



# Overview of Data Quality: Examining the Dimensions, Antecedents, and Impacts of Data Quality

Jingran Wang<sup>2</sup> · Yi Liu<sup>3</sup> · Peigong Li<sup>1</sup> · Zhenxing Lin<sup>1</sup> · Stavros Sindakis<sup>4</sup> · Sakshi Aggarwal<sup>5</sup>

Received: 8 August 2022 / Accepted: 20 December 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

## Abstract

Competition in the business world is fierce, and poor decisions can bring disaster to firms, especially in the big data era. Decision quality is determined by data quality, which refers to the degree of data usability. Data is the most valuable resource in the twenty-first century. The open data (OD) movement offers publicly accessible data for the growth of a knowledge-based society. As a result, the idea of OD is a valuable information technology (IT) instrument for promoting personal, societal, and economic growth. Users must control the level of OD in their practices in order to advance these processes globally. Without considering data conformity with norms, standards, and other criteria, what use is it to use data in science or practice only for the sake of using it? This article provides an overview of the dimensions, subdimensions, and metrics utilized in research publications on OD evaluation. To better understand data quality, we review the literature on data quality studies in information systems. We identify the data quality dimensions, antecedents, and their impacts. In this study, the notion of “Data Analytics Competency” is developed and validated as a five-dimensional formative measure (i.e., data quality, the bigness of data, analytical skills, domain knowledge, and tool sophistication) and its effect on corporate decision-making performance is experimentally examined (i.e., decision quality and decision efficiency). By doing so, we provide several research suggestions, which information system (IS) researchers can leverage when investigating future research in data quality.

**Keywords** Big data analytics · Data quality · Decision-making · Open data · Economic growth · IT-based resources

---

✉ Zhenxing Lin  
linzx@lixin.edu.cn

Extended author information available on the last page of the article

## Introduction

Competition in the business world is fierce, and poor decisions can bring disaster to firms. For example, Nokia's leadership decline in the telecommunications industry resulted from its overestimated brand strength and continued instances that its superior hardware design would win over users long after the iPhone's release (Surowiecki, 2013). Making the right decisions leads to better performance (Goll & Rasheed, 2005; Zouari & Abdelhedi, 2021).

Knowledge is a foundational value in our society. Data must be free and open since they are a fundamental requirement for knowledge discovery. In terms of science and application, the open data idea is still in its infancy. The development of open government lies at the heart of this political and economic movement. The President's Memorandum on Transparency and Open Government, which launched the US open data project in 2009, was followed by the UK government's open data program, which was established in 2011. While public sectors host the bulk of open data activities, open data extends beyond "open government" to include areas such as science, economics, and culture. Open data is also becoming more significant in research and has the ability to enhance public institutional governance (Kanwar & Sanjeeva, 2022; Khan et al., 2021; Šlibar et al., 2021). Thus, open data may be viewed from various angles, providing a range of direct and indirect advantages. For example, the economic perspective makes the case that open data-based innovation promotes economic expansion. The political and strategic viewpoints heavily emphasize political concerns like security and privacy. The social angle focuses on the advantages of data usage for society. The social perspective also looks at how all citizens might see the advantages of open data (Danish et al., 2019; Šlibar et al., 2021).

As was previously said, numerous research discovered that open data activities strive to promote societal ideals and advantages. The following highlights a few instances of social, political, and economic benefits. More openness, increased citizen engagement and empowerment, public trust in government, new government services for citizens, creative social services, improved policy-making procedures, and modeling knowledge advancements are all results of political and social gains. Additionally, there are a number of economic advantages, such as increased economic growth, increased competitiveness, increased innovation, development of new goods and services, and the emergence of new industries that add to the economy (Cho et al., 2021; Ebabu Engidaw, 2021; Šlibar et al., 2021).

In the big data era, data-driven forecasting, real-time analytics, and performance management tools are aspects of next-generation decision support systems (Hosack et al., 2012). A high-quality decision based on data analytics can help companies gain a sustained competitive advantage (Davenport & Harris, 2007; Russo et al., 2015). Data-driven decision-making is a newer form of decision-making. Data-driven decision-making refers to the practice of basing decisions on the analysis of data rather than purely on intuition or experience (Abdouli & Omri, 2021; Provost & Fawcett, 2013). In data-driven decision-making, data is at the core of decision-making and influences the quality of the

decision. The success of data-driven decision-making depends on data quality, which refers to the degree of usable data (Pipino et al., 2002; Price et al., 2013). This research seeks to investigate (1) what kinds of data can be viewed as high-quality, (2) what factors influence data quality, and (3) how data quality influences decision-making.

The scope of the paper revolves around three methodologies used to examine the dimensions of data quality and synthesize those data quality dimensions. The findings in the below section show that data quality has many characteristics, with accuracy, completeness, consistency, timeliness, and relevance considered the most significant ones. Additionally, two important aspects were discovered in the paper that affect data quality, i.e., time constraints and data user experience, which is frequently discussed in the literature review. By doing this, we were able to clearly illustrate the problems with data quality, point out the gaps in the literature, and suggest three key concerns about big data quality.

Moreover, the study's main contributions are beneficial for upcoming academicians and researchers as the literature review emphasizes the benefits of utilizing data analytics tools on firm decision-making performance. There needs to be more research that quantitatively demonstrates the influence of the successful use of data analytics (data analytics competency) on firm decision-making which is fulfilled in our study. This area of research was essential as improving firms' decision-making performance is the overarching goal of data analysis in the field of data analytics. Understanding the factors affecting it is a novel contribution to its field.

Remarkably, the literature review is built by reviewing articles, of which 29 articles related to data quality were considered. By examining the fundamental aspect of data quality and its impact on decision-making and end users, we begin to take the first step towards gaining a more in-depth understanding of the factors that influence data quality. Previous researchers should have focused more on the above areas, and we aim to highlight and enhance the same.

In addition to this, a thorough review of previous works in the same field has helped us track the research gap, and this organized review of previous works was divided into several steps: identifying keywords, analysis of citations, the calibration of the search strings, and the classification of articles by abstracts. Based on all the database searches, we found that these 29 articles best discuss data quality, its dimensions, constructs, and its impact on decision-making and are much more relevant than other articles included in our study. These articles determine the factors that influence data quality, and a framework provided helps illustrate a complete description of the factors affecting data quality.

The paper is organized as follows. The literature review is divided into three sections. In the first section, we review the literature and briefly identify the dimensions of data quality. In the second section, we summarize the antecedents of data quality. In the third section, we summarize the impacts of data quality. We then discuss future opportunities for dimensions of big data quality that have been neglected in the data quality literature. Finally, we propose several research directions for future studies in data quality.

