigation.

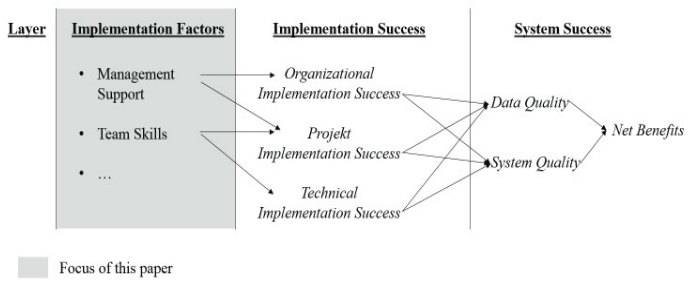

Fig. 1. Simplified model of IS success based on Wixom and Watson [19]

Furthermore, implementation factors can be considered as success factors that needs to be taken into consideration when planning and building a system [22]. For the development of a comprehensive IS success model for the Big Data domain, it needs to be investigated which factors impact the implementation success and consequently determine Big Data System Success. The focus of this paper is the identification of the first layer, namely the implementation factors.

III. DATA COLLECTION AND ANALYTICAL APPROACH

A. Data Collection

Relevant publications were collected based on a systematic literature following the guidelines of Webster and Watson [23]. We screened relevant academic and practical literature to identify appropriate search terms. We then searched in major data bases (EBSCOHost, Google Scholar, IEEE Xplore and Google Search) for papers regarding success factors of Big Data systems using five keywords (Big Data Success, Big Data & Success Factors, Big Data System Success, Big Data & Critical Success Factors, Big Data Project Success). We included industry publications, such as experience reports or case studies, however, PowerPoint presentations were excluded from further review. We identified 89 papers based on limiters in title and abstract. In the next step, we conducted a full text screening and 65 papers had to be excluded since they did not contain any information regarding the research question. Furthermore, we conducted a backward and forward search as recommended by Webster and Watson [23] for the 24 relevant papers. This revealed five additional relevant papers that were added to the sample. In total we found 29 articles.

B. Content Analysis

Content analysis is a scientific research technique to make “replicable and valid interferences from texts” [25, p. 18] to identify patterns, characteristics and trends. Originally developed to analyse speeches and mass media, the content analysis has become an established method in social sciences to analysis written texts (e.g. newspapers) or spoken texts (e.g. interviews). For this purpose, content analysis uses a broad range of techniques, ranging from solely qualitative to highly quantitative methods [26]. Quantitative techniques mainly generate categories based on word frequencies in the text, whereas qualitative techniques generate categories based on a theoretic or inductive guided process [27], [28]. Many studies combine qualitative and quantitative techniques (e.g. descriptive statistics) which is also known as the quantitative analysis of qualitative data [26].

For the rigorous interpretation of the material, clear rules and systematic procedures were developed to analyse the text [25], [27], [28]. The main activity during that systematic process is the condensation and broad description of categories describing the phenomenon under investigation [29]. A key activity to rigorously analysing the material to identify and interpret these categories is the ‘coding’ of the material. Coding refers to the technique that attaches labels to text segments that describe the phenomenon with regard to the research question. If there is enough theory and knowledge about a phenomenon, a deductive category development approach is recommended [28], [29]. Deductive categories are derived from theory and the text material is then analysed using the deductive category system. The inductive category development approach is recommended if there is little or no knowledge about a phenomenon. This approach identifies categories inductively directly from the data [28].

Due to the rigorous rules and procedures, content analysis is a widely accepted method in IS research that has been applied in many studies. Since little research has been done with regard to the factors affecting success of Big Data systems, an inductive procedure is chosen. The next section describes the procedure used in that study to develop the category system.

C. Procedure of Inductive Category Development and Content Analysis

This study applies an inductive category development approach (Fig. 2) based on the elaborations of Mayring [28] and Morris [30].

In the first step, the research question was formulated (What implementation factors impact the implementation success of Big Data systems?). Derived from the research question, we specified the object of content analysis and collected the research material accordingly (refer to Section 2).

In the second step, we have specified the unit of analysis of the coding process. The specification of the unit of analysis has a significant impact on reliability measures [31]. A smaller unit (e.g. word) usually correlates positively with reliability measures, whereas a bigger unit (e.g. theme) impacts reliability negatively. We define 'theme' as unit of analysis, because this approach allows the capturing of sentences-spanning ideas that couldn't be considered if using a smaller unit. To stabilise reliability we did not derive more than one code from one sentence, as recommended by [31], [32].

