Journal of Information Technology Management

Print ISSN: 2008-5893

Online ISSN: 2423-5059

Perspectives of Big Data Quality in Smart Service Ecosystems (Quality of Design and Quality of Conformance)

Markus Helfert

Ph.D., Head of Business Informatics Group, Department of Computing, Dublin City University, Dublin, Ireland. E-mail: markus.helfert@lero.ie

Mouzhi Ge

*Corresponding author, Associate Professor, Department of Computer Systems and Communications, Faculty of Informatics, Masaryk University, Brno, Czech Republic. E-mail: mouzhi.ge@muni.cz

Abstract

Despite the increasing importance of data and information quality, current research related to Big Data quality is still limited. It is particularly unknown how to apply previous data quality models to Big Data. In this paper we review Big Data quality research from several perspectives and apply a known quality model with its elements of conformance to specification and design in the context of Big Data. Furthermore, we extend this model and demonstrate it utility by analyzing the impact of three Big Data characteristics such as volume, velocity and variety in the context of smart cities. This paper intends to build a foundation for further empirical research to understand Big Data quality and its implications in the design and execution of smart service ecosystems.

Keywords: Big data quality, Information quality, Smart cities, Service design, Smart services, Data quality model, Smart service ecosystem.

DOI: 10.22059/JITM.2019.72763

Introduction[1]

In an era of Big Data, organizations are dealing with tremendous amount of data. The data is fast moving, dynamic with many changes and interpretations, and can be originated from a range of various sources such as social networks, unstructured data from different websites or raw feeds from sensors. According to estimates, this type of data contains approximately 85% of potentially valuable information (Das, & Kumar, 2013), which is five times larger than the data used in typical enterprises (Inmon, 2006). Hence, new challenges and opportunities arise along with Big Data (Labrinidis, & Jagadish, 2012). There are some systems that are proposed to process Big Data, while the issues in Big Data still need to be addressed under manual intervention (Yang, & Helfert, 2016). Big Data practitioners are however experience a huge number of data quality problems, which can be time-consuming to solve or even lead to incorrect data analytics. Zhang, Zhang, & Yang (2003) Generally, around 80% of the data engineering effort is consumed in relation to data quality issues. If data quality is not appropriately managed, Big Data will result in even more tasks and challenges and in particular in terms of resources. Therefore, we believe that Big Data Quality (BDQ) should be one of the critical issues related to Big Data research and its applications. Big Data creates not only value in financial terms but also in terms of operational and strategic advantages (Haug, & Arlbjørn, 2010). Thus exploring the value of Big Data and its quality management is crucial to the success of world-leading organizations.

Big Data is typically characterized by the increase in volume, velocity and variety (Laney, 2001; Grover, Chiang, Liang, & Zhang, 2018). As a consequence, BDQ can possibly be affected by the typcial characteristics, volume, velocity and variety. Let us illustrate the challenge with an example from a Smart City context. Smart cities applications present us with an excellent example, as they are characterized by Big Data of high volume, velocity and variety. Many sensor data are used for decision making. In this environment, higher data velocity can result in frequent changes in data specification. For example, in a traffic surveillance information system, the traffic camera is taking a photo every 5 minutes (or even more frequent). Let assume that the data specification for the photo quality is set to be 300 dpi. The traffic photo whose resolution is lower than 300 dpi will be considered as low quality data. When the time interval between two photos is less than 2 minutes, the data specification of photo quality may be lowered because of flow of the traffic photos turns to be fluent. Therefore, as this simple case shows, data specification can be affected by the data velocity, in turn BDQ problems can be caused by using obsolete data specifications.

The aim of this paper is to model and analyze BDQ in smart service ecosystems, as well as derive indications for managing the value of Big Data. We believe that the relationship between Big Data characteristics and the value of the Big Data can be connected by BDQ.

---

1. This paper is an extended version of our paper "Big Data Quality - Towards an Explanation Model", presented in the 21st International Conference on Information Quality, ICIQ 2016, Ciudad Real, Spain, June 22-23, 2016

However, how Big Data characteristics affect the value of Big Data is still unknown. This paper therefore investigates the relationship between the three Big Data characteristics from a quality perspective. We have examined how to adapt traditional data quality research model in the context of Big Data. We believe that the Helfert & Heinrich (2003) model that highlights the importance of conformance to specification and quality of design, is an important contribution to data quality research and builds an excellent foundation for this paper in the context of BDQ. Each of the characteristics in Big Data may affect this quality model and accordingly cause different quality problems. As our research shows, it influences the value of Big Data.