## Literature Review

Data is so essential to modern life that some have referred to it as the “new oil.” A current illustration of the significance of the data is the management of the COVID-19 epidemic. The early detection of the virus, the prediction of its spread, and the evaluation of the effects of lockdowns were all made possible by data gathered from location-based social network posts and mobility records from telecommunications networks, which allowed for data-driven public health decision-making (Dakkak et al., 2021; Shumetie & Watabaji, 2019).

As a result, words like “data-driven organization” or “data-driven services” are beginning to appear with the prefix “data-driven” more frequently. Additionally, according to Google Books Ngram, the word “data-driven” has become increasingly popular during the previous 10 years. Data-driven creation, which has been defined as the organization’s capacity to acquire, analyze, and use data to produce new goods and services, generate efficiency, and achieve a competitive advantage, is a trend that also applies to the development of software (Dakkak et al., 2021; Maradana et al., 2017; Prifti & Alimehmeti, 2017). More and more software organizations are implementing data-driven development techniques to take advantage of the potential that data offers. Facebook, Google, and Microsoft, among other cloud and web-based software companies, have been tracking user behavior, identifying their preferences, and running experiments utilizing data from the production environment. The adoption of data-driven techniques is happening more slowly in software-intensive embedded systems, where businesses are still modernizing to include capabilities for continuous data collection from in-service systems. The majority of embedded systems organizations use ad hoc techniques to meet the demands of the individual, the team, or the customer rather than having a systematic and ongoing method for collecting data from in-service products (Carayannis et al., 2012; Cho et al., 2021; Dakkak et al., 2021; Šlibar et al., 2021). Therefore, Dakkak et al. (2021) discussed the two areas we use to identify the data gathering challenges, which are:

1. *Customer agreement*: The consent between the consumer and the case study organization to obtain and share the information is one of the critical obstacles to data gathering. Since the data is generated by customer-owned products, these products are thought of as producing customer property. Except for basic product configuration data, some clients have tight data-sharing policies that forbid data collection and sharing. Data required during special initiatives like the launch of new goods and features is provided on request, just like data required for fault finding and troubleshooting.

Other clients permit automatic data collection but set restrictions on the types of data that may be gathered, when they can be gathered, how they will be used, and how they will be moved. This is frequently the case with clients who have contracts for services like customer support, optimization, or operations, where data could be used for these reasons and must only be available to those carrying out these tasks. The data itself is now used to evolve these services to become

data-driven, even while consumers with service-specific data collection agreements prohibit the data from being used for continuous software enhancements (Cho et al., 2021; Dakkak et al., 2021).

2. *Technical difficulties:* We have observed some technical challenges related to continuous data collection, including:
  - (a) *Impact on the performance of the product:* While some data can be gathered from in-service products without having any adverse effects on their operations, such as network performance evaluation, other data must be instrumented before collecting due to the adverse effects their collection creates on the product's performance as they require internal resources during the generation and collection times, such as processor and memory power.
  - (b) *Data dependability:* Given the variety of data kinds, it may be misleading to consider one data type in isolation from the quality standpoint. While a single piece of data can be evaluated based on certain quality indicators like integrity, developing a comprehensive picture of data quality necessitates a connection between several data sources (Dakkak et al., 2021; Khan et al., 2021; Šlibar et al., 2021).

## Data Quality Dimensions

Data quality is the core of big data analytics-based decision support. It is not a uni-dimensional concept but a multidimensional concept (Ballou & Pazer, 1985; Pipino et al., 2002). The identified dimensions include accessibility, amount of data, believability, completeness, concise representation, consistent representation, ease of manipulation, free of error, interpretability, objectivity, relevancy, reputation, security, timeliness, understandability, and value-added (Abdouli & Omri, 2021; Pipino et al., 2002). Furthermore, Cho et al. (2021) highlighted that data quality dimensions are constructs used when evaluating data and are criteria or features of data quality that are thought to be crucial for a particular user's task. For instance, completeness (e.g., are measured values present?), conformance (e.g., do data values comply with prescribed requirements and layouts?), and plausibility (e.g., are data values credible?) could all be used to evaluate the quality of data. Since data quality has multiple dimensions, how studies are conducted on the dimensions of data quality and which dimensions are the most popular are two questions we want to review in this section. In the data quality literature, three approaches are commonly used to study data quality dimensions (Wang & Strong, 1996).

The first approach is an *intuitive approach* based on the researchers' past experience or intuitive understanding of what dimensions are essential (Wang & Strong, 1996). The intuitive approach has been used in early studies of data quality (Bailey & Pearson, 1983; Ballou & Pazer, 1985; DeLone & McLean, 1992; Ives et al., 1983; Laudon, 1986; Morey, 1982). For example, Bailey and Pearson (1983) viewed accuracy, timeliness, precision, reliability, currency, and completeness as important dimensions of the data quality of the output information. Ives et al. (1983) viewed relevancy, volume, accuracy, precision, currency, timeliness, and completeness as important dimensions of data quality

for the output information. Ballou and Pazer (1985) also viewed accuracy, completeness, consistency, and timeliness as data quality dimensions. Laudon (1986) used accuracy, completeness, and unambiguousness as essential attributes of data quality included in the information. DeLone and McLean (1992) used accuracy, timeliness, consistency, completeness, relevance, and reliability as data quality dimensions. Studies also argue that inconsistency is important to data quality (Ballou & Tayi, 1999; Bouchoucha & Benammou, 2020). Many studies use an intuitive approach to define data quality dimensions because each study can choose the dimensions relevant to the specific purpose of the study. In other words, the intuitive approach allows scholars to choose specific dimensions based on their research context or purpose.

A second approach is a *theoretical approach* that studies data quality from the perspective of the data manufacturing process. Wang et al. (1995) viewed information systems as data manufacturing systems that work on raw material inputs to produce output material or tangible products. The same can be said of an information system, which acts on raw data input (such as a file, record, single number, report, or spreadsheet) to generate output data or data products (e.g., a corrected mailing list or a sorted file). In some other data manufacturing systems, this data result can be used as raw data. The phrase “data manufacturing” urges academics and industry professionals to look for extra-disciplinary comparisons that can help with knowledge transfer from the context of product assurance to the field of data quality. The phrase “data product” is used to underline that the data output has a value that is passed on to consumers, whether inside or outside the business (Feki & Mnif, 2016; Wang et al., 1995).