In the third step, one researcher has reviewed the research material to inductively derive raw categories from the material. A raw category was coded when a certain text segment revealed insights with regard to success factors of Big Data implementations. The categories were built as closely as possible from the material. This step revealed 54 raw categories

In the fourth step, a proximity matrix of the raw categories was built. The proximity matrix ranges from 100 (perfect similarity) to zero (complete independence) between two categories. To reduce the raw categories, we applied a cluster analysis (refer to Table I for reduced category scheme). Based on the 21 categories remaining, we created a codebook and defined clear coding rules and anchor examples.

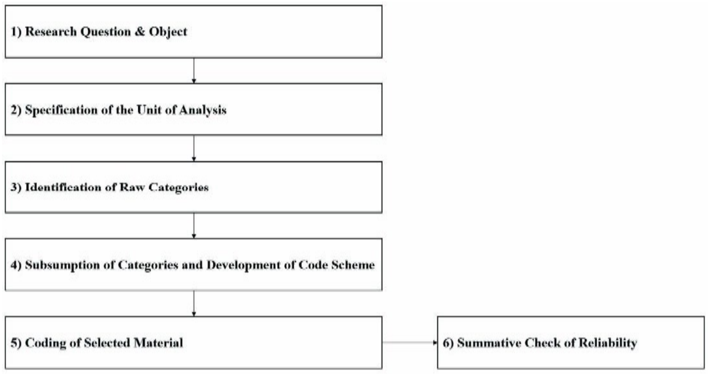

Fig. 2. Procedure of Content Analysis adopted from Mayring [21]

In the fifth step, the author coded the material with the software MAXQDA 11 based on the codebook. In the subsequent step, we conducted a summative check of reliability. Therefore, the co-author was trained during a one day workshop. The codebook was introduced to the co-author and every category was explained. Both author and co-author independently coded one exemplary article. Then, we calculated the intercoder reliability using Holsti's [33] coefficient of reliability that is often used in content analysis studies. The agreement of both coders was 0.65. The disagreement was discussed and two other exemplary articles were coded independently. Intercoder reliability of these articles were above the recommended threshold of 0.7 [28]. Subsequently, the complete research material was coded by the co-author and Holsti's [32] intercoder reliability was calculated for each subcategory. The range of reliability of the subcategories lies between 0.72 and 0.83 and are above the threshold of 0.7.

IV. Results

In total, 21 subcategories (grouped into eight main categories) were identified from the research material based on an inductive coding procedure (refer to Table I for an overview of the category scheme). The first main category *C10 Enterprise Data Management* comprises the subcategories *Master Data Management*, *Data Quality Management*, *Integration Solution* and *Data Security*. *Master Data Management* refers to policies and processes to manage critical data consistently within the organization. *Data Quality Management* refers to roles, processes and responsibilities within the organization to achieve a high level of data quality. *Integration Solution* refers to the architecture to integrate different data sources into the Big Data system, including unstructured and semi-structured data. *Data Security* refers to measures within the organization to protect sensitive data. The main category *C20 Legal* comprises the subcategory *Data Protection* and *Privacy by Design* and refers to the consideration of legal aspects of data protection in the conceptual phase of a Big Data project. The main category *C30 Management* includes the subcategory *Management Priority*, *Sponsorship* as well as *Big Data Strategy Alignment*. *Sponsorship* refers to the financial support of top management, whereas *Management Priority* describe the awareness and clear motivational message of management with regard to Big Data. *Big Data Strategy Alignment* comprises management activities to define a clear vision and strategy how Big Data can be exploited and used within the organization. This also includes the development of a strategy that aligns IT with the business. The main category *C40 Organization* comprises characteristics of the organization that affect the implementation success of Big Data systems within the organization. This includes the subcategory *Close Collaboration between IT & Business* as well as *Organizational Structure*. For instance *Close Collaboration between IT & Business* refers to a high level of cooperation between IT department and business towards the development of Big Data systems. *Organizational Structure* refers to examples found in the research material where to best embed IT and analytical teams in the organization. Depending on the application context, a suitable organizational structure may for instance be cross-functional teams. The main category *C50 Corporate Culture* comprises the subcategory *Data-driven Mindset*, *Change Management Program* and *Freedom for Experimentation*. *Data-driven Mindset* refers to the positive attitude within the organization involving analytics and the exploitation of Big Data. *Change Management Program* describes a systematic change in the business and operational processes as well as the decision styles for use in the exploitation of Big Data. The subcategory *Freedom for Experimentation* refers to the encouragement of employees to experiment and develop creative ideas on how to optimize and enrich operations with the use of Big Data. The main category *C60 Scoping Phase* comprises the subcategories *Focus on Small Projects and Known Questions*, *Specific Business Case*, *Feasibility Study*, *Conduction of Skill Gap Analysis*. The subcategory *Focus on Small Projects and Known Question* comprises examples from the research material that emphasizes starting with a small pilot project that focuses on known business questions. The subcategory *Specific Business Case* refers to the refinement of the exact use case, including the quantification of the business value and clarification of the business question. The subcategory *Feasibility Study* includes test environments and proof of concepts to demonstrate feasibility of the Big Data system. The subcategory *Conduction of Skill Gap Analysis* refers to a systematic comparison of required skills and available skills needed within the organization for the Big Data project.