The remainder of the paper is structured as follows. Section 2 presents a theoretical grounding for BDQ research in the context of smart cities. We further model the BDQ by incorporating the quality concepts of conformance to specifications and quality of design. Subsequently we analyze and demonstrate the impact of BDQ in the context of smart cities. We finally provide insights, discussions and further research directions on how to manage the value of Big Data by managing BDQ, considering quality of conformance and quality of design.

Theoretical Grounding

In the following we ground our work by reviewing key concepts of data quality, Big Data value chain and data flows in smart cities. We first review quality and Big data and then discus the usage and value of Big Data in a value chain. Subsequently we relate Big Data and data quality, and discuss the relationship.

Quality in Big Data

In order to apply the key concepts of data quality in the context of Big Data, we have reviewed data quality from a number of perspectives. A classic definition of data quality is "fitness for use", i.e. the extent to which some data successfully serves the purposes of the user (Wang, & Strong, 1996). Such a definition implies that the concept is contextual or relative. For instance, dimensions of data quality, such as relevance, believability, or usefulness are highly contextual. However, according to (Watts et al. 2009), models of data quality assessment have tended to ignore the impact of contextual quality on information use and decision outcomes. (Wang, 1998) argued that data producing processes could be viewed as producing data products for data consumers, a view shared by many others. More database technical perspectives on quality were also found (Hoxmeier, 1998; Kim, Choi, Hong, Kim, & Lee, 2003).

In order to quantify and scale BDQ, we have considered the BDQ concept from two perspectives: conformance to specifications and quality of design – following earlier work from (Helfert & Heinrich, 2003; Gilmore, 1974) that defines quality as conformance to specifications. This definition is relatively straightforward and frequently used in manufacturing industries. It facilitates measurement and increases measuring efficiency. Organizations can determine the quality of products by measuring how well the product conforms to an established specification. Also, the measuring procedure can be automatically implemented. However, it fails to capture the customer's view on product performance. To compensate for the disadvantage of this definition, (Gronroos, 1983) defines quality as conformance to design. This definition is especially prevalent in marketing research and the service industries. Following this definition, researchers posit that it is the customer who is the ultimate judge of the quality of a product/service. Thus organizations can make a quick response to market changes. However, it is difficult to measure the extent to which a product/service meets and/or exceeds the customer's expectation. Since different customers may assign different values to product/service attributes, coordinating and unifying the various quality results are the principal difficulties facing this definition. Considering both aspects, we consider high BDQ as the data that is conformed to the data specifications and meet the user's requirements.

Big Data and Value Chain

Many authors refer to Big Data with the characteristics of volume, variety and velocity (Laney, 2001). In this regard, we follow the definition of Big Data analytics as technologies (e.g., database and data mining tools) and techniques (e.g., analytical methods) that an enterprise can employ to analyze large scale and complex data for various applications. It is intended to augment the enterprise performance from various perspectives.

Following the concept of (Porters, 1998) value chain, (Miller, & Mock, 2013) propose a value chain for Big Data. The chain includes three main steps of data discovery, data integration and data exploitation. In the traditional view in Data Quality this presents an information manufacturing system, transforming raw data into useful information. (Chaffey, & Wood, 2005) propose a similar model that focuses on the transformation from data to information to knowledge to action and then to results (DIKAR Model). This view resonates to the perspective to view an information manufacturing system (Ge, & Helfert, 2008), and provides a key foundation of data and information quality research.

Usage and Value of Big Data

Following the Big Data Value Chain and information manufacturing perspective, we view data quality from data gathering to its final usage. We follow a framework that we developed as an integrated framework for Information Systems/Information Technology (IS/IT) business value from an information perspective (Borek, Helfert, Ge, & Parlikad, 2011). It relates resources and capabilities to IS/IT utilization in form of decisions and business value. The data usage experience as an intangible asset is divided into two types: that of internal data and that of external data (Kwon, Lee, & Shin, 2014). The internal data refers to any data that are produced internally by a firm as a direct or indirect result of business operations. Those regarding employees, products and services, the production line, management decisions, customer profiles and transaction records, and corporate resources are representative types. External data are obtained from sources over which a firm has little or no control such as additional customer information, the market, competitors, macroeconomics, and those of the firm's natural environment. In the context of Big Data analytics, using such external information may be of high value for corporate decision-making (Chen, Chiang, & Storey, 2012). Since previous publications have indicated positive correlations between high data quality and the value of information (Chen et al., 2012), the value of Big Data can be implied by the impact of BDQ (see Figure 1).