From the data manufacturing standpoint, the quality of data products is decided by consumers. In other words, the actual use of data determines the notion of data quality (Wand & Wang, 1996). Thus, Wand and Wang (1996) posit that the analysis of data quality dimensions should be based on four assumptions: (1) information systems can represent real-world systems; (2) information systems design is based on the interpretation of real-world systems; (3) users can infer a view of the real-world systems from the representation created by information systems; (4) only issues related to the internal view are part of the model (Wand & Wang, 1996). Based on the representation, interpretation, inference, and internal view assumptions, they proposed intrinsic data quality dimensions, including complete, unambiguous, meaningful, and correct data (Wand & Wang, 1996). The theoretical approach provided a more detailed and complete set of data quality dimensions, which are natural and inherent to the data product.

A third approach is *empirical*, which focuses on analyzing data quality from the user’s viewpoint. A tenant of the empirical approach is the belief that the quality of the data product is decided by its consumers (Wang & Strong, 1996). One of the representative studies was done by Wang and Strong (1996), who defined the dimensions and evaluation of data quality by collecting information from data consumers (Wang & Strong, 1996). Data has quality in and of itself, according to intrinsic DQ. One of the four dimensions that make up this category is accuracy. Contextual DQ draws attention to the necessity of considering data quality as a component of the job at hand; that is, data must be pertinent, timely, complete, and of an acceptable volume to provide value. The relevance of systems is highlighted by representational DQ and accessibility DQ. To be effective, a system must display data in a form that

is comprehensible, easy to grasp, and consistently expressed (Ghasemaghahi et al., 2018; Ouechtati, 2022; Wang & Strong, 1996). This study argues that a preliminary conceptual framework for data quality should include four aspects: accessible, interpretable, relevant, and accurate. They further refined their model into four dimensions: (1) intrinsic data quality, which means that data should be not only accurate and objective but also believable and reputable, (2) contextual data quality, which means that data quality must be considered within the context of the task, (3) representational data quality, which means that data quality should include both format of data and meaning of data, and (4) accessible data quality, which is also a significant dimension of data quality from the consumer's viewpoint.

Fisher and Kingma (2001) used these dimensions of data quality to analyze the reasons that caused two disasters in US history. In order to explain the role of data quality in the explosion of *the Challenger* spacecraft and the miss-fire caused by *the USS Vincennes*. Accuracy, timeliness, consistency, completeness, relevancy, and fitness for use were used as data quality dimensions (Barkhordari et al., 2019; Fisher & Kingma, 2001). In their study, accuracy means a lack of error between recorded and real-world values. Timeliness means the recorded value is up to date. Completeness is focused on whether all relevant data is recorded. Consistency means data values did not change in different records. Data relevance means data should relate to special issues, and fitness for use means data should serve the user's purpose (Fisher & Kingma, 2001; Strong, 1997; Tayi & Ballou, 1998). Data quality should depend on purpose (Shankaranarayanan & Cai, 2006). This category of data quality is also used in credit risk management. Parssian et al. (2004) viewed information as a product and presented a method to assess data quality for information products. Researchers mainly focused on accuracy and completeness because they thought these two factors were the most important to decision-making (Parssian et al., 2004; Reforgiato Recupero et al., 2016). They viewed information as a product but still evaluated the factors from the user's viewpoint. Studies also evaluated a model for cost-effective data quality management in customer relationship management (CRM) (Even et al., 2010). Moges et al. (2013) argued that completeness, interpretability, reputability, traceability, easily understandable, appropriate-amount, alignment, and concise representation are important dimensions of data quality in credit risk management (Danish et al., 2019; Moges et al., 2013). These studies view data quality dimensions as involving the voice of data consumers. Examining data quality dimensions from the user's point of view is one of the most critical characteristics of empirical approaches (Even et al., 2010; Fisher & Kingma, 2001; Moges et al., 2013; Parssian et al., 2004; Shankaranarayanan & Cai, 2006; Strong, 1997; Tayi & Ballou, 1998).

Intuitive approaches are the easiest to examine data quality dimensions, and theoretical approaches are supported by theory. However, both approaches overlook the user, the most important judge of data quality. Data consumers are the judges that decide whether data is of high quality or poor quality. However, it is difficult to prove that the results gained from empirical approaches are complete and precise through fundamental principles (Wang et al., 1995; Prifti & Alimehmeti, 2017). Based on the studies reviewed, we summarize data quality dimensions and comparative studies (Table 1). The results indicate that completeness, accuracy, timeliness, consistency, and relevance are the top six dimensions of data quality mentioned in studies.



**Table 1** Summary of data quality dimensions

Dimensions of data quality	Studies	Number
Accessibility	(Wang & Strong, 1996) (Prifti & Alimehmeti, 2017)	2
Access security	(Wang & Strong, 1996)	1
Accuracy/precision	(Morey, 1982) (Ives et al., 1983) (Bailey & Pearson, 1983) (Ballou & Pazer, 1985) (Laudon, 1986) (DeLone & McLean, 1992) (Wang & Strong, 1996) (Wang & Wang, 1996) (Fisher & Kingma, 2001) (Parssian et al., 2004) (Šlibar et al., 2021) (Cho et al., 2021)	12
Alignment	(Moges et al., 2013)	1
Appropriate amount of data/volume	(Ives et al., 1983) (Wang & Strong, 1996) (Moges et al., 2013)	3
Believability	(Wang & Strong, 1996) (Hook et al., 2018)	2
Completeness	(Ives et al., 1983) (Bailey & Pearson, 1983) (Ballou & Pazer, 1985) (Laudon, 1986) (Côte-Real et al., 2020) (Ghasemaghaei et al., 2018) (DeLone & McLean, 1992) (Wang & Wang, 1996) (Wang & Strong, 1996) (Fisher & Kingma, 2001) (Parssian et al., 2004) (Shankaranarayanan & Cai, 2006) (Even et al., 2010) (Moges et al., 2013)	14
Concise representation	(Wang & Strong, 1996) (Moges et al., 2013) (Danish et al., 2019)	3
Consistency	(Ballou & Pazer, 1985) (DeLone & McLean, 1992) (Ballou & Tayi, 1999) (Zhuo et al., 2021) (Wang & Strong, 1996) (Fisher & Kingma, 2001)	6
Ease of understanding	(Wang & Strong, 1996) (Moges et al., 2013)	2
Fitness for use	(Fisher & Kingma, 2001)	1
Interpretability	(Wang & Strong, 1996) (Moges et al., 2013)	2