The main category *C70 Skills* comprises the subcategories *Development of Skills/Training* and *Data Science Skills*. The subcategory *Development of Skills/ Training* refers to educational measures to train and leverage capabilities for the exploitation of Big Data. The subcategory Data Science Skills refers to the existence of interdisciplinary skills within the organization, such as IT, Statistics and Business skills. The main category *C90 Technology* comprises the subcategories *Investments in Technology* and *Adequate Selection of Technology and Algorithms*. The first subcategory *Investments in Technology* refers to sufficient investments in new Big Data technologies. The subcategory *Adequate Selection of Technology and Algorithms* refers to the proper selection of technology for the object under investigation.

The frequencies of codes are shown in brackets (Table 2) and provide an indication of the importance of each factor. The main categories *Scoping Phase*, *Skills* and *Enterprise Data Management* were the most frequently coded main categories in the research material. The subcategory *Specific Business Case, Development of Skills/Training and Data Science Skills* are the three most frequently coded subcategories.

TABLE I. OVERVIEW OF CATEGORY SCHEME

Main Category and Subcategory ***(code frequencies)***		
C10 Enterprise Data Management ***(48)***	**C20 Legal** ***(15)***	**C30 Management** ***(28)***
C11 Master Data Management (20) C12 Data Quality Management (9) C13 Integration Solution (14) C14 Data Security (5)	C21 Data Protection and Privacy by Design *(15)*	C31 Management Priority *(12)* C32 Sponsorship *(3)* C33 Big Data Strategy Alignment *(13)*
C40 Organization ***(24)***	**C50 Corporate Culture** ***(36)***	**C60 Scoping Phase** ***(53)***
C41 Close Collaboration between IT & Business *(12)* C42 Organizational Structure *(12)*	C51 Data-driven Mindset *(19)* C52 Change Management Program *(9)* C53 Freedom for Experimentation *(8)*	C61 Focus on Small Projects and Known Questions *(8)* C62 Specified Business Case *(30)* C63 Feasibility Study *(9)* C64 Conduction of Skill Gap Analysis *(6)*
C70 Skills ***(52)***	**C80 Technology** ***(35)***	
C71 Development of Skills/ Training *(26)* C72 Data Science Skills *(26)*	C81 Investments in Technology *(25)* C82 Adequate Selection of Technology and Algorithms *(10)*	

V. DISCUSSION, LIMITATIONS AND FUTURE RESEARCH

This study has identified various implementation factors that influence the implementation success and consequently determine the overall success of a Big Data systems. Our research shows that there is no single factor, but rather there are multiple factors influencing the Big Data phenomenon under investigation. Researchers from the related Business Intelligence domain have identified various factors impacting the successful implementation and adaption of these systems [22], [34]–[36]. Many factors from the category scheme shown in Table 1, such as *Sponsorship* or *Investments in Technology* were already empirically identified in the BI context. However, our research has revealed Big Data-specific implementation factors that were not mentioned in the BI-related literature. These factors are *Data Protection* and *Privacy by Design*, *Big Data Strategy Alignment*, *Data-driven Mindset*, *Focus on*

Small Projects and Questions, *Specific Business Case*, *Feasibility Study*, *Data Science Skills*. In total, our study has revealed 21 implementation factors. The code frequencies may be interpreted as the importance of the categories with regard to the phenomenon under investigation [25] and give an indication of the relevance.