Big Data Quality Model

In order to develop a theoretical BDQ model, we build on work from (Helfert & Heinrich, 2003) who have proposed a model to describe the impact of DQ on customer relationships. They specify then quality of design and quality of conformance. They propose a (standardized) quality function of data user u at time t to describe the quality of design as $Q_{t,u}^{design}$ (I_t^{spec}, $I_{t,u}^{demand}$) $\in$ (0;1), whereby the value 0 represents no quality and the value 1 represents maximum quality. Second, the other quality function $Q_t^{conform}$ (I_t^{spec}, I_t^{supply}) $\in$ (0;1) describes the quality of conformance between specification and data provided. This function is independent from the data user, whereby the value 0 represents no quality and the value 1 represents maximum quality. In other words, $Q_{t,u}^{design}$ describes the gathering of user requirements thus user dependent, and $Q_t^{conform}$ the implementation and operations of the information system. In this way, data quality management aims to consolidate the best possible the requirements from various users fit into a specification and the best possible information system fulfills the specification. By adopting this quality model, we have proposed our fist conceptual model to describe impacts of BDQ (see Figure 1).

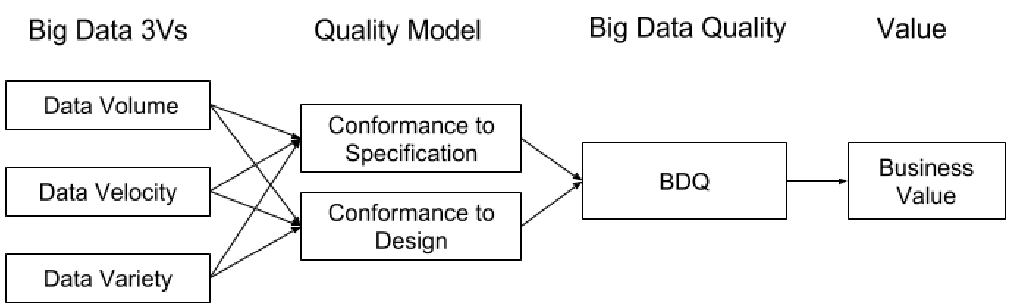

Figure 1. Conceptual model for impact of Big Data Quality

In general, it can be assumed that increasing I_t^{spec} results in higher Q_t^{design} and increasing $I_{t,u}^{demand}$ results in lower Q_t^{design} (exceptions have to be considered at a later stage). Similarly, this applies to quality of conformance $Q_t^{conform}$, whereby increasing I_t^{spec} results in lower $Q_t^{conform}$ and increasing I_t^{supply} results in higher $Q_t^{conform}$. Having formalized the two elements of data quality, data quality management objective function is to maximize the total quality Q_t^{total} over all application areas, which can be described with the optimization variables I_t^{spec}, I_t^{supply} and $I_{t,u}^{demand}$ (Helfert & Heinrich, 2003).

We adopted this model and analyzed the effects of the 3 characteristics of Big Data (Volume, Velocity and Variety) on the quality function. Hypothetically we have postulated the following relationships in Table 1. In this article we do not consider the influence of data demand I_t^{demond} as we assume that it is predetermined. As our discussion illustrates, direct BDQ improvements can be achieved by:

a) an optimization of the specification I_t^{spec} or
b) an (qualitative) increasing of the data provided I_t^{supply}.

In case (a), it improves both quality of design and conformance for Big Data, it requires more sophisticated data specification design and more frequent updates. This can be done by precisely capturing the data demand in a structured way and regularly updating the specification based on the data velocity. For case (b), it includes measurements for increasing the quality of conformance. However, enlarging Big Data does not necessarily increase the overall BDQ. Certain dimensions such as value-added and concise presentation of the BDQ seem to be more important. In future research our theoretical concept can help to illustrate this effect related to data quality criteria.