**Table 1** (continued)

Dimensions of data quality	Studies	Number
Meaningful	(Wand & Wang, 1996) (Feki & Mnif, 2016)	2
Objectivity	(Wang & Strong, 1996) (Ouechtati, 2022)	2
Relevance	(Ives et al., 1983) (Bailey & Pearson, 1983) (DeLone & McLean, 1992) (Wang & Strong, 1996) (Fisher & Kingma, 2001)	5
Reliability	(DeLone & McLean, 1992) (Zouari & Abdelhedi, 2021)	2
Reputation	(Wang & Strong, 1996) (Moges et al., 2013)	2
Timeliness/currency	(Ives et al., 1983) (Bailey & Pearson, 1983) (Ballou & Pazer, 1985) (DeLone & McLean, 1992) (Wang & Strong, 1996) (Fisher & Kingma, 2001)	6
Traceability	(Moges et al., 2013)	1
Unambiguous	(Laudon, 1986) (Wand & Wang, 1996)	2
Value-added	(Wang & Strong, 1996) (Prifti & Alimehmeti, 2017)	2

### Factors that Influence Data Quality (Antecedents)

Several studies try to determine the factors that influence data quality. Wang et al. (1995) proposed a framework that included seven elements that influenced data quality: management responsibility, operation and assurance costs, research and development, production, distribution, personal management, and legal (Wang et al., 1995). This framework provides a complete description of the factors influencing data quality, but it is challenging to implement because of its complexity. Ballou and Pazer (1995) studied the accuracy-timeliness tradeoff and argued that accuracy improves with time and will increase data quality (Ballou & Pazer, 1995; Šlibar et al., 2021). Experience also influenced data quality by affecting the usage of incomplete data (Ahituv et al., 1998). If a decision-maker is familiar with the data, the decision-maker may be able to use intuition to compensate for problems (Chengalur-Smith et al., 1999). Also, studies indicated that information overload would affect data quality by reducing data accuracy (Berghel, 1997; Cho et al., 2021). Later, scholars pointed out that information overload, experience level, and time constraints impact data quality by influencing the way decision-makers use the information (e.g., Fisher & Kingma, 2001). The top ten antecedents of data quality have been identified through a literature review related to the antecedents of data quality. Table 2 presents the summary of antecedents of data quality.

## Impact of Data Quality

The impact of data quality on decision-making and the impact of data quality on end users are two main themes. Studies of the impact of data quality on decision-making frequently use the definition of data quality information (DQI), which is a general evaluation of data quality and data sets (Chengalur-Smith et al., 1999; Fisher et al., 2003). After considering the decision environment, Chengalur-Smith et al. (1999) argued that DQI generates different influences on decision-making in different tasks, decision strategies, and the formation of the DQI context. Later, Fisher et al. (2003) presented the influence of experience and time on the use of DQI in the decision-making process. They developed a detailed model to explain the influence factors between DQI and the decision outcome. Through research, Fisher et al. (2003) argued that (1) experts use DQI more frequently than do novices; (2) managerial experience positively influences the usage of DQI, but domain experience did not have an influence on the usage of DQI; (3) DQI would be useful for managers with little domain-specific experience, and training in the use of DQI by experts would be worthwhile; (4) the availability of DQI will have more influence on decision-makers who feel time pressure than decision-makers who do not feel time pressure (Cho et al., 2021; Fisher et al., 2003). According to Price and Shanks (2011), metadata depicting data quality (DQ) can be viewed as DQ tags. They found that DQ tags can not only increase decision time but can also change decision choices. DQ tags are associated

**Table 2** Summary of antecedents of data quality

Antecedents	Reference	Number
Data user experience	(Ahituv et al., 1998) (Chengalur-Smith et al., 1999) (Fisher & Kingma, 2001) (Cho et al., 2021)	4
Distribution	(Wang et al., 1995) (Šlibar et al., 2021)	2
Information overload	(Berghel, 1997) (Fisher & Kingma, 2001)	2
Legal	(Wang et al., 1995) (Šlibar et al., 2021)	2
Management responsibility	(Wang et al., 1995) (Cho et al., 2021)	2
Operation and assurance costs	(Wang et al., 1995) (Cho et al., 2021)	2
Personal management	(Wang et al., 1995) (Cho et al., 2021)	2
Production	(Wang et al., 1995) (Šlibar et al., 2021)	2
Research and development	(Wang et al., 1995)	1
Timeliness/time pressure	(Mock, 1971) (Ballou & Pazer, 1995) (Fisher & Kingma, 2001)	3

with increased cognitive processing in the early decision-making phases, which delays the generation of decision alternatives (Price & Shanks, 2011). Another study on the impact of data quality on decision-making focused on the implementation of data quality management to support decision-making. The data quality management framework was mainly built on the information product view (Ballou et al., 1998; Wang et al., 1998). Total data quality management (TDQM) and information product map (IPMAP) were developed based on the information product view. Studies of data quality management have focused more on context. For example, Shankaranarayanan and Cai (2006) constructed a data quality standard framework for B2B e-commerce. They proposed three-layer solutions based on IPMAP and IP View, including the DQ 9000 quality management standard, the standardized data quality specification metadata, and the third-party DQ certification issuers (Sarkheyli & Sarkheyli, 2019; Shankaranarayanan & Cai, 2006). The representative research on data quality impact on end-user was proposed by Foshay et al. (2007). They argued that end-user metadata impacts user attitudes toward data in databases, and end-user metadata elements strongly influence user attitudes toward data in the warehouse. They have a similar impact as the “Other factors”: data quality, business intelligence tool utility, and user views of training quality. Together with these other characteristics, metadata factors appear to have a considerable impact on attitudes. This discovery is incredibly important. It implies that metadata plays a significant role in determining whether a user will have a favorable opinion of a data warehouse (Dranev et al., 2020; Foshay et al., 2007) (Table 3).

The study’s “Other factors” do not seem to have much of a direct impact on the utilization of data warehouses. Other elements function similarly to the metadata factors, having an indirect impact on use. Perceived data quality, out of all the other criteria, had the most significant impact on users’ views toward data. Therefore, user perceptions about the data available from the warehouse have a significant impact on how valuable and simple the data warehouse is thought to be to use. Thus, the amount of use of the data warehouse is influenced by perceived usefulness and perceived ease of use to a moderately substantial extent. As a result, it would seem that variables other than perceived usefulness and usability have a role in deciding how widely data warehouses are used. Also, the degree to which end-user metadata quality and use influence user attitudes depend critically on the user experience accessing a data warehouse (Foshay et al., 2007; Zhuo et al., 2021).

