

# Exploring Data Veracity Management in a Post-Truth Business Environment: An Integrative Literature Review and Future Research Direction

Putra Endi Catyanadika

 <https://orcid.org/0000-0003-4452-970X>


*RMIT University, Australia*

Alvedi Sabani

 <https://orcid.org/0000-0001-6628-8092>

*RMIT University, Australia*

Mark A. A. M. Leenders

 <https://orcid.org/0000-0002-6488-8542>

*RMIT University, Australia*

## ABSTRACT

With the ever-increasing volume and variety of data generated, organisations have to ensure their truthfulness and reliability. This paper provides overview of current research on managing data veracity in a business environment where misinformation is growing. A literature analysis from 2002 to 2023 identified three major themes: methods for ensuring data validity, data processing and optimisation, and data veracity in sustainability performance. In addition, the study highlights the gaps in the current research and proposes future research directions to help develop a better understanding of the themes and organisational implications. The study concludes that data veracity is crucial for future organisational research. Nevertheless, further work is required to refine the definition of data veracity to incorporate ‘truthfulness’ better, understand human capabilities to support it, examine firms’ governance of truthfulness and measure data veracity for social impact. The implications of these findings for data management and the development of relevant theories are discussed.

## KEYWORDS

Bibliographic Analysis, Big Data, Data Reliability, Data Veracity, Literature Review, Truthfulness

## INTRODUCTION

In recent years, there has been unprecedented growth in globally generated data. The rapid development of information technology, IoT, and data-driven organizations has accelerated the reliance on data for daily business operations (Parviainen et al., 2017; Sawadogo & Darmont, 2021). The advancement of information technology hastens the process of acquiring, storing, organizing, sharing, and visualizing data in diverse formats, which increases its volume, velocity, and variety (Cappa et al., 2021). In fact, data growth worldwide, which reached 33 zettabytes in 2018, is projected to reach 125 zettabytes in 2025, with an estimated 40% growth annually (Tang et al., 2020). Consequently, the increasing demand for systems, tools, capabilities, and storage to manage data has resulted in a

DOI: 10.4018/JDM.361726

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

projected global market value exceeding US\$203 billion. This indicates the substantial growth in data that is reshaping the business landscape (Tao et al., 2019).

The prevalence of data sharing has significantly influenced how data is accessed and used, presenting both benefits and risks for organizations. The integrity and accuracy of data have become increasingly important as businesses rely on data for decision-making processes. However, the ease of large-scale data generation and sharing, primarily through social media and fake news, has raised concerns about its credibility and quality. Several studies highlight the prevalence of misinformation and data manipulation on social media platforms, severely compromising the accuracy of available information (Conroy et al., 2015; Nakov & Da San Martino, 2021). The emergence of the filter bubble effect, wherein information is selectively presented to suit specific objectives, further exacerbates the distortion of reality and the dissemination of biased content (Amrollahi, 2021).

Post-truth society has emerged in political sciences, referring to the widespread concern about disputes over public truth claims (Suiter, 2016). The term even received the Oxford Word of the Year status from Oxford University Press in the context of Brexit and the Trump election in 2016. The implications for data collection and management are profound. Therefore, the goal of this research is to explore the body of knowledge regarding data veracity in this context.

Businesses tend to outsource their data collection processes as a cost-cutting measure (Hamlen & Thuraisingham, 2013). While this approach provides financial benefits, it introduces uncertainty regarding the accuracy and quality of the data obtained from third-party vendors. Cohen (2017) highlighted the risk of inaccurate data when vendors cannot guarantee its accuracy, potentially resulting in flawed decision-making. In addition to accuracy concerns, sharing information with external parties poses significant risks to data security. Potential data breaches and unauthorized access can compromise the confidentiality and integrity of sensitive information (Hamlen & Thuraisingham, 2013; Pandey et al., 2020). These risks are compounded by the rapid spread of fake news and other forms of deceptive information, which distorts understanding and undermines the accuracy of data analytics. Thus, there is an increasing need to develop robust measures to combat data deception and fake news while navigating the complex, data-driven business world.

Recognizing proper data quality and management systems has changed our understanding of data management and its essential characteristics. Big data is characterized by increased data volume, velocity, and variety, necessitating data veracity to yield valuable insights (Geerts & O'Leary, 2022; Ghasemaghahi, 2021; Rubin & Lukoianova, 2014). The goal of veracity is to extract maximum value from data processing by maintaining the accuracy, credibility, and reliability of the data (Rubin & Lukoianova, 2014). Increasing and preserving data veracity can be achieved through different approaches, including fact-checking or veracity assessment tools via data processing, continuous monitoring of data flow, or experts who can assess valuable information from big data (Cappa et al., 2021; García Lozano et al., 2020).

Adhering to the principle of data veracity is essential for organizations to provide honest and unbiased data that is conducive to effective decision-making and operational performance (Chae et al., 2014; Gelman, 2010; Janssen et al., 2017; Rubin & Lukoianova, 2014). Additionally, it can mitigate the risk of potential data scams, propaganda, or disinformation attempts (Evans et al., 2021), promoting trust and confidence among stakeholders (La Torre et al., 2020; Shafieizadeh & Tao, 2020). Moreover, ensuring the credibility and reliability of data is an ethical obligation, contributing to the creation of a better society (Rubin, 2022; Sanney et al., 2020). Making data veracity a priority in data management objectives is essential for achieving successful data quality.

There has been an increase in research efforts on veracity systems and their assessment methods since the introduction of big data management and the changing paradigm of its characteristics in 2012 (Brynjolfsson & McAfee, 2012; Schroeck et al., 2012). The research attempts focus on various aspects, including the definition and concept of veracity (Ramachandramurthy et al., 2015; Rubin & Lukoianova, 2014; Schroeck et al., 2012), the development of applications related to veracity (Fuller et al., 2009; Sampson et al., 2016; Shi & Weninger, 2016), tools and methods for assessing veracity

(García Lozano et al., 2020; Nakov et al., 2021; Shaar, Nikolov, et al., 2020), and the implementation of veracity in the context of big data (Cappa et al., 2021; Geerts & O’Leary, 2022; Shams & Solima, 2019). Numerous studies have explored the evolution of veracity across different fields, including healthcare (Lyu et al., 2020; Mønsted, 2019; Reimer & Madigan, 2019), journalism (Eembijamil et al., 2017; Hassan et al., 2015; Jamil et al., 2015; Liu et al., 2016), supply chain (Alshawabkeh et al., 2022; Viet et al., 2021), and computer science (Moke et al., 2021; Revathy & Canessane, 2018; Shukla et al., 2020). The wide range of research endeavors across various disciplines demonstrates the significance of understanding and developing data veracity.

Data veracity for business organizations is a crucial area for focused study. Organizations, driven by digitalizing business processes, are collecting and sharing enormous amounts of data for their daily operations (Sestino et al., 2020). The gathered information presents various opportunities and advantages, such as supporting decision-making processes (Constantiou & Kallinikos, 2015; Janssen et al., 2017), improving the value creation process and competitive advantages (R. Lin et al., 2020), and enhancing supply chain management (Viet et al., 2021). However, managing data poses challenges for businesses due to the complexity and uncertainty associated with gathered data, potentially resulting in significant costs for the firms (Gärtner & Hiebl, 2018; Rubin & Lukoianova, 2014).

Effective data management is essential to extract value from data resources. Valuable data is obtained when the benefits of data processing exceed the costs of data organization (Cappa et al., 2021; Chesbrough et al., 2018). In this context, data veracity enhances the accuracy of processed data for business activities, minimizing the risk of misinformation, reducing the cost associated with handling non-valuable data, and improving overall firm performance (Cappa et al., 2021; Figueira & Oliveira, 2017; Saha & Srivastava, 2014).

Considering the vital role of data veracity in business and management, there may be a need to identify additional research development in this field. The emerging and scattered literature, as well as ongoing developments and implementation related to veracity, may warrant a systematic review attempt. This review would contribute to understanding the content and depth of existing theories, highlighting potential gaps for further exploration (Xiao & Watson, 2019). In addition, a systematic review of the research in this field can provide insight into veracity in business and management, helping to identify knowledge gaps and areas for development.

This study began with a preliminary observation to uncover studies and theories on data veracity development. A search of the Scopus database for registered journals and proceedings from the past 20 years revealed a growing interest in this topic. However, a comprehensive review of the literature published in the business and management field is lacking. The observation also found that research focused on data veracity tends to flourish in other fields, such as computer science or engineering. Therefore, conducting a thorough literature review in the business area is crucial, especially considering the increasing number of publications and the importance of data veracity for daily business activities.

To address these gaps, this study conducted literature reviews on publications related to data veracity in the fields of management and business knowledge. The purpose of the literature review was to examine the research cluster to determine the current state of research, identify articles that significantly contributed to the topic, and explore potential directions for future research. To achieve these objectives, this study aimed to answer the following research questions:

- How have studies about veracity in business and management evolved in the last 20 years?
- What are the key research clusters and common themes in publications on veracity within business and management?
- Based on the analysis results, what future research agenda can be proposed to contribute to the development of research on veracity in business and management?

Following the order of these research questions, the article is structured as follows. The following section presents a conceptual background of veracity theories and prior bibliographic studies

addressing veracity and big data. The next sections discuss the study's methodology and results, comparing publications about veracity in business, management, and research regarding veracity. This is followed by a presentation of results from cluster analysis, citation, co-citation, and bibliographic coupling. The article explores each analysis, as well as the presentation of proposed research directions. Finally, the article presents a conclusion and limitations of the study.

## CONCEPTUAL BACKGROUND

### Data Veracity

Although the term “veracity” has been used since 2012, its exact definition remains a subject of debate and discussion (García Lozano et al., 2020). Various studies describe veracity in relation to data reliability, accuracy, and credibility, which support the quality of data processing outcomes (García Lozano et al., 2020; Schroeck et al., 2012). Other studies view veracity as an uncertain aspect of big data, attributed to complexity (Rubin & Lukoianova, 2014), biases, and variances in data (Himanen et al., 2019) or errors and inconsistencies in data quality (Saha & Srivastava, 2014). These perspectives often emphasize the potential negative consequences associated with the growth of big data.

Rubin and Lukoianova (2014) viewed veracity as a development process to improve data quality. They emphasized data transformation from subjective to objective, deceptive to truthful, and implausible to credible. Another perspective is defined from an organizational point of view, which is related to digitalization and resource allocation strategies designed to extract value from big data through information analysis (Mithas et al., 2013). In this context, investing in effective data analytics is seen as a means for firms to enhance veracity (LaValle et al., 2011). Managerial efforts include hiring employees with expertise in data analytics or using data analytics in managements' decision-making processes (Cappa et al., 2021).

Despite the various interpretations of veracity dimensions in prior research, terms like data accuracy, truthfulness, trust, credibility, objectivity, and quality are consistently mentioned and closely associated with the concept of data veracity (García Lozano et al., 2020; Hariri et al., 2019; Rubin & Lukoianova, 2014).

Veracity is often associated with the notion of data value. Improved veracity implementation enhances a firm's opportunities to extract valuable information and knowledge from its data (Cappa et al., 2021; Erevelles et al., 2016). However, effectively managing big data demands a substantial investment, requiring significant resources, systems, and devices (Tallon, 2013). The cost associated with big data can also arise when extensive amounts of data are stored within a database but remain underutilized due to an organization's inability to effectively use large datasets (Kimble & Milolidakis, 2015; Rubin & Lukoianova, 2014).

Given the substantial costs involved, big data management is expected to yield valuable advantages that exceed its overall expenses (Björkdahl & Holmén, 2019; Chesbrough et al., 2018; Wamba et al., 2017). In this context, efforts toward veracity ensure that interpretations of big data are reliable, timely, relevant, and insightful (Del Vecchio et al., 2018; Erevelles et al., 2016). The value derived from veracity actions can stem from the optimal use of the data management system and its ability to filter unused data and information.