Our research has some limitations that should be acknowledged. Software vendors, consulting firms and scholars regularly publish to Big Data related topics and consequently a new search query may lead to a different research material. One major limitation of this study is the relatively small data set, because only a few publications have specifically addressed the phenomenon under investigation. Another limitation is the use of Holsti's [33] simple coder agreement that is predominantly used in content analysis studies. Holsti's [33] reliability may have some biases, because it doesn't capture the incidentally occurring agreement between two coders. Another limitation is the interpretation of code frequencies as an indicator of importance, because the code frequency can be biased if an article focuses on specific topics. The research topic is mainly discussed by practitioners. Consequently, a significant amount of resources have been spent to identify relevant sources from practitioners using the google search. Thus, we couldn't include the ACM Digital Library, Springer and Science Direct database in our search process and thus the literature search process is of exploratory nature.

REFERENCES

The following sources have been identified through the systematic literature search: [2], [7], [37]–[63].

[1] P. M. Hartmann, M. Zaki, N. Feldmann, and A. Neely, "Big Data for Big Business? A Taxonomy of Data-Driven Business Models Used by Start-Up Firms," *Cambridge Serv. Alliance*, 2014.

[2] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. Byers, "Big Data: The next frontier for innovation, competition, and productivity," *McKinsey Glob. Inst.*, 2011.

[3] P. Cato, P. Gölzer, and W. Demmelhuber, "Impact of Big Data on Value Creation," in *GSTF ICT-BDCS 2015 Proceedings*, 2015.

[4] Economist Intelligence Unit, "Big data: Harnessing a game-changing asset," *Econ.*, 2011.

[5] Infogroup, "Big Data' s Big Step : Analytics Takes Center Stage for Marketers in 2014," *Infogr. Res.*, 2013.

[6] N. Kabir and E. Carayannis, "Big Data , Tacit Knowledge and Organizational Competitiveness," *J. Intell. Stud. Bus.*, vol. 3, pp. 54–62, 2013.

[7] A. Koronios, J. Gao, and S. Selle, "Big Data Project Success – A Meta Analysis," in *PACIS 2014 Proceedings*, 2014.

[8] E. Ularu, F. Puican, A. Apostu, and M. Velicanu, "Perspectives on Big Data and Big Data Analytics," *Database Syst. J.*, vol. 4, no. 2012, 2012.

[9] Gartner, "Big Data," 2012. [Online]. Available: http://www.gartner.com/it- glossary/big-data/. [Accessed: 01-Jan-2015].

[10] P. Zikopoulos, D. DeRoos, D. Parasuraman, T. Deutsch, J. Giles, and D. Corrigan, *Harness the Power of Big Data: The IBM Big Data Platform*. McGraw Hill, 2013.

[11] T. Dapp and V. Heine, "Big data: The untamed force," *Dtsch. Bank Res.*, 2014.

[12] R. Gupta, H. Gupta, and M. Mohania, "Cloud Computing and Big Data Analytics: What Is New from Databases Perspective?," in *Big data analytics : first international conference, BDA 2012, New Delhi, India, December 24 - 26, 2012;*, 2012, pp. 42–61.

[13] M. J. Culnan, "Chauffeured versus End User Access to Commerical databases: The Effects of Task and Individual Differences," *MIS Q.*, vol. 7, no. 1, pp. 55–67, 1983.

[14] J. Baroudi, M. Olson, and B. Ives, "An Empirical Study of the Impact of User Involvement on System Usage and Information Satisfaction," *Commun. ACM*, vol. 29, no. 3, pp. 232–238, 1986.

[15] R. W. Zmud, "An Empirical Investigation of the Dimensionality of the Concept of Information," *Decis. Sci.*, vol. 9, no. 2, pp. 187–195, 1978.

[16] W. H. DeLone and E. McLean, "Information Systems Success: The Quest for the Dependent Variable," *Inf. Syst. Res.*, vol. 3, no. 1, pp. 60–95, 1992.

[17] N. Urbach, S. Smolnik, and G. Riempp, "The State of Research on Information Systems Success - A Review of Existing Multidimensional Approaches," *Bus. Inf. Syst. Eng.*, vol. 1, no. 4, pp. 315–325, Jul. 2009.