**Table 3** Summary of data quality impacts

Data quality impacts	Studies	Number
Decision-making	(Chengalur-Smith et al., 1999) (Fisher et al., 2003) (Shankaranarayanan & Cai, 2006) (Price & Shanks, 2011)	4
Impact on end-user	(Foshay et al., 2007) (Zhuo et al., 2021) (Dranev et al., 2020)	3

## Discussion

Through the literature review of these 29 articles related to data quality in the IS field, we found that data quality has multiple dimensions, and that *completeness* (Bailey & Pearson, 1983; Ballou & Pazer, 1985; Côte-Real et al., 2020; DeLone & McLean, 1992; Even et al., 2010; Fisher & Kingma, 2001; Ives et al., 1983; Laudon, 1986; Moges et al., 2013; Parssian et al., 2004; Shankaranarayanan & Cai, 2006; Šlibar et al., 2021; Wand & Wang, 1996; Wang & Strong, 1996; Zouari & Abdelhedi, 2021), *accuracy* (Bailey & Pearson, 1983; Ballou & Pazer, 1985; Dakkak et al., 2021; DeLone & McLean, 1992; Fisher & Kingma, 2001; Ghasemaghaei et al., 2018; Ives et al., 1983; Juddoo & George, 2018; Laudon, 1986; Morey, 1982; Parssian et al., 2004; Safarov, 2019; Wand & Wang, 1996; Wang & Strong, 1996), *timeliness* (Bailey & Pearson, 1983; Ballou & Pazer, 1985; Cho et al., 2021; Côte-Real et al., 2020; DeLone & McLean, 1992; Fisher & Kingma, 2001; Ives et al., 1983; Šlibar et al., 2021; Wang & Strong, 1996), *consistency* (Ballou & Pazer, 1985; Ballou & Tayi, 1999; Cho et al., 2021; Dakkak et al., 2021; DeLone & McLean, 1992; Fisher & Kingma, 2001; Ghasemaghaei et al., 2018; Wang & Strong, 1996), and *relevance* (Bailey & Pearson, 1983; Côte-Real et al., 2020; Dakkak et al., 2021; DeLone & McLean, 1992; Fisher & Kingma, 2001; Ives et al., 1983; Klein et al., 2018; Šlibar et al., 2021; Wang & Strong, 1996) are the top five data quality dimensions mentioned in studies.

However, existing studies focused on multiple dimensions of traditional data quality and did not address the new dimensions of big data quality. In traditional data, timeliness is important; however, one new attribute of big data is its real-time delivery. So, we do not know whether timeliness will still play an important role in the dimensions of big data. Volume is also a new attribute of big data. Three papers address the volume of data (Ives et al., 1983; Moges et al., 2013; Šlibar et al., 2021; Wang & Strong, 1996). One reason that traditional data quality highlights the role of volume is that it is hard to get enough data. However, in the era of big data, there are enormous amounts of data, and volume is no longer a big issue. Therefore, we do not know whether volume will still be an important attribute of big data quality. Value-added is also one of the traditional data quality dimensions (Dakkak et al., 2021; Sarkheyli & Sarkheyli, 2019; Wang & Strong, 1996). Big data's value is still, however, uncertain, whether value will be a new important attribute of big data quality needs to be studied further. Recent studies indicate that volume, variety, velocity, value, and veracity (5 V) are five common characteristics of big data (Cho et al., 2021; Firmani et al., 2019; Gordon, 2013; Hook et al., 2018). Nevertheless, few studies investigated the impacts of the 5 V dimensions on big data quality.

Also, existing studies identified several factors that influenced data quality, such as *time pressure* (Ballou & Pazer, 1995; Cho et al., 2021; Côte-Real et al., 2020; Fisher & Kingma, 2001; Mock, 1971), *data user experience* (Ahituv et al., 1998; Chengalur-Smith et al., 1999; Cho et al., 2021; Fisher & Kingma, 2001), and *information overload* (Berghel, 1997; Fisher & Kingma, 2001; Hook et al., 2018). There are, however, few studies explaining what new factors influence big data quality. For example, existing studies talk about time pressure (Ballou & Pazer, 1995; Fisher & Kingma, 2001; Mock, 1971; Zhuo et al., 2021) more

than information overload (Berghel, 1997; Fisher & Kingma, 2001; Klein et al., 2018; Sarkheyli & Sarkheyli, 2019). But big data volume is expansive, and data is obtained in real-time. This will cause information overload problems more serious than time pressure. Variety dimensions of big data mean that the structures of big data are various, which could cause problems when unstructured data is converted into structured data. Human error or system error may also constitute new factors influencing big data quality. Most research related to data quality considers how data quality impacts decision-making. No studies discussed the unknown impact of big data quality. Recent studies indicate that future decisions will be based on data analytics, and our world is data-driven (Davenport & Harris, 2007; Juddoo & George, 2018; Loveman, 2003).

Based on the literature review and the research gaps identified, we propose several future research directions related to data quality within the big data context. First, future studies on data quality dimensions should focus more on the 5 V dimensions of big data quality to identify new attributes of big data quality. Furthermore, future research should examine possible changes in the other quality dimensions, such as accuracy and timeliness. Secondly, future research should focus on identifying the new impacts of big data quality on decision-making by answering how big data quality influences decision-making and finding other issues related to big data quality (Davenport & Patil, 2012; Safarov, 2019). Third, future research should investigate various factors influencing big data quality. Finally, any future research should also actively investigate how to leverage a firm's capabilities to improve big data quality.

There is some proof that adopting data analytics tools can help businesses become better at making decisions. Studies showed that many businesses that invested in data analytics were unable to fully utilize these capabilities. A study that quantitatively demonstrates the influence of successful use of data analytics (data analytics competency) on firm decision-making is lacking, despite the academic and practitioner literature emphasizing the benefit of employing data analytics tools on firm decision-making effectiveness. We, therefore, set out to investigate how this impact operates. Understanding the elements affecting it is a novel addition to the data analytics literature because increasing firms' decision-making performance is the ultimate purpose of data analysis in the realm of data analytics. In this research, we filled this knowledge gap by using Huber's (1990) theory of effects of advanced IT on decision-making and Bharadwaj's (2000) framework of key IT-based resources to describe and justify data analytics competency for enterprises as a multidimensional formative index, as well as to create and validate a framework to predict the role of data analytics competency on firm decision-making performance (i.e., decision quality and efficiency). The two aforementioned initiatives represent fresh characteristics that have not yet been discussed in IS literature.

Furthermore, in this work, various techniques were used to identify the data quality aspects of user-generated wearable device data. Literature analysis and survey were done to comprehend the issues associated with data quality for investigators and their perspectives on data quality dimensions. Domain specialists chose the right dimensions based on this information (Cho et al., 2021; Ghasemaghahi et al., 2018).

## **Completeness**

In this analysis, the contextual and data quality characteristics of breadth and density completeness were thought to be crucial for conducting research. It is critical to evaluate the breadth and completeness of data sets, especially those gathered in a bring-your-own-device research environment. The number of valid days required within a specific data collection period or the frequency with which the data must be present for the individual data to be included in the analysis is another way that researchers can define completeness. Further research is required to establish how completeness is defined in research studies because recently launched gadgets have the capacity to evaluate numerous data types and gather data continuously for years (Côte-Real et al., 2020).