The concept of big data has traditionally been represented by the 3Vs of volume, velocity, and variety (Kitchin & McArdle, 2016; Laney, 2001). However, as the importance of data and information quality gains recognition, the concept has evolved to include veracity and value, resulting in the acknowledgment of the 5Vs of big data (García Lozano et al., 2020; Reimer & Madigan, 2019; Rubin & Lukoianova, 2014). Moreover, other views have extended the characteristics to the following 6Vs: (1) volume; (2) variety; (3) velocity; (4) variability; (5) veracity; and (6) value (Gandomi & Haider, 2015; Ristevski & Chen, 2018).

Introducing data into a system increases its volume, variety, and velocity, heightening its uncertainty and making it more complex to manage (Kaul et al., 2017; Reimer & Madigan, 2019a;

Rubin & Lukoianova, 2014). Veracity, therefore, plays a crucial role by improving data accuracy, credibility, objectivity, truthfulness, precision, and reliability (García Lozano et al., 2020; Rubin & Lukoianova, 2014; Schroeck et al., 2012). Veracity ensures that the retrieved data adds value while minimizing data risk by addressing the complexity of the volume, velocity, and variety (Geerts & O’Leary, 2022). It is, therefore, vital for delivering valuable information and knowledge to support decision-making processes (Cappa et al., 2021; Erevelles et al., 2016).

Although multiple definitions have been proposed, this study states its own definition of veracity to set the scope and direction of its analyses. The exploration revealed that the term “veracity” was used in various contexts, with both positive and negative perspectives toward data quality. The negative perspective originated from studies that viewed veracity as the consequence of big data complexity (e.g., Himanen et al., 2019; Saha & Srivastava, 2014). Other studies emphasized positive characteristics, viewing data veracity as a quality or capability of data to overcome the complex challenges (e.g., Cappa et al., 2021; García Lozano et al., 2020; Rubin & Lukoianova, 2014; Schroeck et al., 2012). Some studies connected veracity to strategies or processes that enhance data quality, not merely focusing on data characteristics alone. This indicates a shift from negative to positive outcomes, which highlights the processes that improve integrity (Mithas et al., 2013; Rubin & Lukoianova, 2014; Tallon, 2013).

This study views veracity as data characteristics, strategies, and processes that change negative characteristics to positive ones. The authors reviewed and compiled various definitions for the term “veracity,” as represented in Table 1, to capture their adopted definition. In this study, veracity encompasses the complexity of data quality characteristics, such as variance, deception, and bias. It focuses on organizational efforts to gain values from these challenges by prioritizing data accuracy, truthfulness, objectivity, and trust.

### **Similar Research in the Area of Data Veracity**

A preliminary study was conducted to address similar literature reviews on big data and data veracity, tracking the evolution of knowledge in this field and identifying the novelties of this study compared to prior publications. The authors reviewed the Scopus database for journals and proceedings that conducted a literature review analysis for veracity topics from 2002 to 2023. The result of the preliminary observation is summarized in Table 2 and described further in this section.

Several literature review studies for big data have been conducted in the last 20 years. One publication analyzed the literature on the relationship between big data and cognitive computing using publications from Scopus, Web of Science, and the DBLP database (Gupta et al., 2018). The result proposed a map integrating the characteristics of cognitive computing and big data for future research. The second literature review study identified the emerging management area supported by the big data system (Kushwaha et al., 2021). The result proposed 10 areas of emerging management disciplines in which big data analytics are extensively used, including healthcare management, crisis management, governance management, dynamic capabilities management, and decision support systems. The third registered publication proposed a new understanding of big data by analyzing 60 significant works in the field, identifying various scenarios that offer a different perspective to exploring big data (Lugmayr et al., 2017). The research proposed a five-trait framework for cognitive big data that goes beyond its common characteristics. The fourth publication conducted a bibliographic analysis to identify the development of theories regarding big data and dynamic capabilities (Rialti et al., 2019). The study proposed the following essential topics: (1) supply chain management; (2) knowledge management; (3) decision-making; and (4) business process data analytics. Several other publications address aspects within big data implementation, including bibliographic studies focused on the development of research in spatial decision support systems (Keenan & Jankowski, 2019) and big data analytics (Ardito et al., 2019). Additionally, another study addresses the relationship between big data analysis and organizational performance (Batistič & van der Laken, 2019).

Table 1. Definitions of data veracity from previous studies

No	References	Definitions of Data Veracity	Mentioned Dimensions	Focus	Context
1	Reimer and Madigan (2019)	Accuracy and fidelity of particular datasets	accuracy, fidelity	Data characteristics	Accuracy and reliability of data used for clinical decision-making and research
2	García Lozano et al. (2020)	Reliability, accuracy, and credibility to support the data quality outcome	reliability, accuracy, credibility	Data characteristics	Quality and reliability assessment of information that is usually ambiguous, bias, or deceptive
3	Rubin and Lukoianova (2014)	Transforming data characteristic to objective, truthful, and credible	objectivity, truthfulness, credibility	Data processing	Development process to improve data from subjective, deceptive, and implausible to objective, truthful, and credible
4	Schroeck et al. (2012)	Reliability level of datasets	uncertainty, unpredictability, data value	Data characteristics	Seeking data value inside the uncertainty of big datasets
5	Gandomi and Haider (2015)	Unreliability of data sources	uncertainty, data value, data mining	Data processing	Utilizing data mining tools to seek values from uncertain data
6	Himanen et al. (2019)	Quality and trustworthiness from bias and variance within the datasets	uncertainty, bias, variances, trustworthiness	Data characteristics	Ensuring data generated is accurate and credible in the context of data-driven science
7	Saha and Srivastava (2014)	Inconsistencies and uncertainties inherent in large datasets	inconsistency, data quality rules, efficiency, accuracy	Data characteristics	Scalable and efficient technique to identify data errors
9	Cappa et al. (2021)	Reliability, accuracy, and trustworthiness of data being analyzed	reliability, accuracy, insightfulness, trustworthiness	Data processing	Firms' ability to capture value from big data analyses by ensuring that data used is vast, diverse, reliable, and insightful
10	Janssen et al. (2017)	Manipulation and noise in big datasets	incomplete, inconsistent, noisy data	Data characteristics	Challenge in big data due to vast amounts of information from various sources
11	Assiri (2020)	Process to define the mark of quality and determine quality of datasets	accuracy, consistency, timeliness, trustworthiness	Data processing	Higher level of veracity, allowing firms to focus on data that is reliable, relevant, and timely
12	Hariri et al. (2019)	Data quality represented in accuracy and consistency that affected analytics outcome	accuracy, consistency, completeness, trustworthiness, uncertainty mitigation	Data characteristics	Veracity is critical big data challenge due to risk of incompleteness, noise, and inconsistency of data

The authors also searched the Scopus database for a bibliographic study addressing veracity in the management and business fields. Two prior bibliographic studies were considered in this attempt, as the research methods and topics bring similarities to this study. The first is a bibliographic analysis to identify the development of theories and practices regarding veracity assessment of online data (García Lozano et al., 2020). The paper identifies effective methods and techniques for veracity

**Table 2. Similar research identified from preliminary observation**

No	Publications	Databases	Methods and Tools	Focus Area
1	Lugmayr et al. (2017)	More than 60 literature sources	Expert think tank discussions	New insights of big data
2	Ardito et al. (2019)	Web of Science	Bibliometric analysis by VOSviewer	Big data analytics
3	Gupta et al. (2018)	Scopus, Web of Science, Digital Bibliography & Library Project (DBLP)	Manual reviews, community-detected algorithm	Big data and cognitive computing
4	Keenan and Jankowski (2019)	Web of Science	Bibliometric analysis by VOSviewer	Spacial decision support systems
5	Batistič and van der Laken (2019)	ISI Web of Knowledge	Bibliometric analysis by VOSviewer	Big data analytics
6	Rialti et al. (2019)	Web of Science	Bibliometric analysis by VOSviewer	Big data and dynamic capabilities
7	García Lozano et al. (2020)	IEEE, Web of Science, ACM, Science Direct, Scopus	Manual review with designated review protocols	Veracity assessment
8	García et al. (2020)	Web of Science	Bibliometric analysis by VOSviewer	Fake news
9	Kushwaha et al. (2021)	50 top-listed journals from ABDC	Bibliometric analysis by VOSviewer	Application of big data in management disciplines

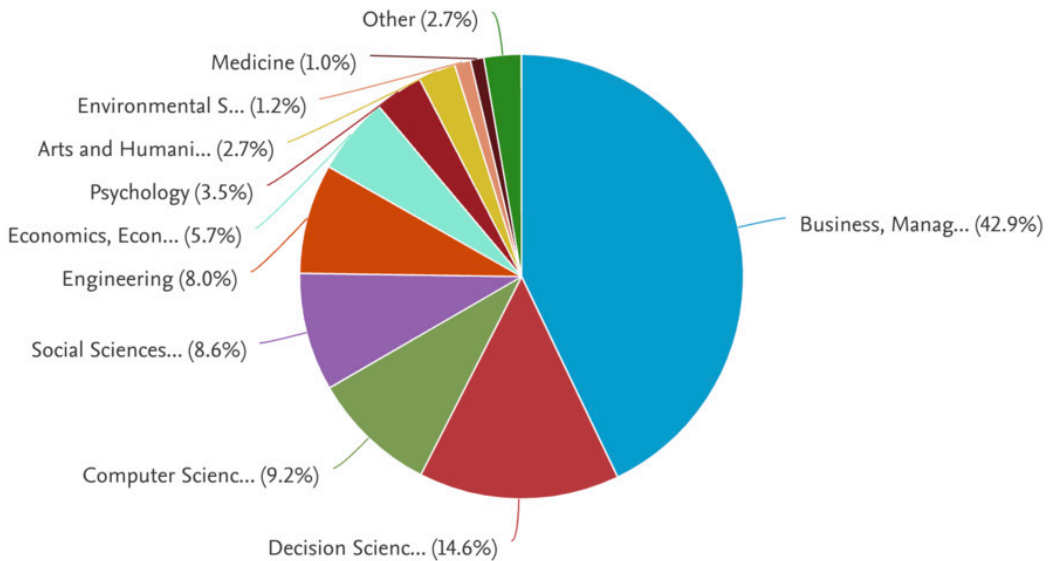
assessment, proposing research gaps that may be used as a basis for future research directions. The second addressed the term “fake news,” as well as the development of research regarding its use in the research community (García et al., 2020). This study offers insight into the development of veracity, particularly toward fake news prevention through informal education.

First, from the preliminary observation of bibliographic research on big data, the authors found that while various studies have explored big data management across various areas, only a few literature studies explicitly address big data veracity as a primary focus. Databases from well-reputed academic sources like Scopus and Web of Science were selected as primary sources for publication, indicating their credibility. Second, the authors found that VOSviewer analysis was often adopted as an analysis tool. Third, the authors determined that literature review research on data veracity for business and management still needs to be conducted. These findings, therefore, motivated the authors to address gaps and conduct a literature review study to address data veracity in the fields of business and management.

## METHODOLOGY

This study employed a bibliometric approach to acquire the core elements of research on data veracity in publications related to business and management. Using a bibliometric approach aids in keeping track of research frontiers in the designated topic, outlining future research directions and potential exploration, and assisting in making informed research decisions based on available literature (Katoch, 2022). Bibliometric mapping also captures the interdisciplinarity of the research subject, allowing users to understand the relationship between two or more research keywords in the investigated topic (Pauna et al., 2018). The bibliometric analysis in this study used VOSviewer software version 1.6.20, which facilitated the generation and exploration of the graphical representations derived from bibliometric network data (Pauna et al., 2018).

Figure 1. Fields addressed in datasets for analysis



Data for the analysis were extracted from the Scopus database, focusing on selected publications from 2002 to 2023. A specific search query was applied to identify publications containing general veracity research and veracity in the business and management field. Articles also addressed other fields, as many articles in the database were multidisciplinary, including decision science, computer science, and social science. The scope of fields addressed is visualized in Figure 1.