Data Flow in Smart City Ecosystems

Smart Cities are innovative cities that use ICT to achieve certain objectives, such as increased efficiencies in urban operations and services as well as improving the quality of life of their citizens. Smart Cities can also be seen as entities, with organizational aspects, governance and innovation capabilities. Smart governance requires metropolitan governments to adopt a set of guiding principles to shape city growth and provide patterns for internal and external stakeholder relations. Therefore, Smart Cities can be seen as a multi-layered and multidimensional issue (Anttiroiko, Valkama, & Bailey, 2014).

Various information systems and services place more emphasis on data governance and regulations within the public sector. Anttiroiko et al. (2014) point out that a Smart City is a city with smooth information processes, facilitation of creativity and innovativeness, and smart and sustainable solutions promoted through service platforms. The fundamental idea behind this approach is that smart information and communication systems are needed to build smart social and public systems, which help to achieve the goals within cities and help improve urban life. However, many case studies show that Smart Cities are difficult to realize. Cities are complex using many individual systems, involving many stakeholders and aiming to fulfill multiple aims and goals. How to integrate, plan and maintain these various

systems is yet an open challenge. At the same time, cities are slowly moving to the adoption of smarter technologies and thus transformational, and planning aspects are important.
Designing a smart service is a system engineering approach of determining the required enterprise capabilities and subsequently designing the organization, processes, services, information, and technologies to provide those capabilities (Giachetti, 2010). To manage and organize the complexity, we are utilizing models as logical artifacts and representations, usually described as ecosystems. Therefore, ecosystems are used to construct blueprints of an enterprise for organizing system components, interfaces, processes, services and business capabilities and much more. Ecosystems in the wider context of information systems are often used to model aspects of a system, especially a computer, network, software, application, services, business, and project-development ecosystems. By adopting the (IEEE Standard 2007), smart city ecosystem can be defined as the compressive organization of a system that includes its components, their relationships to each other, and to the environment, and the principles guiding its design and evolution.

Impact of Big Data quality in smart cities

In order to position Big Data in the smart service ecosystems, in this section we first revisit the smart service ecosystem, which is the main pipeline in Smart Cities. There are many definitions and constructs associated with Smart Cities. Due to the emphasis on Information and Communication Technologies (ICT) in our work, we have adopted the definition of a Smart City from International Telecommunications Union as: “An innovative city that uses ICTs and other means to improve quality of life, efficiency of urban operations and services, and competitiveness, while ensuring that it meets the needs of present and future generations with respect to economic, social, environmental as well as cultural aspects” (ITU, 2014).

Nowadays, the global population have become more urban than rural, for example, current forecasts indicate that global metropolitan area populations will increase by 84% to 6.3 billion by 2050, continuing the trend of urbanization. This results in the challenges of broadly meeting the demand of modern societies to ensure the quality of life, sustainability, and economic growth. Proposals to address these challenges with technology are usually associated with the term ‘Smart City’. In addition to technological advances, many emerging service industries create greater competitive advantages to be cost-effective and innovative (Anttiroiko et al., 2014). Among others, transportation, and environment, major bases of urban planning include the implementation of information systems, and providing a smart ICT environment.

With the rapid development of urban areas, we describe an ecosystem for smart service, which is developed to provide an integrated paradigm for considering the complex systems that deliver smart services to a geographic region and its constituent political entities. The ecosystem illustrates layers of components within domains that are combined to provide services for the enterprise; those services will be then consume, governed and supported by the stakeholders, and information flows which traverse the metropolitan enterprises and facilitate interaction with the environment. Figure 2 presents a high-level, conceptual view of the ecosystem of smart services. Data flow across different domains is particularly important to smart services (Pourzolfaghar, Bastidas, & Helfert, 2019).