[18] W. H. DeLone and E. R. McLean, "The DeLone and McLean Model of Information Systems Success: A Ten-Year Update," *J. Manag. Inf. Syst.*, vol. 19, no. 4, pp. 9–30, 2003.

[19] B. H. Wixom and H. Watson, "An empirical investigation of the factors affecting data warehousing success," *MIS Q.*, vol. 25, no. 1, pp. 17–41, Mar. 2001.

[20] P. F. Clay, A. R. Dennis, and D.-G. Ko, "Factors Affecting the Loyal Use of Knowledge Management Systems," in *HICSS 2005 Proceedings*, 2005.

[21] H. Chen, R. Chiang, and V. Storey, "Business Intelligence and Analytics: From Big Data to Big Impact," *MIS Q.*, vol. 36, no. 4, pp. 1165–1188, 2012.

[22] B. Dinter, C. Schieder, and P. Gluchowski, "Towards a Life Cycle Oriented Business Intelligence Success Model," in *AMCIS 2011 Proceedings*, 2011.

[23] J. Webster and R. Watson, "Analyzing the Past to Prepare for the Future: Writing a Literature Review," *MIS Q.*, vol. 26, no. 2, pp. xiii–xxiii, Jun. 2002.

[24] J. Webster and R. Watson, "Analyzing the Past to Prepare for the Future: Writing a Literature Review," *MIS Quarterlyuarterly*, vol. 26, no. 2, pp. xiii–xxiii, 2002.

[25] K. Krippendorff, *Content Analysis: An Introduction to Its Methodology*, 2nd ed. Thousand Oaks: Sage Publications, 2004.

[26] D. L. Morgan, "Qualitative Content Analysis: A Guide to Paths not Taken," *Qual. Health Res.*, vol. 3, no. 1, pp. 112–121, Feb. 1993.

[27] K. Klenke, *Qualitative research in the study of leadership*. Bingley: Emerald group publishing, 2008.

[28] P. Mayring, "Qualitative Content Analysis," *Forum Qual. Sozialforsch. / Forum Qual. Soc. Res.*, vol. 1, no. 2, 2000.

[29] S. Elo and H. Kyngäs, "The qualitative content analysis process.," *J. Adv. Nurs.*, vol. 62, no. 1, pp. 107–15, Apr. 2008.

[30] R. Morris, "Computerized Content Analysis in Management Research: A Demonstration of Advantages & Limitations," *J. Manage.*, vol. 20, no. 4, pp. 903–931, Aug. 1994.

[31] D. M. Steininger, M. Trenz, and D. J. Veit, "Building Taxonomies in IS and Management – A Systematic Approach Based on Content Analysis," in *International Conference on Wirtschaftsinformatik*, 2013, no. March, pp. 1441–1455.

[32] H. H. Kassarjian, "Content Analysis in Consumer Research," *J. Consum. Res.*, vol. 4, no. 1, pp. 8–18, 1977.

[33] O. R. Holsti, *Content Analysis for the Social Sciences and Humanities*. Reading: Addison-Wesley, 1969.

[34] H. Watson, D. L. Goodhue, and B. H. Wixom, "The benefits of data warehousing: why some organizations realize exceptional payoffs," *Inf. Manag.*, vol. 39, no. 6, pp. 491–502, 2002.

[35] K. Ramamurthy, A. Sen, and A. P. Sinha, "An empirical investigation of the key determinants of data warehouse adoption," *Decis. Support Syst.*, vol. 44, no. 4, pp. 817–841, Mar. 2008.

[36] H.-G. Hwang, C.-Y. Ku, D. C. Yen, and C.-C. Cheng, "Critical factors influencing the adoption of data warehouse technology: a study of the banking industry in Taiwan," *Decis. Support Syst.*, vol. 37, no. 1, pp. 1–21, Apr. 2004.

[37] J. Gao, A. Koronios, and S. Selle, "Towards A Process View on Critical Success Factors in Big Data Analytics Projects," *AMCIS 2015 Proc.*, 2015.

[38] P. Cato, P. Gölzer, and S. Brumm, "Erfolgsfaktoren von Big Data Projekten," 2015. [Online]. Available: http://intelligence.de/news/erfolgsfaktoren-von-big-data-projekten.html. [Accessed: 20-May-2015].

[39] G. Slinger and R. Morrison, "Will Organization Design Be Affected By Big Data?," *J. Organ. Des.*, vol. 3, no. 3, pp. 17–26, 2014.