There were no prior attempts to explore this topic. Thus, a comparison was made between search results for veracity across all registered areas and veracity, specifically in business and management. Including veracity research from all fields entails incorporating publications from various study disciplines. This includes those in which the term “veracity” may not relate to big data, such as cartographic veracity (Lloyd & Lilley, 2009), political veracity (Zvokuomba & Batisai, 2020), and veracity in chemistry (Barba-Montoya et al., 2018). The search results for veracity research across all registered fields on the Scopus database yielded 5,953 articles. The article selection, review process, and analyses conducted are illustrated in Figure 2.

The use of the search term “veracity” was decided after several trial processes using keywords like “data veracity,” “information veracity,” and “veracity management.” Previous attempts limited the search results, causing many relevant publications to be excluded due to missing keywords. This study used the keyword “veracity,” applying filters to capture a broader range of publications on the topic and minimize the risk of excluding relevant articles.

The initial set of 5,953 articles underwent a filtering process, focusing on articles registered in business, management, and accounting. This resulted in a refined total of 357 publications. All articles were manually scrutinized to eliminate those unrelated to data veracity, resulting in a final selection of 245 articles for analysis. The manual selection was necessary as some articles contained the word “veracity” in the abstract but were unrelated to the context of data veracity. For example, specific abstracts included sentences like “in order to test the veracity of the model ...,” which pertained to hypothesis testing rather than data veracity. The article selection and review process are depicted in Figure 2, while the search framework for analysis is presented in Table 3.



Table 3. Search framework for analysis

Database	Scopus
Analysis	Co-occurrence, citation, co-citation and bibliographic coupling analysis
Search Query (TITTLE-ABS-KEY)	TITLE-ABS-KEY (veracity) AND PUBYEAR > 2001 AND PUBYEAR < 2024 AND (LIMIT-TO (DOCTYPE, 'ar') OR LIMIT-TO (DOCTYPE, 'cp')) AND (LIMIT-TO (PUBSTAGE, 'final')) AND (LIMIT-TO (SRCTYPE, 'j') OR LIMIT-TO (SRCTYPE, 'p')) AND (LIMIT-TO (LANGUAGE, 'English')) AND (LIMIT-TO (SUBJAREA, 'BUSI'))
Document types	Article and conference papers
Time interval	2002-2023
Language	English
Subject area	Business, management, and accounting
Corpus	245 documents in final set

## RESULTS AND DISCUSSION

### Descriptive Analysis

Publications on veracity published between 2002 and 2023 retrieved from Scopus displayed an upward trend. The most productive year was 2019, with 478 publications. There was a slight decrease in the number of publications in 2020, rebounding in 2021. The decline in 2020 could be attributed to challenges posed by COVID-19, which hampered research activities due to social restrictions (Fell et al., 2020). The pandemic affected the productivity of academic researchers in various ways, including reduced working hours, limited access to data, a decline in grants and sponsors, and increased time spent with family while working from home (Myers et al., 2020). In addition, several studies addressed issues related to fake news and growing concerns regarding the quality of information during the COVID-19 pandemic (Apuke & Omar, 2021; Y. C. Lee et al., 2022; Rocha et al., 2021).

Figure 2. Selection steps and analyses methods

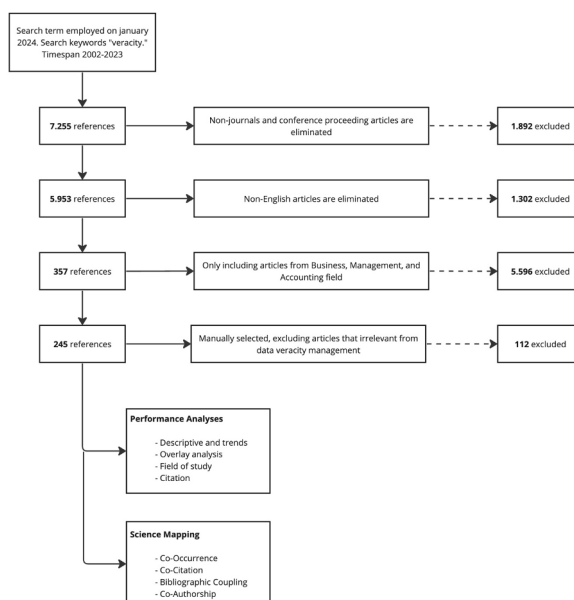
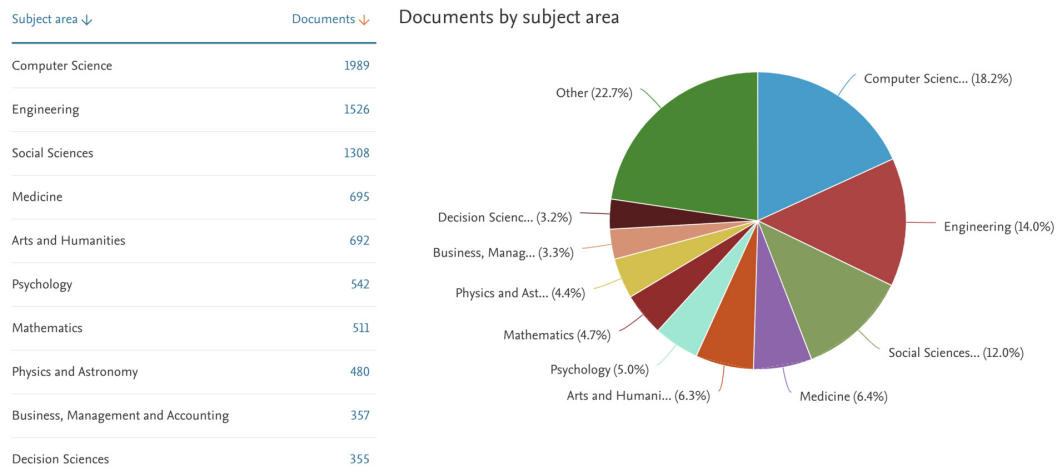


Figure 3. Percentage of field of research in veracity topic



Since the pandemic, there has been growing interest in data quality issues. This includes using machine learning systems for managing COVID-19 clinical data and employing spatial data to map known information about the virus in specific areas (Ali et al., 2022; Deng et al., 2021; Leung et al., 2020). This rise in interest has resulted in an increased number of publications since 2021. It is anticipated to contribute to sustained growth in research on data veracity.

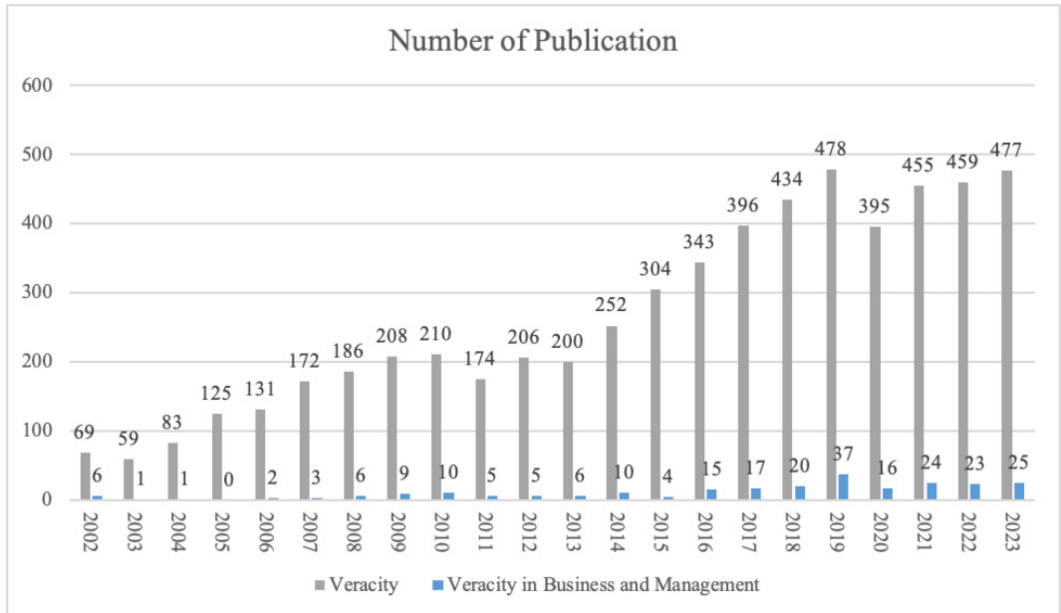
The overall number of articles published on this topic is 5,953 articles. However, only 245, or 4.12% of that number, are articles related to data veracity in business and management. This suggests that research efforts need to be more widespread.

The primary publication fields on veracity topics are computer science (1,769), engineering (1,382), and social science (1,125). Figure 2 illustrates the comparison between the number of publications on veracity in business, management, and accounting with research in other fields. In Figure 3, the graph compares the number of publications about veracity in the business and management field compared to veracity in all fields of study. The graph uses two colors to differentiate the trend between overall veracity publications (grey) and those in the business and management field (blue). It illustrates a significant and rapid increase in the number of veracity publications across all fields of study since 2012. Conversely, the number of publications related to data veracity management has shown relatively modest growth, although there is an increasing trend. Based on the two figures, it is evident that publications covering veracity in business and management are relatively minor and, thus, need more attention.

Bibliographic analyses conducted with VOSViewer generate visualization networks consisting of nodes and lines connecting keywords and colors to differentiate the cluster. Each node represents keywords; and the size of the node indicates the frequency of the appearance of each keyword. Bigger nodes indicate a higher number of the keyword's appearance. The line among nodes represents the pairing of two keywords. Nodes that are located near each other indicate a strong relationship. On the other hand, the same color nodes show a cluster of closely related keywords (van Eck & Waltman, 2014).

The first science mapping conducted is co-occurrence analysis. Co-occurrence looks for significant terms and keywords in the selected database by examining the frequency of pairs of keywords that appeared in the title, abstract, or keywords list (van Eck & Waltman, 2014). Lines connecting nodes imply the frequency of paired keywords appearing together in one publication, whereas the clusters represent a group of keywords with similar topics or research themes.

Figure 4. Number of publications on veracity topics from 2002 to 2023 (Business and management field vs. all fields)



A co-occurrence analysis was conducted on 245 publications on veracity in the management and business field. The analysis, which required a minimum of five occurrences, resulted in four clusters and 27 keywords. Figure 5 presents a network visualization of this analysis, where key terms with the most vital links are “big data” (total link strength 94), “veracity” (total link strength 51), and “data analytics” (total link strength 43). The most frequently occurring keywords in the network are “big data,” “veracity,” and “information management.” The top 10 keyword occurrences from the analysis are summarized in Table 4.

Figure 6 shows the findings of the overlay analysis on the database of business and management publications related to data veracity. Recent publications containing keywords like “fake news,” “misinformation,” “fake detection,” “social media,” and “machine learning” have been identified.

### Cluster Analysis

Applying natural language processing calculation, VOSviewer identifies keywords with strong similarities and presents them in clusters within the network. The analysis of 245 publications on data veracity in business and management revealed four significant clusters. See Figure 4.

The first keyword cluster network indicated a strong connection among machine learning, AI, social networking, and deception. Articles with these keywords addressed data validation techniques to detect poor data. Major articles focused on automatic fact-checking tools, such as adopting a computational path-based approach (Shi & Weninger, 2016) and linguistic-based cues (Fuller et al., 2009). One article conducted literature reviews to investigate computerized veracity assessment methods (García Lozano et al., 2020). Social media also became a central area of discussion within this cluster, indicated by articles grouped among keywords like fake news, information, and social media. One study found that social media search features are ineffective at assessing the truthfulness of rumors spread on these platforms (Eismann, 2021). Other research explored how conspiracy theories become the focus of attention, addressing how users’ cognitive principles deal with this growing problem (Connolly et al., 2019; Johar, 2022). Several research areas included two keyword

Table 4. Rank of keyword occurrence in data veracity management topic

Rank	Keywords	Occurrences	Total Link Strengths
1	big data	44	73
2	veracity	14	38
3	information management	14	17
4	social networking (online)	11	21
5	decision-making	10	23
6	big data analytics	10	18
7	data analytics	9	32
8	artificial intelligence	9	16
9	volume	8	31
10	velocity	8	31

groups (Liu et al., 2016; Shrivastava & Kumar, 2021), indicating the importance of validity tools and systems that effectively handle data acquired through social media.