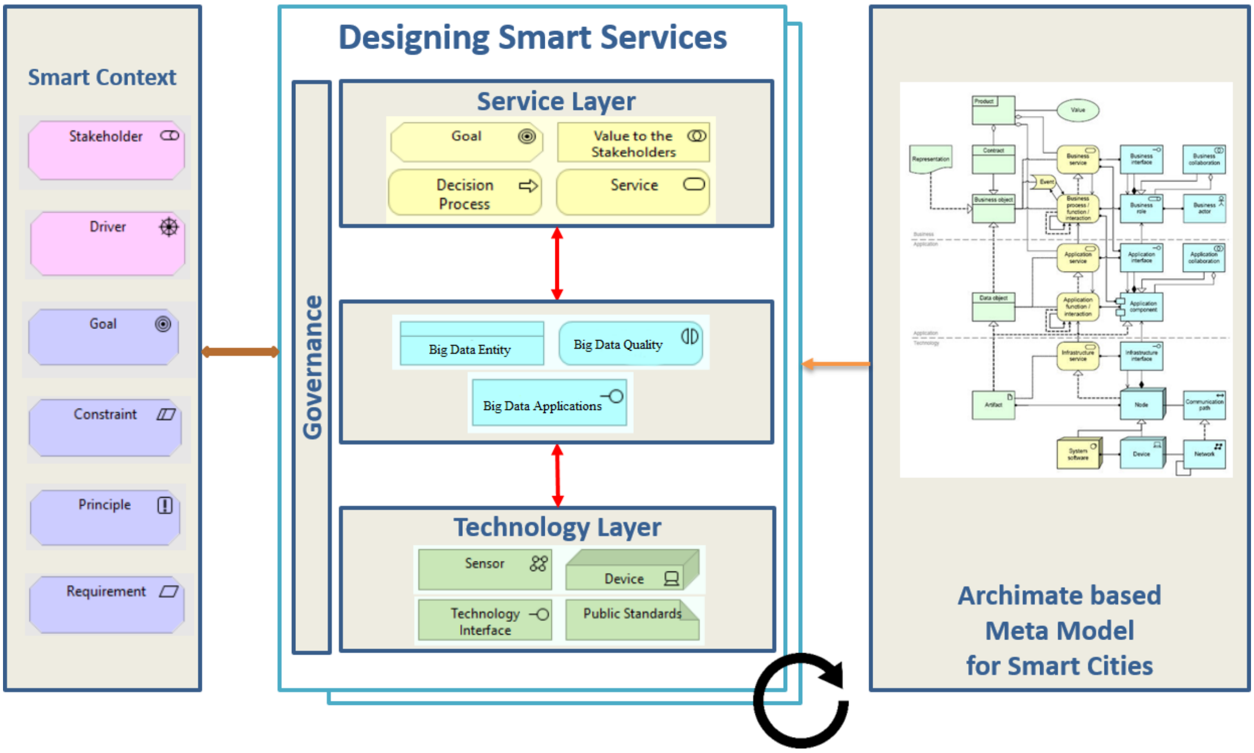

Figure 2. Smart Service Ecosystem (Source: http://scrita.lero.ie/)

Understanding the movement of information is fundamental to understanding and describing operational relationships between the urban enterprise and its environment. Further, understanding and managing data flows are essential to the evolution of digital services in the urban enterprise to efficiently shape and rebalance the service portfolio with changing stakeholder and environmental requirements. An example of capturing relevant social media content as well as direct feedback from service performance can provide useful metrics for service and process improvement programs, and to inform service portfolio balancing programs.

It can be seen that Big Data serves as layer that connects technologies and services. Thus in smart cities, Big Data not necessarily just means I have more data I_t^{supply} – thus better, but actually it depends on the data specification as well as data demand. Increasing the data volume can fulfill the data shortage or overfill the data according to the I_t^{spec} . The data velocity is the speed at which the data is created, stored, analyzed and visualized. Compared to traditional data quality that may only design the I_t^{spec} once, BDQ may require more updates on data specification. This can effect of the life cycle of data timeliness. Furthermore, variety means the different types of data. Once the data variety increases, the data consistency becomes more dynamic and determining the consistency can be also complicated. From our discussion, it can be seen that when measuring BDQ, the data quality criteria that are used to measure the traditional data quality will vary. Therefore, it is critical to consider the feasibility of BDQ model when using Big Data analytics to create business value (Ge, O'Brien, & Helfert, 2017).

Table 1. Impact of Data Quality in Big Data

Big Data	I_t^{supply}	I_t^{spec}	**Impact on BDQ**
Volume Increase	Increase	-	$Q_t^{conform}$ more challenging
Velocity Increase	-	Need for more I_t^{spec} updates	$Q_t^{conform}$ and Q_t^{design} more challenging
Variety Increase	-	Complexity of I_t^{spec} increases	Q_t^{design} more challenging

Table 1 shows the proposition that when we increase the Volume of Big Data, the data supply also increases. As a consequence, assuring $Q_t^{conform}$ becomes more challenging. In our example related to Smart Cities, as the volume of traffic sensor data increases, quality assurance mechanisms and data cleansing have to be increased accordingly.