[40] J. R. Galbraith, "Organization Design Challenges Resulting From Big Data," *J. Organ. Des.*, vol. 3, no. 1, pp. 2–13, 2014.

[41] A. H. Church and S. Dutta, "The Promise of Big Data for OD: Old Wine in New Bottles or the Next Generation of Data-Driven Methods for Change?," *OD Pract.*, vol. 45, no. 4, pp. 23–31, 2013.

[42] S. F. Wamba, S. Akter, A. Edwards, G. Chopin, and D. Gnanzou, "How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study," *Int. J. Prod. Econ.*, vol. 165, pp. 234–246, 2015.

[43] M. Adolph, "Big data, its enablers and standards," *PIK - Prax. der Informationsverarbeitung und Kommun.*, vol. 37, no. 3, pp. 197–204, 2014.

[44] K. Bottles, E. Begoli, and B. Worley, "Understanding the Pros and Cons of Big Data Analytics," *Physician Exec.*, vol. 40, no. 4, pp. 6–12, 2014.

[45] Avanade, "Unlocking the business value of social, mobile, cloud and big data," 2013.

[46] T. H. Davenport, "At the Big Data Crossroads: turning towards a smarter travel experience," *AMADEUS White Pap.*, 2013.

[47] Open Data Center Alliance, "Open Data Center Alliance: Big Data Consumer Guide," *Open Data Cent. Alliance White Pap.*, 2012.

[48] J. Willkommer, M. May, D. Haller, H. Stange, C. Wass, and T. Fasching, "Big Data: Vorsprung durch Wissen," *TechDivision White Pap.*

[49] B. Schmarzo, *Big Data: Understanding How Data Powers Big Business*. Indianapolis, IN: John Wiley & Sons, 2013.

[50] T. Davenport and J. Dyché, "Big Data in Big Companies," *Int. Inst. Anal. White Pap.*, 2013.

[51] S. Conway, U. Homann, and M. Wise, "CIO considerations for Big Data: Obtaining real business value from Big Data," *Microsoft White Pap.*, 2012.

[52] M. Schroeck, R. Shokley, J. Smart, D. Romero-Morales, and P. Tufano, "Analytics: Big Data in der Praxis," *IBM Inst. Bus. Value White Pap.*, 2012.

[53] V. Gopalkrishnan, D. Steier, H. Lewis, J. Guszcza, and J. Lucker, "Big Data 2.0: New business strategies from Big Data," *Deloitte Rev.*, 2013.

[54] A. Hems, A. Soofi, and E. Perez, "How innovative oil and gas companies are using big data to outmaneuver the competition," *Microsoft White Pap.*, 2013.

[55] A. Bhargava, "A Dozen Ways Insurers Can Leverage Big Data for Business Value," *Tata Consult. White Pap.*, 2011.

[56] R. Adduci, D. Blue, G. Chiarello, J. Chickering, D. Mavroyiannis, S. Mirchandani, J. Solimando, and D. Woods, "Big Data : Big Opportunities to Create Business Value," *EMC White Pap.*, 2011.

[57] R. Hammell, "Open data: Driving growth, ingenuity and innovation," *Deloitte White Pap.*, 2012.

[58] BITKOM, "Management von Big-Data-Projekten," 2013.

[59] D. Dutta and I. Bose, "Managing a Big Data project: The case of Ramco Cements Limited," *Int. J. Prod. Econ.*, vol. 165, pp. 293–306, 2015.

[60] R. Conrads, "In sieben Schritten zum erfolgreichen Big-Data-Projekt," *Informatik-Spektrum*, vol. 37, no. 2, pp. 127–131, 2014.

[61] V. Frehe, T. Kleinschmidt, and F. Teuteberg, "Big Data in Logistics – Identifying Potentials through Literature, Case Study and Expert Interview Analyses," *Inform. 2014 Proc.*, pp. 173–186, 2014.

[62] T. J. Gabel and C. Tokarski, "Big Data and Organizational Design: Key Challenges Await the Survey Research Firm," *J. Organ. Des.*, vol. 3, no. 1, pp. 37–45, 2014.

[63] H. V. Jagadish, J. Gehrke, A. Labrinidis, Y. Papakonstantinou, J. M. Patel, R. Ramakrishnan, and C. Shahabi, "Big data and its Technical Challenges," *Commun. ACM*, vol. 57, no. 7, pp. 86–94, 2014.