The second cluster showed big data as the dominant keyword, indicated by the largest dot in the network. The first connected group consisted of “big data” and “big data analysis,” which were adopted by articles addressing the handling and analytics process of big data. This included a study on big data computing analytics (Gupta et al., 2018). The governance of big data management to improve quality was reviewed by another study within this cluster, highlighting the challenge of the velocity, veracity, and volume of the data (Janssen et al., 2017). Other articles focused on data analytics with various objectives, such as the analysis of customer reviews (Salehan & Kim, 2016), employee-customer relationships (Lam et al., 2017), and innovation attempts (Troisi et al., 2018). Other connections to big data keywords were identified, though these associations appeared to be

Figure 5. Network visualization theme for data veracity management

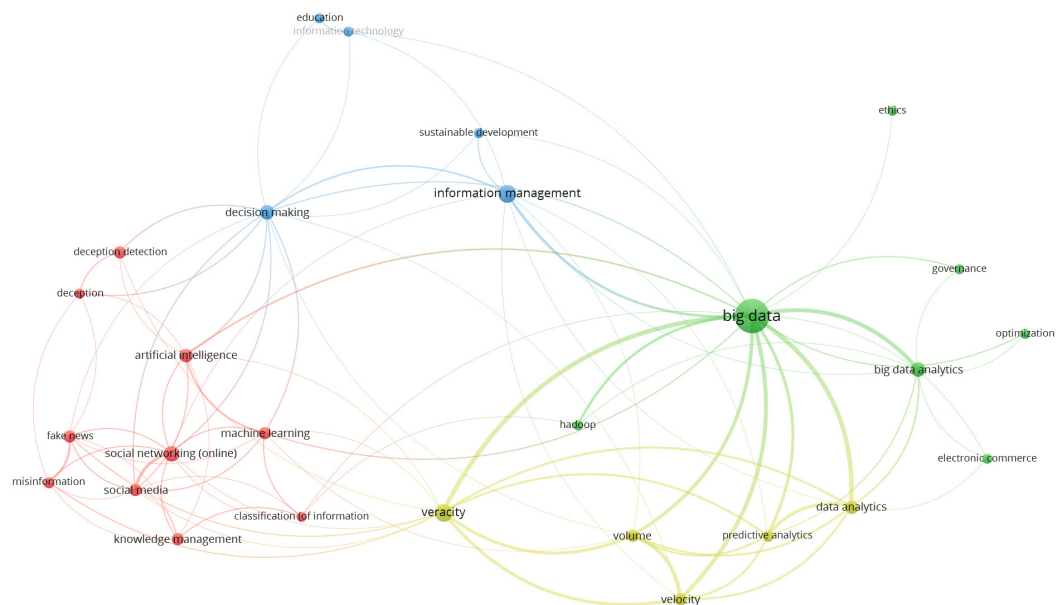
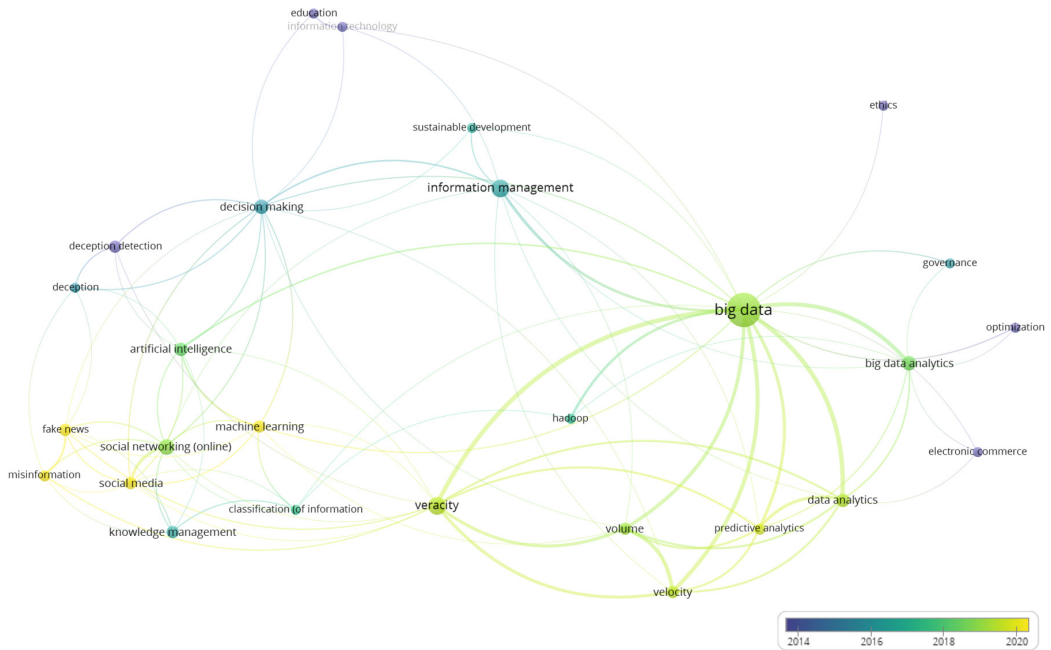


Figure 6. Overlay visualization of publications on data veracity management



weak. The linked keywords included electronic commerce, ethics, governance, and optimization. This implies that while major areas of study in big data veracity focus on data and processing analytics, other aspects—such as governance and ethics—remain underexplored.

The third cluster revealed extensive research on data veracity in the context of sustainability. Key themes included decision-making, education, information management, information technology, and sustainable development. The link found in this group implied weak connections among the keywords, pointing to an area of research that has been attempted but remains underexplored. Several articles addressed sustainability issues, including the role of big data in the circular economy and energy sectors (Jabbour et al., 2019; Riasetiawan et al., 2021). Both studies found the positive roles of data management systems and tools in supporting sustainability strategy, particularly for resource optimization and decision-making based on company data. Additionally, several studies in this cluster addressed deception in sustainability, exploring practices like greenwashing and environmental audits (Akturan, 2018; Castka et al., 2020; Mateo-Márquez et al., 2022; Yu et al., 2020). While minor, research in this cluster also addressed the decision-making process involving data and information management (Chi & Liao, 2022; Jensen et al., 2010).

The final cluster displayed a strong network connected to the second cluster, spotlighting the optimization of data processing. The keywords linked big data to characteristics like veracity, velocity, and volume. Connections also existed between big data keywords, data analysis, and predictive analysis. The strong connection between big data keywords and the fourth cluster indicated that major research in the selected datasets addressed big data and its V characteristics, including veracity. The research involved in this cluster primarily addressed implementing data optimization in various sectors, such as media, healthcare, humanitarian aid, and supply chain (Bell et al., 2021; Genevès et al., 2018; Viet et al., 2021). Big data management was found to improve supply chain performance (Alshawabkeh et al., 2022; Richey et al., 2016). One paper examined the impact of combining diverse data, bringing challenges to big data usage for humanitarian purposes (Bell et al., 2021). In terms of contributions to firm performance, one study found that data veracity did not directly impact firm performance unless

Figure 7. Citation map



it generated valuable data (Ghasemaghahi, 2021). Results from this study implied the importance of veracity strategies to focus on value generation, such as accuracy and reliability.

The study inside this cluster also implied two essential results. First, the strong connection between big data and its characteristics, including veracity, suggests that veracity is often viewed as a key feature of big data and is majorly included in related research. Second, the strong connection between big data, data analysis, and predictive analysis keywords indicates that the articles in this cluster have a strong connection with the second cluster. This is supported by the similarity of research themes that address big data analysis and processing. However, the study in this cluster is dominated by the implementation of data processing and analysis in various sectors as compared to the second cluster.

### Citation, Co-Citation, and Bibliographic Coupling Analysis

Citation analysis was conducted to identify the most cited articles in the database, which indicates publications with high impact on this topic. Figure 7 shows the cluster based on citation analysis, with a more significant node indicating the high number of citations of the articles. Of the 245 articles, 177 are cited at least once. The most cited article to date, “Predicting the Performance of Online Consumer Reviews: A Sentiment Mining Approach to Big Data Analytics,” has 380 citations (Salehan & Kim, 2016). The second most-cited article, with 346 citations, is “Factors Influencing Big Data Decision-Making Quality” by Janssen et al. (2017). The third most-cited article is “Unlocking the Circular Economy Through New Business Models Based on Large-Scale Data: An Integrative Framework and Research Agenda” by Jabbour et al. (2019), with 213 citations to date. The high number of clusters and few direct links (three) in Figure 5 indicates the dispersion among articles.

Table 5 represents the geographical distributions of the articles regarding the origin countries and organizations where the articles were published. The United States was the country with the most publications on data veracity management (62), followed by China (29) and India (26). The United States also came in first for the most-cited country, with 2,275 citations, followed by India (464) and France (443).

The co-citation analysis has been conducted in this study, looking for references cited by selected articles or databases (van Eck & Waltman, 2014). The authors conducted the co-citation analysis using VOSviewer software from 245 observed articles, limiting the references to those cited at least three times. This resulted in 16 selected references.

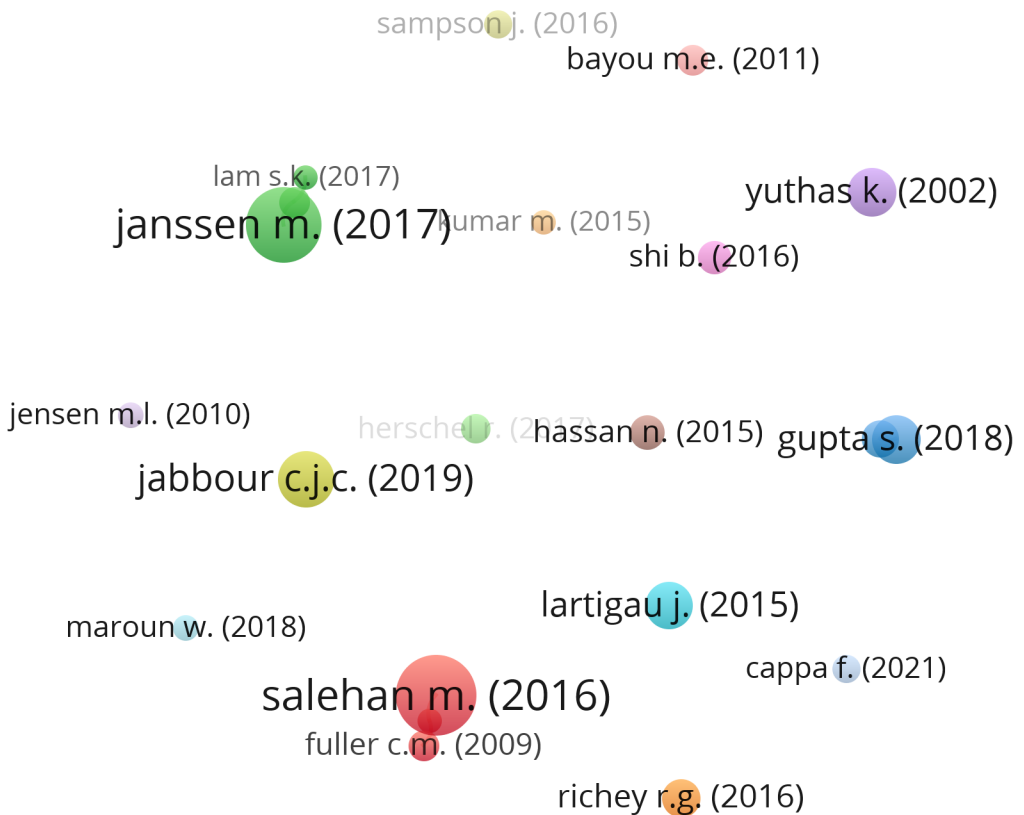
Figure 8 visualizes the network output from the co-citation analysis. Three clusters are generated, indicating three thematic similarities identified from co-citation patterns inside the selected database (van Eck & Waltman, 2014). The selected references are summarized in Table 6.