## **Conformance**

While value, relational, and computational conformity are all seen as crucial aspects of wearable device data, data administration and quality evaluation present difficulties. Only the data dictionary and relational model particular to the model, brand, and version of the device can be used to evaluate the value and relational conformity, and only in cases where this data is publicly available.

## **Plausibility**

Plausibility fits with researchers' demands for precise data values. For example, the data might be judged implausible when step counts are higher than expected but associated heart rate values are lower than usual. Before beginning an investigation, researchers frequently and arbitrarily create their own standards to determine the facts' plausibility. However, creating a collection of prospective data quality criteria requires extensive topic expertise and expert time. Therefore, developing a knowledge base of data quality guidelines for user-generated wearable device data would not only help future researchers save time but also eliminate the need for ad hoc data quality guidelines (Cho et al., 2021; Dakkak et al., 2021).

## **Conclusion**

### **Theoretical Implications**

In this paper, we reviewed literature related to data quality. Based on the literature review, we presented previous studies on data quality. We identified the three approaches used to study the dimensions of data quality and summarized those data quality dimensions by referencing to work of several works of research in this area. The results indicated that data quality has multiple dimensions, with accuracy,

completeness, consistency, timeliness, and relevance viewed as the important dimensions of data quality. Also, through the literature review, we identified two important factors that influence data quality: time pressure and data user experience, which are frequently mentioned throughout the literature. We also identified the impact of data quality through present studies related to data quality impacts. We found that many studies examined the impact of data quality on decision-making. By doing so, we depicted a clear picture of issues related to data quality, identified the gaps in existing research, and proposed three important questions related to big data quality.

A study that quantitatively demonstrates the influence of the successful use of data analytics (data analytics competency) on firm decision-making is lacking, despite the academic and practitioner literature emphasizing the benefit of employing data analytics tools on firm decision-making performance (Bridge, 2022). We, therefore, set out to investigate how this impact operates. Understanding the elements affecting it is a novel addition to the data analytics literature because increasing firms' decision-making performance is the overarching goal of data analysis in the realm of data analytics.

Surprisingly, although the size of the data dramatically improves the quality of firm decision-making, it has no discernible effect on firm decision efficiency, according to later investigations. This indicates that while having large amounts of data is a great resource for businesses to use to increase the quality of company decisions, it does not increase the speed at which they can make decisions. The difficulties in gathering, managing, and evaluating massive amounts of data may be to blame. Decision quality and decision efficiency were highly impacted by all other first-order constructs, including data quality, analytical ability, subject expertise, and tool sophistication.

Limitations in our review of literature do exist. Although we summarized the data quality dimensions, antecedents, and impacts through our literature review, we may have overlooked other data quality dimensions, antecedents, and impacts due to the limited number of papers we reviewed. In order to have a comprehensive understanding of data quality, we also suggest that further research needs to be conducted through the review of more papers related to data quality to reveal more dimensions, antecedents, and impacts.

## **Managerial Implications**

The findings of this study have significant ramifications for managers who use data analytics to their benefit. Organizations that make sizable investments in these technologies do so primarily to enhance decision-making performance through the use of data analytics. As an outcome, managers must pay close attention to strengthening data analytics competency dimensions to improve firm decision-making performance as a result of using these tools. This is because they have now adequately explained a large portion of the variance in decision-making performance (Ghasemaghahi et al., 2018). The use of analytical tools may fail to enhance organizational decision-making performance without such competency.

Companies could, for instance, invest in training to enhance employees' analytical skills in order to improve firm decision-making. When employees are equipped with the skills needed to carry out the demands of their jobs, the quality of their



work is increased. Furthermore, managers must ensure that staff members who utilize data analytics to make important choices have the necessary domain expertise to correctly use the tools and interpret the findings. When purchasing data analytics tools, managers can use careful selection procedures to ensure the chosen means are powerful enough to support all the data required for carrying out current and upcoming analytical jobs. Therefore, managers must invest in data quality to speed up information processing and increase the efficiency of business decisions if they want to increase their data analytics proficiency.

## Ideas for Future Research

It is essential to recognize the limits of this study, as with all studies. First, factors other than data analytics proficiency can have an impact on how well a company makes decisions. Future research is also necessary to better understand how other factors (such as organizational structure and business procedures) affect the effectiveness of company decision-making. Second, open data research is a new area of study, and the current assessment of open data within existing research has space for improvement, according to the initial literature review. This improvement can target a variety of open data paradigm elements, including commonly acknowledged dataset attributes, publishing specifications for open datasets, adherence to specific policies, necessary open data infrastructure functionalities, assessment processes of datasets, openness, accountability, involvement or collaboration, and evaluation of economic, social, political, and human value in open data initiatives. Because open data is, by definition, free, accessible to the general public, nonexclusive (unrestricted by copyrights, patents, etc.), open-licensed, usability-structured, and so forth, its use may be advantageous to a variety of stakeholders. These advantages can include the creation of new jobs, economic expansion, the introduction of new goods and services, the improvement of already existing ones, a rise in citizen participation, and assistance in decision-making. Consequently, the open data paradigm illustrates how IT may support social, economic, and personal growth.

Finally, the decisions that were made were not explicitly covered by this study. Future study is necessary to examine the impact of data analytics competency and each of its dimensions on decision-making consequences in particular contexts, as the relative importance of data analytics competency and its many dimensions may change depending on the type of decision being made (e.g., recruitment processes, marketing promotions).

**Funding** The paper was supported by the H-E-B School of Business & Administration, the University of the Incarnate Word, and the Social Science Foundation of the Ministry of Education of China (16YJA630025).