With the increase in the Velocity of Big Data, we argue that the data specification should become more dynamic. This is due to the fact of update delay. In turn $Q_t^{conform}$ and Q_t^{design} are more challenging to address. In our example related to Smart Cities, when the velocity of traffic sensor data increases, the quality requirements of the data could also be changed such less resolution with pictures and less strict between data capturing intervals, all the changes should be reflected in the data specification. However, when the velocity of traffic sensor data changes, there might be a delay for updating the data specification, and the system may filter out the useful traffic data with the current data velocity.

Further if the Variety of Big Data is increased, it will result in a more complicated I_t^{spec} and designing the Q_t^{design} gets more challenging. In our example related to Smart Cities, the traffic data can be obtained from sensors, traffic cameras, driver's report over the telephone or Internet, or certain notification from a construction site. The data from different sources with different formats can be either structured or unstructured. Thus when the variety increases, I_t^{spec} becomes more complicated and Q_t^{design} gets more challenging due to potential data inconsistency.

In the context of smart cities, we can observe that when the three characteristics of Big Data change, how they impact the quality of conformance and design. It provides a theoretical guideline for Big Data practitioners to assure the BDQ.

Conclusions

In this paper we have presented a data quality model in the context of Big Data. We have described the concept of BDQ and the Big Data value chain. Many of the data quality dimensions have been discussed in the literature (Wang & Strong, 1998), however little research or insights into dimensions and DQ in the Big Data has been conducted yet. Thus some dimensions in DQ like completeness, timeliness, need to be re-considered or re-defined for the context of Big Data. The issues have been related to theoretical perspectives of data quality and the resource-based view on organizations. We have introduced a conceptual model that differentiates between quality of design and quality of conformance. This model has been applied and described within the context of BDQ. We argue that the essential Vs of Big Data (Volume, Velocity and Variety) impact Data Quality and in turn the Value of Big Data. By applying the theoretical model in smart cities and services, we conclude the following:

- Research is often focused on designing services, however usually do not consider the implementation of the services in an ecosystem.
- Therefore, Quality of Design and Quality of Conformance cannot be viewed in isolation.
- Our model allows to analyze both aspects and understand its relation. This in turn helps to design services of high quality that at the same time can be executed with a high level of conformance quality.

We believe that this research is the first contribution that highlights the relation between quality of conformance and quality of design, and applies this into Smart Service Ecosystems. It has both practical contributions and theoretical value. The theoretical value is grounded in the formalization of two distinct but related quality aspects (quality of design and quality of conformance). The practical value results from its application to complex services systems, as illustrated in this paper to a smart cities service ecosystem.

In our further research we aim to develop a simulation environment that allows us to model, refine and test this theoretical model in the context of Smart City Ecosystems. We aim to model instances of architectural representations and analyze the impact of various levels of design quality of services and various levels of conformance quality. We aim to extend our simulation approach for data manufacturing systems and use virtual machines to build a Big Data infrastructure that allows us to test the proposition in Table 1. Virtual machines will act as data sources generating large amount of "real-time" sensor data. We are able to set parameters such as volume, velocity and variety for the data generation. Other virtual machines will be used to integrate data and analyze the data by, for example, a Hadoop cluster and analytics tools. Further research should also aim to verify our quality model, and indeed further research that differentiates quality of design clearly form quality of conformance is required.

Acknowledgement: This work was supported, in part, by Science Foundation Ireland grant 13/RC/2094 and co-funded under the European Regional Development Fund through the Southern & Eastern Regional Operational Programme to Lero - the Irish Software Research Centre (www.lero.ie).

References

Anttiroiko, A., Valkama, P., & Bailey, S. (2014). Smart Cities in the New Service Economy: Building Platforms for Smart Services, *AI & Society*, 29(3), 323–334.

Borek, A., Helfert, M., Ge, M., & Parlikad, A.K.N. (2011). An information oriented framework for relating IS/IT resources and business value. *In Proceedings of the International Conference on Enterprise Information Systems (ICEIS).* Beijing, China.

Chaffey, D., & Wood, S. (2005). Business Information Management: Improving Performance Using *Information Systems.* Pearson Education Ltd., Upper Saddle River.

Chen, H., Chiang, R.H.L., & Storey, V.C. (2012). Business intelligence and analytics: from Big Data to big impact. *MIS Quarterly*, 36(4), 1165–1188.

Das, T.K., & Kumar, P.M. (2013). Big data analytics: A framework for unstructured data analysis. *International Journal of Engineering Science & Technology*, 5(1), 153-156.