Figure 8, which displays the results of the co-citation analysis, shows that three major reference clusters were predominantly cited by articles in the database. Cluster 1 included publications that cited topics related to big data and big data analytics, such as the influence of big data on firm performance (Aker et al., 2016; Müller et al., 2018; Wamba et al., 2017), big data for business analytics (Chen et al., 2012; Erevelles et al., 2016; Kwon et al., 2014), and the concept, methods, and challenges of big data (Gandomi & Haider, 2015; Janssen et al., 2017; Sivarajah et al., 2020). In Cluster 2, the publications cited focused on online reviews, specifically on the deception and manipulation of customer online reviews (Luca & Zervas, 2016; Mayzlin et al., 2012) and the impact of online reviews on customers’ decision-making and judgment (Chevalier & Mayzlin, 2006; Tversky & Kahneman, 2005). Cluster 3

Table 5. Ten countries with the most publications

Countries	Number of Publications	Number of Citations
United States	62	2,275
China	29	279
India	26	464
Australia	21	214
United Kingdom	18	317
France	11	443
Germany	10	182
Canada	10	207
Netherlands	6	349
Singapore	6	29

Figure 8. Authors' co-citation map



centered on the topic of deception, with most-co-cited publications discussing judgment and detection of deception (Bond et al., 1990; Bond & Depaulo, 2006; Buller & Burgoon, 1996).

Table 6. Topics of most-co-cited references

Clusters	Articles	Keywords
Cluster 1	Big data analytics and firm performance: Effects of dynamic capabilities (Wamba et al., 2017)	big data analytics, big data analytics capability, business values, dynamic capabilities, firm performance, econometric analysis, firm performance, resource-based view, data quality management, IT capability, data usage, e-government, decision-making quality, big data definition, unstructured data analytics, predictive analytics, consumer analytics
	Critical analysis of big data challenges and analytical methods (Sivarajah et al., 2017)	
	The effect of big data and analytics on firm performance: An econometric analysis considering industry characteristics (Müller et al., 2018)	
	Data quality management, data usage experience and acquisition intention of big data analytics (Kwon et al., 2014)	
	Factors influencing big data decision-making quality (Janssen et al., 2017)	
	Beyond the hype: Big data concepts, methods, and analytics (Gandomi & Haider, 2015)	
	Big data consumer analytics and the transformation of marketing (Erevelles et al., 2016)	
	Business intelligence and analytics: From big data to big impact (Chen et al., 2012)	
	How to improve firm performance using big data analytics capability and business strategy alignment? (Akter et al., 2016)	
Cluster 2	Promotional reviews: Empirical investigation of online review manipulation (Mayzlin et al., 2012)	online reviews, word-Of-mouth, consumer decision-making, consumer behavior, manipulation, marketing, advertising and media, organizational studies, judgement, subjective assessment
	The effect of word of mouth on sales-online book reviews (Chevalier & Mayzlin, 2006)	
	Fake it till you make it: Reputation, competition, and Yelp review fraud (Luca & Zervas, 2016)	
	Judgement under uncertainty: Heuristics and biases (Tversky & Kahneman, 2005)	
Cluster 3	Interpersonal deception theory (Buller & Burgoon, 1996)	deception theory, interpersonal communication, credibility, truthfulness, deception judgement, cultural studies, nonverbal behavior
	Accuracy of deception judgments (Bond & DePaulo, 2006)	
	Lie detection across cultures (Bond et al., 1990)	

Supporting the co-citation analysis, this research also conducted a bibliographic coupling analysis. The bibliographic coupling analysis addresses articles that cited another publication to determine the list of overlapping references used in the designated topic (van Eck & Waltman, 2014). The bibliographic coupling analysis was conducted using VOSviewer software. It selected articles with 50 or more citations. This setup focuses on the articles that contributed to the largest source of references for data veracity management.

As in Figure 9, the analysis result showed four clusters and 10 connected articles. Each has a more significant similarity of references with other articles in the clusters. The article with the most significant bibliographic coupling, “Big Data for Creating and Capturing Value in the Digitalized Environment: Unpacking the Effects of Volume, Variety, and Veracity on Firm Performance,” is authored by Cappa et al. (2021). The study has more significant similarities in references as compared to the other articles. It also has the most articles that combine other articles in the domain, with 18 total link strengths. Articles with greater bibliographic coupling in each cluster should also be identified to address the similarities of the topic in the cluster, looking at its reference similarities. The dominant articles and authors are presented in Table 7.



Figure 9. Map of bibliographic coupling



Table 7. Clusters of bibliographic coupling analysis

Cluster	Authors	Research
1	Gupta et al. (2018)	Big data with cognitive computing: A review for the future
	Kushwaha et al. (2021)	Applications of big data in emerging management disciplines: A literature review using text mining
	Jabbour et al. (2019)	Unlocking the circular economy through new business models based on large-scale data: An integrative framework and research agenda
2	Cappa et al. (2021)	Big data for creating and capturing value in the digitalized environment: Unpacking the effects of volume, variety, and veracity on firm performance
	Lam et al. (2017)	Leveraging frontline employees' small data and firm-level big data in frontline management: An absorptive capacity perspective
	Ghasemaghaei and Calic (2019)	Does big data enhance firm innovation competency? The mediating role of data-driven insights
3	Janssen et al. (2017)	Factors influencing big data decision-making quality
	Richey et al. (2016)	A global exploration of big data in the supply chain
4	Salehan and Kim (2016)	Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics
	Chatterjee (2020)	Drivers of helpfulness of online hotel reviews: A sentiment and emotion mining approach

### Co-Authorship Analysis

An authorship analysis was conducted to identify the level of cooperation among scholars in the designated field of study. First, the authors identified the number of articles published by each author. Table 8 lists the most productive authors in the designated field, with Schweitzer and Xu listed as the top two (three publications). The top 10 productive authors showed that they each produced two or three papers on this topic. Out of the 598 registered authors, the majority published only one paper in this field.

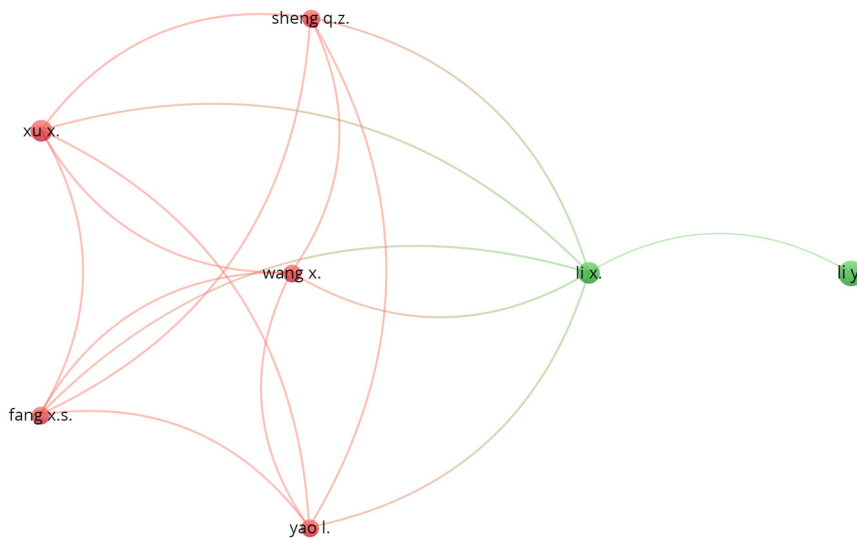
The authors also conducted a co-authorship analysis using VOSViewer to identify collaboration attempts between registered authors. Figure 10 represents the author collaboration network in the database. They selected authors with a minimum of two publications related to veracity research, resulting in a total of 23 out of 598 registered authors. The network shows that among the 23 authors, only 7 authors are connected in several registered publications.

An authorship analysis based on the country of origin was also conducted (see Figure 11). The countries were derived by the origin of the country where the articles were published and recorded in the Scopus database. The network shows that the United States (62), China (29), India (26), Australia (21), and the United Kingdom (18) are identified as the countries with the highest number of publications. Regarding collaboration among countries, the network shows that collaboration

Table 8. Most productive authors

Author's Name	Number of Publications	Country of Origin
Schweitzer, M. E.	3	United States
Xu, X.	3	China
Dillard, J. F.	2	United States
Fang, X. S.	2	China
Ghasemaghaei, M.	2	Canada
Gupta, S.	2	France
Harding, N.	2	Australia
Kar, A. K.	2	India
Li, X.	2	Australia
Nakov, P.	2	United Arab Emirates

Figure 10. Authorship network



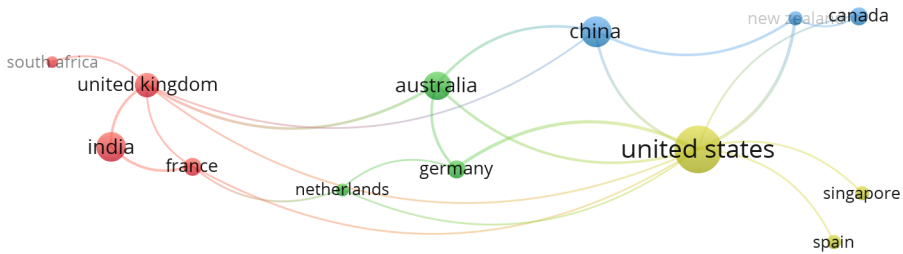
between various countries and continents has occurred, with the strongest collaboration between the United States and Germany, resulting in three publications.

## DISCUSSION

### Development of Veracity Research in the Business and Management Field

This study employed a bibliographic analysis to examine research trends in business and management related to veracity over the past two decades, emphasizing the number of publications. The development of research between all registered topics on veracity and those in the business and management fields was compared. Finally, an overlay network analysis was used to identify the most discussed topics within the designated timeline.

Figure 11. Authorship network based on country of origin



The number of publications on data veracity and data veracity research related to the business and management field were calculated and compared. The objective was to discern trends in data veracity research to date. The findings are illustrated in Figure 2, indicating that while publications on veracity continue to rise, there is a need for increased focus on business and management. The dominance of computer science and engineering publications may also suggest that the development of veracity studies may be more oriented toward tools and applications rather than managerial aspects. This finding aligns with previous research assertions advocating for expanding data veracity research beyond the focus on systems and tools (e.g., Meel & Vishwakarma, 2020; Zhang et al., 2018).

There are significant opportunities for exploring aspects related to business and management, considering the continuous digitalization of business processes and data-driven business practices and the growing complexity involved in maintaining data quality (Côte-Real et al., 2020; Trabucchi & Buganza, 2019). Several publications in this study have proposed findings that could potentially serve as supporting theories for further research in business and management. These findings cover diverse areas like the decision-making process (Janssen et al., 2017), marketing (Salehan & Kim, 2016), supply chain (Castka et al., 2020), human resources (Lam et al., 2017), accounting and finance (Maroun, 2018), and management issues like sustainability (Jabbour et al., 2019), innovation (Ghasemaghaei, 2021), and ethics (Herschel & Miori, 2017).