**Data Availability** The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

## References

- Abdoul, M., & Omri, A. (2021). Exploring the nexus among FDI inflows, environmental quality, human capital, and economic growth in the Mediterranean region. *Journal of the Knowledge Economy*, 12(2), 788–810.
- Ahituv, N., Igbaria, M., & Sella, A. V. (1998). The effects of time pressure and completeness of information on decision making. *Journal of Management Information Systems*, 15(2), 153–172.
- Bailey, J. E., & Pearson, S. W. (1983). Development of a tool for measuring and analyzing computer user satisfaction. *Management Science*, 29(5), 530–545.
- Ballou, D. P., & Pazer, H. L. (1985). Modeling data and process quality in multi-input, multi-output information systems. *Management Science*, 31(2), 150–162.
- Ballou, D. P., & Pazer, H. L. (1995). Designing information systems to optimize the accuracy-timeliness tradeoff. *Information Systems Research*, 6(1), 51–72.
- Ballou, D. P., & Tayi, G. K. (1999). Enhancing data quality in data warehouse environments. *Communications of the ACM*, 42(1), 73–78.
- Ballou, D., Wang, R., Pazer, H., & Tayi, G. K. (1998). Modeling information manufacturing systems to determine information product quality. *Management Science*, 44(4), 462–484.
- Barkhordari, S., Fattahi, M., & Azimi, N. A. (2019). The impact of knowledge-based economy on growth performance: Evidence from MENA countries. *Journal of the Knowledge Economy*, 10(3), 1168–1182.
- Berghel, H. (1997). Cyberspace 2000: Dealing with information overload. *Communications of the ACM*, 40(2), 19–24.
- Bharadwaj, A. S. (2000). A resource-based perspective on information technology capability and firm performance: An empirical investigation. *MIS Quarterly*, 169–196.
- Bouchoucha, N., & Benammou, S. (2020). Does institutional quality matter foreign direct investment? Evidence from African countries. *Journal of the Knowledge Economy*, 11(1), 390–404.
- Bridge, J. (2022). A quantitative study of the relationship of data quality dimensions and user satisfaction with cyber threat intelligence (Doctoral dissertation, Capella University).
- Carayannis, E. G., Barth, T. D., & Campbell, D. F. (2012). The Quintuple Helix innovation model: Global warming as a challenge and driver for innovation. *Journal of Innovation and Entrepreneurship*, 1(1), 1–12.
- Chengalur-Smith, I. N., Ballou, D. P., & Pazer, H. L. (1999). The impact of data quality information on decision making: An exploratory analysis. *IEEE Transactions on Knowledge and Data Engineering*, 11(6), 853–864.
- Cho, S., Weng, C., Kahn, M. G., & Natarajan, K. (2021). Identifying data quality dimensions for person-generated wearable device data: Multi-method study. *JMIR mHealth and uHealth*, 9(12), e31618.
- Côrte-Real, N., Ruivo, P., & Oliveira, T. (2020). Leveraging internet of things and big data analytics initiatives in European and American firms: Is data quality a way to extract business value? *Information & Management*, 57(1), 103141.
- Dakkak, A., Zhang, H., Mattos, D. I., Bosch, J., & Olsson, H. H. (2021, December). Towards continuous data collection from in-service products: Exploring the relation between data dimensions and collection challenges. In *2021 28th Asia-Pacific Software Engineering Conference (APSEC)* (pp. 243–252). IEEE.
- Danish, R. Q., Asghar, J., Ahmad, Z., & Ali, H. F. (2019). Factors affecting “entrepreneurial culture”: The mediating role of creativity. *Journal of Innovation and Entrepreneurship*, 8(1), 1–12.
- Davenport, T. H., & Harris, J. G. (2007). *Competing on analytics: The new science of winning*. Harvard Business Press.
- Davenport, T. H., & Patil, D. J. (2012). Data scientist. *Harvard Business Review*, 90(5), 70–76.
- DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information Systems Research*, 3(1), 60–95.
- Dranev, Y., Izosimova, A., & Meissner, D. (2020). Organizational ambidexterity and performance: Assessment approaches and empirical evidence. *Journal of the Knowledge Economy*, 11(2), 676–691.
- EbabaEngidaw, A. (2021). The effect of external factors on industry performance: The case of Lalibela City micro and small enterprises, Ethiopia. *Journal of Innovation and Entrepreneurship*, 10(1), 1–14.


- Even, A., Shankaranarayanan, G., & Berger, P. D. (2010). Evaluating a model for cost-effective data quality management in a real-world CRM setting. *Decision Support Systems*, 50(1), 152–163.
- Feki, C., & Mnif, S. (2016). Entrepreneurship, technological innovation, and economic growth: Empirical analysis of panel data. *Journal of the Knowledge Economy*, 7(4), 984–999.
- Firmani, D., Tanca, L., & Torlone, R. (2019). Ethical dimensions for data quality. *Journal of Data and Information Quality (JDIQ)*, 12(1), 1–5.
- Fisher, C. W., & Kingma, B. R. (2001). Criticality of data quality as exemplified in two disasters. *Information & Management*, 39(2), 109–116.
- Fisher, C. W., Chengalur-Smith, I., & Ballou, D. P. (2003). The impact of experience and time on the use of data quality information in decision making. *Information Systems Research*, 14(2), 170–188.
- Foshay, N., Mukherjee, A., & Taylor, A. (2007). Does data warehouse end-user metadata add value? *Communications of the ACM*, 50(11), 70–77.
- Ghasemaghahi, M., Ebrahimi, S., & Hassanein, K. (2018). Data analytics competency for improving firm decision making performance. *The Journal of Strategic Information Systems*, 27(1), 101–113.
- Goll, I., & Rasheed, A. A. (2005). The relationships between top management demographic characteristics, rational decision making, environmental munificence, and firm performance. *Organization Studies*, 26(7), 999–1023.
- Gordon, K. (2013). What is big data? *Itnow*, 55(3), 12–13.
- Hook, D. W., Porter, S. J., & Herzog, C. (2018). Dimensions: Building context for search and evaluation. *Frontiers in Research Metrics and Analytics*, 3, 23.
- Hosack, B., Hall, D., Paradice, D., & Courtney, J. F. (2012). A look toward the future: Decision support systems research is alive and well. *Journal of the Association for Information Systems*, 13(5), 3.
- Huber, G. P. (1990). A theory of the effects of advanced information technologies on organizational design, intelligence, and decision making. *Academy of Management Review*, 15(1), 47–71. <https://www.jstor.org/stable/258105>
- Ives, B., Olson, M. H., & Baroudi, J. J. (1983). The measurement of user information satisfaction. *Communications of the ACM*, 26(10), 785–793.
- Juddoo, S., & George, C. (2018). Discovering most important data quality dimensions using latent semantic analysis. In *2018 International Conference on Advances in Big Data, Computing and Data Communication Systems (icABCD)* (pp. 1–6). IEEE.
- Kanwar, A., & Sanjeeva, M. (2022). Student satisfaction survey: A key for quality improvement in the higher education institution. *Journal of Innovation and Entrepreneurship*, 11(1), 1–10.
- Khan, R. U., Salamzadeh, Y., Shah, S. Z. A., & Hussain, M. (2021). Factors affecting women entrepreneurs' success: A study of small-and medium-sized enterprises in emerging market of Pakistan. *Journal of Innovation and Entrepreneurship*, 10(1), 1–21.
- Klein, R. H., Klein, D. B., & Luciano, E. M. (2018). Open Government Data: Concepts, approaches and dimensions over time. *Revista Economia & Gestão*, 18(49), 4–24.
- Laudon, K. C. (1986). Data quality and due process in large interorganizational record systems. *Communications of the ACM*, 29(1), 4–11.
- Loveman, G. (2003). Diamonds in the data mine. *Harvard Business Review*, 81(5), 109–113.
- Maradana, R. P., Pradhan, R. P., Dash, S., Gaurav, K., Jayakumar, M., & Chatterjee, D. (2017). Does innovation promote economic growth? Evidence from European countries. *Journal of Innovation and Entrepreneurship*, 6(1), 1–23.
- Mock, T. J. (1971). Concepts of information value and accounting. *The Accounting Review*, 46(4), 765–778.
- Moges, H. T., Dejaeger, K., Lemahieu, W., & Baensens, B. (2013). A multidimensional analysis of data quality for credit risk management: New insights and challenges. *Information & Management*, 50(1), 43–58.
- Morey, R. C. (1982). Estimating and improving the quality of information in a Mis. *Communications of the ACM*, 25(5), 337–342.
- Ouechati, I. (2022). Financial inclusion, institutional quality, and inequality: An empirical analysis. *Journal of the Knowledge Economy*, 1–25. <https://doi.org/10.1007/s13132-022-00909-y>
- Parssian, A., Sarkar, S., & Jacob, V. S. (2004). Assessing data quality for information products: Impact of selection, projection, and Cartesian product. *Management Science*, 50(7), 967–982.
- Pipino, L. L., Lee, Y. W., & Wang, R. Y. (2002). Data quality assessment. *Communications of the ACM*, 45(4), 211–218.
- Price, R., & Shanks, G. (2011). The impact of data quality tags on decision-making outcomes and process. *Journal of the Association for Information Systems*, 12(4), 1.