Gilmore, H.L. (1974). Product conformance cost. *Quality Progress*, 7, 16-19.

Gronroos, C. (1983). *Strategic management and marketing in the service sector*. Marketing Science Institute Massachusetts, USA. Available at: https://www.msi.org/reports/strategic-management-and-marketing-in-the-service-sector/.

Giachetti, R. (2010). *Design of Enterprise Systems: Theory, Architecture, and Methods*. 1st edition. Boca Raton, CRC Press.

Ge, M., & Helfert, M. (2008). Data and Information Quality Assessment in Information Manufacturing Systems. *International Conference on Business Information Systems*, pp. 380-389.

Ge, M., O'Brien, T., & Helfert, M. (2017). Predicting Data Quality Success - The Bullwhip Effect in Data Quality. *Perspectives in Business Informatics Research - 16th International Conference*, BIR 2017, Copenhagen, Denmark, August 28-30, 2017.

Grover, V., Chiang, R.H., Liang, T.P., & Zhang, D. (2018). Creating strategic business value from big data analytics: A research framework. *Journal of Management Information Systems*, 35(2), 388-423.

Helfert, M., & Heinrich, B. (2003). Analyzing Data Quality Investments in CRM – A model-based approach. *In Eppler, M.J., Helfert, M (eds): Eighth International Conference on Information Quality (IQ 2003), 7th to 9th November*, MIT Sloan School of Management, Cambridge, pp. 80-95.

Hoxmeier, J.A. (1998). Typology of database quality factors. *Software Quality Journal*, 7(3-4), 179-93.

Haug, A., & Arlbjørn, J.S. (2010). Barriers to master data quality. *Journal of Enterprise Information Management*, 24(3), 288-303.

IEEE Standard (2007). *Systems and software engineering. Recommended practice for architectural description of software-intensive systems*, IEEE.

Inmon, B. (2006). DW 2.0: Architecture for the Next Generation of Data Warehousing. *Information Management*, 16(4), p.8.

ITU (2014). Smart sustainable cities: An analysis of definitions, Focus Group Technical Report. https://www.itu.int/en/ITU-T/focusgroups/ssc/Documents/Approved_Deliverables/TR-Definitions.docx

Kim, W., Choi, B.J., Hong, E.K., Kim, SK., & Lee, D. (2003). A taxonomy of dirty data. *Data Mining and Knowledge Discovery*, 7(1), 81-99.

Kwon, O., Lee, N., & Shin, B. (2014). Data quality management, data usage, experience and acquisition intention of Big Data analytics. *International Journal of Information Management*, 34, 387–394.

Labrinidis, A., & Jagadish, H.V. (2012). Challenges and opportunities with big data. *Proceedings of the VLDB Endowment*, 5(12), 2032-2033.

Laney, D. (2001). *3D Data Management: Controlling Data Volume, Velocity, and Variety, Technical report.* META Group.

Miller, H.G., & Mork, P. (2013). From Data to Decisions: A Value Chain for Big Data. *IT Professional*, 15(1), 57-59.

Pourzolfaghar, Z., Bastidas, V., & Helfert, M. (2019). Standardisation of Enterprise Architecture Development for Smart Cities. *Journal of the Knowledge Economy*. https://doi.org/10.1007/s13132-019-00601-8

Zhang, S., Zhang, C., & Yang, Q. (2003). Data preparation for data mining. *Applied artificial intelligence*, 17(5-6), 375-381.

Wang, R.Y., & Strong, D.M. (1996). Beyond accuracy: what data quality means to data consumers. *Journal of Management Information Systems*, 12(4), 5-34.

Wang, R.Y. (1998). A product perspective on total data quality management. *Communications of the ACM*, 41(2), 58-65.

Yang, Q., & Helfert, M. (2016). Revisiting arguments for a three layered data warehousing architecture in the context of the Hadoop platform. *In: The 6th International Conference on Cloud Computing and Services Science (CLOSER 2016)*, 23-25 Apr 2016, Roma, Italy. ISBN 978-989-758-182-3.

Bibliographic information of this paper for citing:

Helfert, Markus, & Ge, Mouzhi (2018). Perspectives of Big Data Quality in Smart Service Ecosystems (Quality of Design and Quality of Conformance). *Journal of Information Technology Management*, 10(4), 72-83.