The overlay network analysis was conducted to trace the evolution of research focused on data veracity, using the publication year as a parameter. Figure 5 illustrates the shifting focus of study in the network from 2014 to 2022 when most publications were released. The green nodes and lines in the network indicate research on big data and veracity prevalent from 2017 to 2019. During this period, most publications in this area discussed the characteristics of big data and methods or tools for promoting veracity in big data systems. In contrast, publications from 2014 to 2017 covered various topics, including education, ethics, electronic commerce, and deception detection, as indicated by the blue network. During this timeframe, the topics covered reflected concerns about big data and its quality. For instance, studies examined the effectiveness of online customer review data in providing valuable information for readers (Salehan & Kim, 2016).

Ethical considerations related to the sharing and use of big data were explored (Herschel & Miori, 2017). In publications spanning from 2018 to 2022, the focus shifted toward issues of information deception, with keywords like misinformation, fake news, and social media. Highly cited publications in this domain included research on the development of sarcasm detection in social media text (Shrivastava & Kumar, 2021), an investigation into the impact of accounting disclosure on stock performance (Eachempati et al., 2021), and the identification of rumor detection and correction on the Twitter platform (Eismann, 2021).

The overlay analysis showed a transition in the focus of issues within the veracity research topic over time. Initial publications on this topic focused on identifying and emphasizing the significance of big data, along with strategies to maintain its quality across various sectors. Research attempts to address this issue have increased, particularly following the introduction of big data management in

2012 (Brynjolfsson & McAfee, 2012). During the peak publication period for big data and veracity research, specifically from 2018 to 2019, there was a notable proliferation of models and tools aimed at maintaining big data veracity. Subsequently, these models were increasingly implemented across various data systems.

Beginning in 2019, publications shifted their focus to the issue of deception and fake news in big data systems and social media. This trend may be attributed to growing concerns about the widespread dissemination of misinformation and false narratives on the Internet, posing a significant challenge for data users who seek reliable information (Johar, 2022; Nakov & Da San Martino, 2021).

## Research Themes

Research publications within each cluster were examined to identify themes and prevalent research topics discussed to date. The four clusters were categorized into three major research themes: (1) T1 - data validation (Cluster 1); (2) T2 - data processing optimization (Clusters 2 and 4); and (3) T3 - data veracity in the sustainability domain (Cluster 3). These themes were generated from the co-occurrence network (see Figure 12). Proposed themes from keywords and dominant publications from each cluster are listed in Table 9.

The analysis results indicated that research on data validation and social media were predominantly grouped together as a main theme in Cluster 1. The connection between Clusters 2 and 4 showed an interesting pattern, in which a strong connection appears among keywords in the clusters, particularly involving big data. The topics in Clusters 2 and 4 address big data processing techniques and their implementation through various sectors. Considering the strong connection and similarity of the topic within each cluster, this research proposes a combined theme from Clusters 2 and 4.

The last theme, specifically for Cluster 3, addresses data veracity for sustainability context. The cluster's network was weak, indicating that limited research was involved in this cluster. However, the cluster shows that this research topic provides distinct connected keywords that need to be proposed in their own theme. The detailed outcomes and interpretation of selected themes are provided in the following subsection.

### *Theme 1 - Data Validation*

Cluster 1 featured publications about veracity assessment, presenting various approaches for data validation methods. The significance of data validation has increased in recent years with studies addressing data management. This is driven by the growing emphasis on data quality and value generation (Gao et al., 2016).

The expansion of data volume, velocity, and variety poses potential challenges for businesses. Additionally, data growth is impacted by the increasing number of connected devices to the Internet and data-sharing activities within or outside firms' databases (Gandomi & Haider, 2015). This growth, however, entails additional data maintenance. Companies worldwide are projected to spend US\$12.73 per employee on data storage, with an anticipated increase to 16.61 cents by 2028. This signifies a substantial investment in storing and managing a company's big data (Statista, 2023). Considering the substantial expenditure on big data management, companies must ensure that the data yields valuable advantages for their business. In this context, the data validation process plays a crucial role in supporting data management by sorting and selecting valuable data while eliminating unused data. This optimization aims to streamline the volume of data within the system.

The results from Cluster 1 suggest a close relationship between veracity efforts and data validation. Several studies confirm that veracity plays a crucial role in enhancing the quality of big data and contributes to converting data volume, velocity, and variety into data value (Ramachandramurthy et al., 2015; Rubin & Lukoianova, 2014). Veracity-focused publications on data validation address the tools and methods for conducting fact-checking on gathered data, with an emphasis on automated tools.

Automated fact-checking tools are essential in managing big data given the diverse range of information gathered, specifically from sources like social media and other open platforms (García

Table 9. Themes, keywords, and dominant publications from each cluster

Clusters	Themes	Keywords	Publications	References	Contexts
1	Data validation methods	machine learning, AI, social networking, deception	Discriminative predicate path mining for fact checking in knowledge graphs	Shi and Weninger (2016)	A computational method for fact-checking statements using a discriminative path-based approach
			Decision support for determining veracity via linguistic-based cues	Fuller et al. (2009)	Examines automated text-based deception detection methods, which utilize linguistic cues to classify statements
			Veracity assessment of online data	García Lozano et al. (2020)	Reviews the current literature on computerized veracity assessment methods
			A pragmatic and intelligent model for sarcasm detection in social media text	Shrivastava and Kumar (2021)	Proposing a novel sarcasm detection model based on Google BERT, designed to handle large, fast, and complex data
			Reuters tracer: A large scale system of detecting & verifying real-time news events from twitter	Liu et al. (2016)	A system designed to help journalists quickly discover and verify news events on Twitter by using machine learning models
			Diffusion and persistence of false rumors in social media networks: Implications of searchability on rumor self-correction on Twitter	Eismann (2021)	Simulations to show that while search features on social media like Facebook and Twitter can make it harder to judge the truth of rumors, interactions between connected users can help spread accurate information and corrections
			Untangling the web of misinformation and false beliefs	Johar (2022)	Explore cognitive principles that shape beliefs and suggest strategies for preventing false beliefs in the misinformation crisis
			Communicating to the public in the era of conspiracy theory	Connolly et al. (2019)	Synthesizes research on conspiracy theories and public administration, highlighting the rapid spread and widespread belief in conspiracy theories
2	Data processing and implementation	big data, big data analytics	Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics	Salehan and Kim (2016)	Investigates predictors of readership and helpfulness in online consumer reviews (OCRs) using sentiment mining for big data analytics
			Factors influencing big data decision-making quality	Janssen et al. (2017)	Highlighting that the veracity, variety, and velocity of big data necessitate relational and contractual governance mechanisms to ensure data quality and contextualization

continued on following page

Table 9. Continued

Clusters	Themes	Keywords	Publications	References	Contexts
			Big data analytics with swarm intelligence	Cheng et al. (2016)	Exploration of the potential application of swarm intelligence techniques in big data analytics, discussing the correspondence between the characteristics of big data and the features of swarm intelligence algorithms
			Big data with cognitive computing: A review for the future	Gupta et al. (2018)	Reviews the past, present, and future of big data and cognitive computing through a systematic literature review using Scopus, DBLP, and Web of Science databases, highlighting that the field is still in its early stages
			Big data management: The case of Mulino Bianco's engagement platform for value co-creation	Troisi et al. (2018)	Demonstrates that engagement platforms help acquire and manage integrated, clean, and trustworthy data for decision-making and innovation
		big data, velocity, veracity, volume	Understanding the impact of big data on firm performance: The necessity of conceptually differentiating among big data characteristics	Ghasemaghaei (2021)	Examines the impact of the main characteristics of big data (data volume, data velocity, and data variety) on firm performance using the resource-based view, and explores the mediating roles of data value and data veracity
			Exploring future challenges for big data in the humanitarian domain	Bell et al. (2021)	Explores the challenges of using big data in the humanitarian sector to support United Nations Sustainable Development Goal 17, showing that the assumption that diverse data can be easily combined is problematic due to differences in data detail
			Big data and health care system using mlearning	Manoharan (2019)	Examines how analyzing data from people of different ages and lifestyles can help predict future illnesses by collecting information on symptoms, disease stages, treatments, side effects, and outcomes
			A global exploration of big data in the supply chain: Global exploration of big data	Richey et al. (2016)	Defines big data in supply chain management through four key dimensions (volume, velocity, variety, and veracity) based on insights from supply chain managers across six nations, and identifies critical success factors and future research opportunities

*continued on following page*

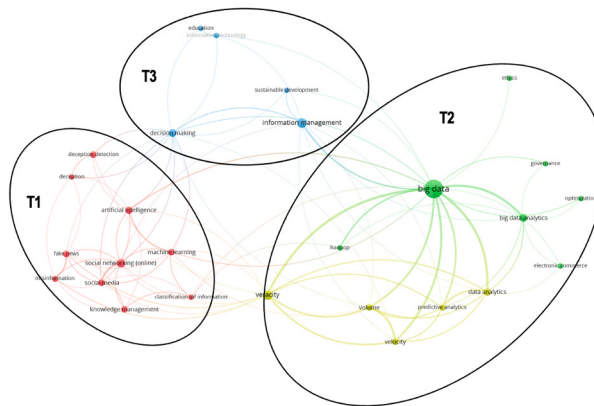
Table 9. Continued

Clusters	Themes	Keywords	Publications	References	Contexts
			The mediating role of supply chain management on the relationship between big data and supply chain performance using SCOR model	Alshwabkeh et al. (2022)	Investigates the impact of big data on supply chain performance, mediated by supply chain management activities using the SCOR model, and finds that big data significantly improves both supply chain management and performance
3	Data veracity in the sustainability domain	information management, decision making, sustainable development	Unlocking the circular economy through new business models based on large-scale data: An integrative framework and research agenda	Jabbour et al. (2019)	Integrating the circular economy with large-scale data management, highlighting ReSOLVE-based CE models, stakeholder roles in sustainability, and the 4Vs of LD
			An international empirical study of greenwashing and voluntary carbon disclosure	Mateo-Márquez et al. (2022)	Introduces a method for identifying greenwashing firms based on carbon performance and disclosure, contributing to the literature by applying institutional theory to explain corporate greenwashing behavior
			Data model and analysis for big data mapping and management in the energy data platform	Riasetiawan et al. (2021)	Focuses on developing a data model for big data storage in the energy sector, addressing the challenges posed by the volume, variety, veracity, and velocity of energy data
			On veracity in corporate sustainability claims: Why society should be cautious	Ngwakwe and Netswera (2013)	Critiques the veracity of corporate sustainability claims and warns of the societal risks posed by false claims, including hidden carbon emissions, and increased socio-economic inequality
			Technology dominance in complex decision making: The case of aided credibility assessment	Jensen et al. (2010)	Explores the theory of technology dominance in the context of credibility assessment, testing a decision aid designed to help both novice and professional users evaluate message veracity

Lozano et al., 2020). However, relying solely on manual selection by experts may be insufficient to address the pace of retrieved information (Shi & Weninger, 2016). Within the cluster, there is a prevalent emphasis on automatic fact-checking approaches (Genevès et al., 2018; Hassan et al., 2015; Sampson et al., 2016; Shi & Weninger, 2016; Twyman et al., 2014). Several studies also explore semi-automated tools that combine human and computer abilities for fact-checking (Jensen et al., 2010; Martín et al., 2022).

Regarding the research objective, the data sources selected for analysis are primarily social media platforms (Chi & Liao, 2022; Liu et al., 2016; Lukasik et al., 2019; Martín et al., 2022; Sampson et al., 2016; Shrivastava & Kumar, 2021). The cluster also includes other data sources, such as firms' audit

Figure 12. Research themes groups



data (Eachempati et al., 2021; C. C. Lee et al., 2013), knowledge sources like Wikipedia (Chernyavskiy et al., 2021; Shi & Weninger, 2016), and chat applications (Zhou & Zhang, 2007). The increasing number of publications about data validation on social media platforms suggests that contemporary businesses rely heavily on open-source platforms like social media for data and information. However, it also underscores the importance of considering information quality issues related to social media (García Lozano et al., 2020; Muninger et al., 2019; Sivaraman et al., 2020).