- Price, D. P., Stoica, M., & Boncella, R. J. (2013). The relationship between innovation, knowledge, and performance in family and non-family firms: An analysis of SMEs. *Journal of Innovation and Entrepreneurship*, 2(1), 1–20.
- Prifti, R., & Alimehmeti, G. (2017). Market orientation, innovation, and firm performance—An analysis of Albanian firms. *Journal of Innovation and Entrepreneurship*, 6(1), 1–19.
- Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big Data*, 1(1), 51–59.
- Reforgiato Recupero, D., Castronovo, M., Consoli, S., Costanzo, T., Gangemi, A., Grasso, L., ... & Spampinato, E. (2016). An innovative, open, interoperable citizen engagement cloud platform for smart government and users' interaction. *Journal of the Knowledge Economy*, 7(2), 388–412.
- Russo, G., Marsigalia, B., Evangelista, F., Palmaccio, M., & Maggioni, M. (2015). Exploring regulations and scope of the Internet of Things in contemporary companies: A first literature analysis. *Journal of Innovation and Entrepreneurship*, 4(1), 1–13.
- Safarov, I. (2019). Institutional dimensions of open government data implementation: Evidence from the Netherlands, Sweden, and the UK. *Public Performance & Management Review*, 42(2), 305–328.
- Sarkheyli, A., & Sarkheyli, E. (2019). Smart megaprojects in smart cities, dimensions, and challenges. In *Smart Cities Cybersecurity and Privacy* (pp. 269–277). Elsevier.
- Shankaranarayanan, G., & Cai, Y. (2006). Supporting data quality management in decision-making. *Decision Support Systems*, 42(1), 302–317.
- Shumetie, A., & Watabaji, M. D. (2019). Effect of corruption and political instability on enterprises' innovativeness in Ethiopia: Pooled data based. *Journal of Innovation and Entrepreneurship*, 8(1), 1–19.
- Šlibar, B., Oreški, D., & BegičevićRedep, N. (2021). Importance of the open data assessment: An insight into the (meta) data quality dimensions. *SAGE Open*, 11(2), 21582440211023176.
- Strong, D. M. (1997). IT process designs for improving information quality and reducing exception handling: A simulation experiment. *Information & Management*, 31(5), 251–263.
- Surowiecki, J. (2013). "Where Nokia Went Wrong," Retrieved February 20, 2020, from <http://www.newyorker.com/business/currency/where-Nokia-went-wrong>
- Tayi, G. K., & Ballou, D. P. (1998). Examining data quality. *Communications of the ACM*, 41(2), 54–57.
- Wand, Y., & Wang, R. Y. (1996). Anchoring data quality dimensions in ontological foundations. *Communications of the ACM*, 39(11), 86–95.
- Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems*, 12(4), 5–33.
- Wang, R. Y., Lee, Y. W., Pipino, L. L., & Strong, D. M. (1998). Manage your information as a product. *MIT Sloan Management Review*, 39(4), 95.
- Wang, R. Y., Storey, V. C., & Firth, C. P. (1995). A framework for analysis of data quality research. *IEEE Transactions on Knowledge and Data Engineering*, 7(4), 623–640.
- Zhuo, Z., Muhammad, B., & Khan, S. (2021). Underlying the relationship between governance and economic growth in developed countries. *Journal of the Knowledge Economy*, 12(3), 1314–1330.
- Zouari, G., & Abdelhedi, M. (2021). Customer satisfaction in the digital era: Evidence from Islamic banking. *Journal of Innovation and Entrepreneurship*, 10(1), 1–18.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

## Authors and Affiliations

Jingran Wang<sup>2</sup> · Yi Liu<sup>3</sup> · Peigong Li<sup>1</sup> · Zhenxing Lin<sup>1</sup> · Stavros Sindakis<sup>4</sup>  · Sakshi Aggarwal<sup>5</sup>

Jingran Wang  
jrwang1@sina.com

Yi Liu  
yiliu5@uiwtx.edu

Peigong Li  
pgli@xmu.edu.cn

Stavros Sindakis  
sindakis.stavros@ac.eap.gr

Sakshi Aggarwal  
s.aggarwal@iseeg.org

- <sup>1</sup> School of Accounting, Shanghai Lixin University of Accounting and Finance, 2800 Wenxiang Rd, Songjiang District, Shanghai 201620, China
- <sup>2</sup> Sogang Business School, Sogang University, 35 Baekbeom-Ro, Mapo-Gu, Seoul, South Korea
- <sup>3</sup> H-E-B School of Business & Administration, University of the Incarnate Word, 4301 Broadway, San Antonio, TX 78209, USA
- <sup>4</sup> School of Social Sciences, Hellenic Open University, 18 Aristotelous Street, Patras 26335, Greece
- <sup>5</sup> Institute of Strategy, Entrepreneurship and Education for Growth, Athens, Greece