### Theme 2 - Data Processing Optimization

The changing nature of big data highlights its focus on managing the growing volume of retrieved data and extracting valuable advantages for business operations (Geerts & O’Leary, 2022; Rubin & Lukoianova, 2014). Cluster 2 publications predominantly address the central theme of processing and analytics techniques to provide valuable information for decision-makers (Gandomi & Haider, 2015; Janssen et al., 2017). Big data analytics is driven by the necessity to analyze extensive and rapidly generated datasets, aiming to acquire intelligence that supports firms’ competitive performance and identifies opportunities (Gandomi & Haider, 2015; Mikalef et al., 2019).

Big data analytics provides various capabilities to support decision-making processes, including descriptive, inferential, predictive, explanatory, causal, and mechanistic approaches (Janssen et al., 2017). Data analytics processing techniques range from natural language processing to machine learning and syntax and lexical analysis (García Lozano et al., 2020; Janssen et al., 2017).

However, large datasets characterized solely by high volume, variety, and velocity (without ensuring high data veracity) may increase data uncertainties and complexity. These challenges are considered significant in pursuing successful big data analytics (Dai et al., 2020; Geerts & O’Leary, 2022; Hariri et al., 2019). In this context, ensuring data veracity becomes crucial for enhancing information quality, a vital requirement for accurate analytics that supports effective decision-making.

Several aspects of data veracity, such as data accuracy and reliability, are highlighted as essential factors contributing to high-quality data analysis outcomes (Cappa et al., 2021; Janssen et al., 2017; Richey et al., 2016). The emphasis of this research on aligning data veracity with big data analytics underscores the significance of initiatives aimed at maintaining or improving the quality of data that will be analyzed for decision-making processes or other managerial purposes.

The cluster analysis results emphasize the widespread implementation of big data and its veracity across business sectors. Among the extensively studied areas, the news or journalism sector stands out, with numerous publications addressing efforts to combat fake news and improve the speed of information processing (e.g., Hassan et al., 2015; Liu et al., 2016; Nakov & Da San Martino, 2021; Nguyen et al., 2020; Shaar et al., 2020). This finding indicates the growing concern about the



prevalence of fake news on contemporary information platforms, particularly online (George et al., 2021; Wu et al., 2022).

Social media has become an essential source of information that people and organizations use to quickly retrieve and share information (Liu et al., 2016; Sivarajah et al., 2020; Troisi et al., 2018). Businesses often use social platforms during emergency events, driving users to access data that is free and efficient to share (Khaldarova & Pantti, 2016; Mirbabaie et al., 2021). Nevertheless, the freedom of information sharing is followed by uncertainty about the quality of the information being shared. Several published studies highlight the drawbacks of information quality in social media, which must be overcome through veracity attempts. These challenges include the absence of gatekeepers to ensure the credibility of shared news and the potential for misinterpretation when news is retrieved and reshared on social media platforms (George et al., 2021; Nakov & Da San Martino, 2021; Suthanthiradevi & Karthika, 2019). The proposed veracity practices are primarily related to the development of fact-checking tools and their potential implementation on social media platforms (Chi & Liao, 2022; Eismann, 2021; Lukasik et al., 2019; Sampson et al., 2016; Shrivastava & Kumar, 2021; Suthanthiradevi & Karthika, 2019; Zubiaga et al., 2018). Additionally, other publications discuss veracity approaches to capture customer reviews on online platforms (Martin-Fuentes et al., 2018; Salehan & Kim, 2016).

While fact-checking remains a prominent research theme in the context of data veracity in business practices, some publications focus on the veracity of accounting and audit data. The increasing complexity of accounting data calls for initiatives to improve the auditing process and ensure the transparency of the processed data (Bayou et al., 2011). Much of the registered research related to accounting and audit addresses the importance of data objectivity and the risks of fraud in firms' reports. Factors influencing the objectivity of firms' disclosure include sentiments (Eachempati et al., 2021), conflicting external and internal evidence (Trotman & Wright, 2012), and the quality of the internal control system (Gal & Akisik, 2020). Automated veracity tools may improve the auditing process, addressing concerns about the potential lack of objectivity and consistency in assessments by human auditors when evaluating a company's reports (Castka et al., 2020).

Papers focused on accounting practices highlight that veracity characteristics, mainly transparency and objectivity of audit files and systems, are crucial issues in accounting. Calling for veracity improvement in accounting was motivated by various factors outlined in the registered studies, including challenges associated with misleading financial reports (Bayou et al., 2011; Emeseh & Songi, 2014), the need for enhancement in assurance methods (Castka et al., 2020; Maroun, 2018; Wang et al., 2022), and communication issues during the auditing process (Harding & Trotman, 2017; C. C. Lee et al., 2013).

### *Theme 3 - Data Veracity in Sustainability Performance*

Several publications also address the veracity of firms' sustainability performance, extending the scope beyond extensive research in accounting and auditing. In contemporary business practices, organizations are increasingly accountable to the shareholders for their financial performance and socio-environmental impact (Ackers, 2017). This responsibility requires firms to provide sustainability reports to stakeholders. Several studies in the field of accounting and auditing veracity address the complexity of sustainability reports, recognizing factors that may affect their veracity (Castka et al., 2020; Wanner & Janiesch, 2019). These studies highlight that sustainability strategies, such as the circular economy, require integration and collaboration among stakeholders involved in the innovation process (Despeisse et al., 2017). This collaborative process entails extensive data gathering and sharing activities among stakeholders, underscoring the need for a reliable veracity system to ensure data quality for innovation (Jabbour et al., 2019). In supporting the innovation process, veracity has been shown to enhance innovation competency by providing highly credible and reliable data, thereby increasing descriptive, predictive, and prescriptive insight (Ghasemaghahi & Calic, 2019).

Another study also highlighted concerns about the quality of sustainability assurance, revealing the potential for firms to conceal unethical practices. Some firms may intentionally provide exaggerated reports of their sustainability performance to gain advantages. This phenomenon is known as greenwashing (Kuruppu & Milne, 2010; Wanner & Janiesch, 2019). In several countries, businesses consider corporate social responsibility (CSR) a voluntary initiative. However, this may be abused to uphold a company's positive image and reputation (Emeseh & Songi, 2014). This prompts concerns about the reliability of reports received by stakeholders. The perceived credibility of the information provided is contingent on the completeness and accuracy of the data, which is affected by the level of variety and veracity of the provided data (Wanner & Janiesch, 2019).

There is also a call for regulatory pressures, which have been found to be more effective in curbing greenwashing practices than relying on voluntary initiatives (Mateo-Márquez et al., 2022). The results of the increasing research on sustainability performance highlight two critical issues related to data veracity. First, implementing sustainability strategies requires big data management and veracity systems due to the extensive collaboration and integration involved. Second, there is a significant emphasis on ensuring the veracity of sustainability reports through a robust assurance system, including legal consequences to address misreporting and greenwashing practices.

The business applications of data veracity, as explored in the documented publications, extend beyond the previously mentioned areas, particularly in recent publications. These include topics like data veracity for marketing and customer relationship purposes (Azimi et al., 2022; Chatterjee, 2020; Kollmer et al., 2022; Salehan & Kim, 2016), healthcare (Genevès et al., 2018; Segers, 2022; Stone et al., 2021), humanitarian efforts (Bell et al., 2021), and supply chain management (Alshawabkeh et al., 2022; Viet et al., 2021). Given the increasing emphasis on data quality and the adoption of big data systems in various business sectors, it is anticipated that research on developing and implementing data veracity theories will continue to increase.

Sustainability, accounting, and social media emerge as significant publication topics in data veracity practices. Understandably, accounting and social media emerge as prominent concerns in data veracity research due to their susceptibility to fake data and deception. However, emerging studies connecting data veracity with sustainability offer new insights into this research domain. While the acknowledged value of big data veracity lies in any valuable data that supports firms in decision-making and financial advantages (Geerts & O'Leary, 2022; Rubin & Lukoianova, 2014), the shift toward sustainability emphasizes the value of veracity for both the company's benefit and for socio-environmental advancement.

## Key Publications

The authors conducted various analyses using VOSviewer software to identify impactful publications in data veracity in the business and management field. Analyses included co-citation, bibliographic coupling, and citation, helping to shed light on publications that were adopted as references and offering clues about the future direction of the research topic that would be studied. The selected publications were manually reviewed through their content and roles as references for other studies.

Using co-citation and bibliographic coupling analyses, this study focused on a list of publications that served as primary sources for developing publications related to data veracity. The results from co-citation analysis indicate that publications related to big data, online reviews, and deception issues have become the primary sources for developing publications on data veracity.

Data veracity is commonly identified as one characteristic of big data, making it a central focus in research on big data (Cappa et al., 2021; Gandomi & Haider, 2015; Rubin & Lukoianova, 2014). It has also been noted that data quality impacts big data analytics and its value in terms of firm performance (Akter et al., 2016; Müller et al., 2018), with veracity playing a pivotal role in this context. Additionally, publications on online reviews served as the foundation for research on data veracity in the database, which is crucial as the credibility and reliability of online reviews have

been questioned (Luca & Zervas, 2016; Mayzlin et al., 2012). This finding suggests that information gathered from online platforms should be thoroughly evaluated and verified, highlighting the need for methods to assess the veracity of online information.

Lastly, publications on deception and manipulation have also influenced research on data veracity in the database, indicating that these issues provide a basis for studying data veracity. The complexity of data handling and processing may cause an increased risk of data deception and misrepresentation, which can have a negative impact on data analytics. The publications included in the cluster emphasize the interpersonal representation of deception and culture, indicating the higher human roles in influencing the veracity of managed data (Bond et al., 1990; Bond & Depaulo, 2006; Tversky & Kahneman, 2005). In fact, human judgment in assessing data plays an important role, although the result can be subjective. This serves as a basis for calling for research on the impact of psychological factors in supporting the rapid development of technical methods and tools to foster data veracity in the business environment.

In comparison, the bibliographic coupling analysis result showed that the publication by Cappa et al. (2021) has the most vital link. The topic covered by this study links big data characteristics with data value. This was also addressed by other publications in the list (e.g., Geerts & O'Leary, 2022; Ghasemaghahi, 2021; Rubin & Lukoianova, 2014; Shams & Solima, 2019) through the use of similar sets of references. The result of the bibliographic coupling analysis may serve as a basis for predicting the significance of big data. It could be a key area of research, given the number of articles that touch on its value.

This aligns with the findings of the co-occurrence analysis, which identified big data as the most discussed keyword in the database. However, as the outcome of both co-citation and bibliographic coupling analyses were able to identify the thematic and intellectual connections inside the research landscape of data veracity in business and management fields, the networks generated from both analyses in Figures 9 and 10 also showed weak links with only a few polarized references in each cluster. This may indicate that the results of both co-citation and bibliographic coupling analyses in this study may only portray a group of interrelated studies. It may imply that the connection among references is yet to give polarized dominant trends about publication topics. It would be essential to keep updating the trend about data veracity topics based on future publications to address the trends of research topics.

The most-cited article, presented by Salehan and Kim (2016), provided a perspective on customers' sentiments to perceive product reviews read in online marketplaces. The study stated that positive sentiment in the review title would influence users. A neutral polarity of the review would give perceived helpfulness. In comparison, the longevity of the review would influence both attention and helpfulness. The authors proposed the results as considerable factors in developing an automated review classification system for online review datasets, helping firms filter valuable customer reviews and assisting customers in obtaining a qualified review. The study has been cited by 380 publications, with most focusing on online customer review data quality in various platforms and sectors. Several of the citing publications for this article addressed the reviewers' trustworthiness of online reviews about hospitality products (Chatterjee et al., 2023), as well as the identification of critical factors of adaptive online clothing customers via online customer reviews (Li et al., 2023).

Citing publications in this article strengthen the argument that an online customer review is one significant source of information for customers or companies to determine product strategies, as stated by prior studies (e.g., de Maeyer, 2012; Zhu et al., 2010). However, customers' freedom to write reviews may also bring trustworthiness and issues related to information quality (Banerjee et al., 2017). Calling for attempts to improve the veracity of the reviews would be essential to bring valuable information to review users.

The second most-cited article on the list, by Janssen et al. (2017), addressed essential factors of big data decision-making quality, with 346 citations. The study revealed the importance of the routinization process and gradual understanding of the analytic system to get valuable advantages from

the big data analytics system. The study emphasizes that effective big data analytics is an evolving process requiring knowledge transfer, employee skills and competencies, and proper integration. Quality of data management is also mentioned as a crucial factor for consideration. In fact, data volatility may hamper the analytic process.

This research has been cited by several studies addressing similar topics. Dubey et al. (2019) explored factors that influence big data and predictive analytics to improve cost and operational performance. Mikalef et al. (2020) identified a relationship between big data analytics capability and competitive performance.

Exploring the essential roles of big data for business performance, data quality would become an essential factor for consideration. Uncompleted or unstructured data would obstruct the analytics process, potentially generating improper decisions (Janssen et al., 2017). Poor data quality also brings costly consequences for firms, leading to billions of lost dollars each year (Ghasemaghaei & Calic, 2019; Hazen et al., 2014).

Data quality is closely linked to data veracity, as intrinsic data quality emphasizes that data should be credible and unbiased (Ghasemaghaei & Calic, 2019; Webb & Webb, 2004). Data development of research related to data veracity and analytics should continue to address data quality for optimal business analytics processes.

Jabbour et al. (2019) was the third most-cited article, with 213 citations. The research aims to integrate the circular economy with big data characteristics, proposing a framework for further study. Following the increasing number of publications on big data and sustainability performance, circular economies have become a widely discussed topic. The circular economy requires innovation and collaboration to raise data-handling complexity. The paper contributes to clarifying key characteristics of big data that need to be addressed at each phase of a circular economy strategy, as well as identifying critical stakeholders affected by the strategies' implementation.

Like big data management, studies about sustainability continue to gain attention in the business and management field (Landrum & Ohsowski, 2017; Wichaisri & Sopadang, 2018). In line with the cluster analysis results, many citations in this article imply that sustainability may become one of the most discussed themes on data veracity. This focus on sustainability issues may bring various developments for data veracity from a sustainability perspective, as well as open possibilities for research novelties in the area.

## Authorship

The development of the study of data veracity can be explored by looking at the contributions of authors in this field of knowledge. Co-authorship analysis using VOSViewer was used to identify the number of authors involved, their publications, and their collaborations. The analysis revealed that a total of 598 authors have contributed to this field to date. However, it was also noted that most authors have only produced one publication in this area, with only 23 authors having published more than two. This suggests that more authors should focus on research related to data veracity in the business and management field.

It was observed that the study is still in its early stages, as the number of publications has been increasing steadily since 2014 (see Figure 3), indicating a growing interest in this topic among academia. This presents more opportunities for authors to explore this topic, conduct in-depth investigations, respond to findings from published articles, and collaborate with other interested authors.

The existing collaboration attempts also indicate that there is a need for more attention to be given to collaborative research among authors. Figure 9 shows that only seven authors with at least two publications are connected within the network. A review of publications by these authors revealed that two proceedings have been published based on their collaboration, both of which address truth discovery using scale reduction (X. Wang et al., 2015, 2016). This finding emphasizes the highly recommended nature of collaboration among researchers interested in this topic, allowing them

to explore more areas of research and cover a wider range of topics related to data veracity in the business and management field.

Despite the lack of collaboration attempts, cross-country collaborative publications, as shown in Figure 10, demonstrate that collaboration among authors from different countries has been increasing. However, most of the partnerships involve just two or three publications. Most of these collaborations have been conducted between developed countries, highlighting the need for more collaboration between developed and developing countries.

Conducting cross-country collaborations may prove to be challenging due to barriers like language, cultural differences, and geopolitics, which may affect the collaboration attempts (Schubert & Schubert, 2020). Nevertheless, cross-country collaborative research would benefit the development of this field, increasing the impact of research outcomes (Turner & Baker, 2020).

## **Gaps and Possible Future Research**

Analysis of this study has identified a substantial number of publications focusing on data validation, particularly in the context of automated fact-checking applications. The significance of data veracity in adding value to data is a recurring theme, including in data processing and analytics. Numerous publications also discuss the application of data veracity across various business sectors, with specific attention given to social media, accounting, and sustainability strategies.

Despite the increasing research on big data and veracity, this study underscores the need for more attention to publications related to data veracity in the business and management domain. This suggests potential opportunities to address research topics that require further exploration and attention in future research. Reflecting on findings from this study, several gaps are identified that could contribute to future research agendas.

### *Data Veracity Definition and Characteristics*

An issue that requires attention is the definition of data veracity, which remains subject to further discussion for consensus. This is underscored by the study by García Lozano et al. (2020). Previous research has associated data veracity with the quality of data that contributes to its value, such as data accuracy, reliability, credibility, and objectivity (García Lozano et al., 2020; Geerts & O’Leary, 2022; Schroeck et al., 2012). Other studies highlighted attributes of veracity data that can negatively affect data quality, such as uncertainty, incompleteness, and inconsistency (Gandomi & Haider, 2015; Himanen et al., 2019; Saha & Srivastava, 2014). The findings of this study support the view that data veracity contributes to improving data quality, as the term “veracity” is commonly used in research to signify a supporting factor for data quality.

The research outlined in this study primarily focused on improving data quality by implementing veracity tools. Despite the increasing research on data veracity, conflicting definitions of the term may create confusion for researchers in future studies. It is, therefore, recommended that a dedicated literature review is conducted to address the definition of veracity and to clarify divergent perspectives on this topic.

In addition to defining data veracity, there is a need for research specifically addressing the characteristics of data veracity itself. Previous research has predominantly viewed data veracity as one aspect of big data management, limiting the focus of data veracity discussions to improving veracity within the context of big data and positioning it solely as a support system for enhancing big data quality. However, critical aspects of data veracity, such as data truthfulness and authenticity, remain insufficiently explored in current studies on the subject.

While truthfulness and honesty are universally considered ethical standards, some businesses may prioritize maximizing monetary benefits over ensuring data accuracy. Evidence indicates that several companies prefer faster, more up-to-date data over accurate information (Cohen, 2017). The increasing trend of businesses relying on external data vendors for faster data gathering and processing introduces risks related to the accuracy and authenticity of the data (Hamlen & Thuraisingham, 2013).

Further discussion is required on companies' perceptions of the trade-off between data velocity and data veracity, as it affects their priorities regarding pursuing data veracity.

### *Human Capabilities on Data Veracity*

The dominant research in Clusters 1, 2, and 4, which focuses on data validation and processing, highlights automated data management tools to enhance data veracity.

The development of AI algorithms can introduce bias and deception risks due to poor quality or poor training (Akter et al., 2022; Nadeem et al., 2022; Xia et al., 2023). Insufficient data not only impacts the process of AI development, but also leads to negative consequences and unethical conduct toward users and society (Siau & Wang, 2020). This underscores the importance of ensuring good data quality for AI to develop algorithms designed to improve performance (Peng & Bhaskar, 2023). Merely investing in automated tools or AI systems is insufficient; additional support, such as enhancing human digital and information literacy capabilities and implementing good corporate governance, is necessary to maintain data quality.

The prevalence of automated verification tools raises concerns about human roles in ensuring data veracity. While automated tools are acknowledged to be more consistent and objective compared to human reviews, the involvement of humans as the ultimate determinant remains essential to support data veracity. The critical factors in big data management, which include technology, also hinge on human capabilities, such as effective leadership, talent management, decision-making, and organizational culture (Brynjolfsson & McAfee, 2012).

The limitations of AI can be analyzed from the perspective of the data, information, knowledge, and wisdom model (Ackoff, 1989; Baškarada & Koronios, 2013; Rowley, 2007). According to this model, the intelligence aspect of data processing can transform data into information and generate knowledge, a task that computers, robots, or automated systems can perform. However, the human wisdom factor remains indispensable for making sound decisions based on knowledge. While computerized data analytics provide alternatives, it is humans who decide which alternative to implement. This implies that, regardless of how advanced the installed veracity system is, the data quality will not thrive if the data users make an inappropriate judgement.

Human wisdom is distinctive in judgement and ethics, shaped by diverse experiences, perceptions, ideologies, or education. This uniqueness is crucial for making valuable decisions, particularly in situations marked by uncertainty (Hariri et al., 2019). Dealing with uncertainties requires cognitive, perceptual, or psychosocial capabilities, which may extend beyond the capacity of computer algorithms (Korherr & Kanbach, 2021; Mondal & Samaddar, 2021; Verganti et al., 2020). In situations where expected values vary (such as mass customization requests), human capabilities for uniqueness may support sound practical decisions. This is a task beyond the scope of automated intelligence systems, which are confined by their designated algorithms (Mondal & Samaddar, 2021; Verganti et al., 2020). Human ethics also can guide humans in assessing potentially malicious information risk, as individuals with strong moral beliefs would not compromise misuse behavior (C. Lin et al., 2020).

Consequently, data processing should not solely focus on automating systems to make the processing more efficient. It should also integrate human capabilities and ethics to make accurate judgements and valuable decisions (Nahavandi, 2019).

Several potential research themes need to be considered to enhance the discussion. Exploring the integration and collaboration between humans and AI in managing data veracity presents an intriguing topic, offering insights into the role each plays in improving big data veracity. The theme of integrating humans and AI has become central to Industry 5.0, emphasizing the need for sustainable collaborations across various business sectors (Demir et al., 2019; Nahavandi, 2019). Examining the collaboration between humans and veracity tools may suggest effective strategies for addressing big data issues. For instance, one study highlighted the need for employees to undergo knowledge transfer and adapt to change when a data analytics system is implemented (Janssen et al., 2017), prompting further investigation into human adaptation to data veracity systems and tools. Another potential

theme involves exploring human skills and capabilities in promoting data veracity. Humans assume various roles in big data management, such as analysts, data scientists, system developers, and system managers (De Mauro et al., 2018). Their ability to verify the credibility, objectivity, and truthfulness of data is crucial for generating valuable data. Businesses can leverage human skills and AI to ensure they work with reliable and trustworthy information. Furthermore, an investigation into how humans can be trained to collaborate with AI systems is needed to ensure maximum effectiveness.

### *The Governance of Data Veracity*

Following the focus on technical areas and tool development, there is a growing need to explore the human roles in ensuring data veracity. Company management practices for handling and maintaining data also needs to be explored.

Organizations regard data as vital for business performance and innovation. The complexity involved in handling data can alter the company's perspective on governing its organization. By following the value chain, a company might prioritize data governance to support its employees, processes, policies, and technologies to leverage its data assets (Abbasi et al., 2016; Khatri & Brown, 2010).

Establishing a function within the organizational structure to manage the flow of data is, therefore, essential. This role, led by a chief data officer, entails the unified organization of data (Porter & Heppelmann, 2015). Considering the large amount of data transferred from various sources, effective data governance requires a support system for managing the veracity of the data. Issues like cybersecurity threats and fake data must be addressed to ensure the effective use of data in a firm's daily operations (Yang et al., 2019).

Approaches to ensure data veracity can be explored by focusing on data governance. Systems supporting data veracity can be integrated into organizational structures, incorporating authority and divisions designed to maintain data veracity (Mullon & Ngoepe, 2019). Alternatively, a project-based strategy involving annual planning and continuous monitoring can be employed to improve data veracity over a designated period (De Haes et al., 2013). Another option is a risk-based approach, acknowledging that data veracity issues may arise from uncertainties and risks (Geerts & O'Leary, 2022). This approach involves identifying barriers and implementing governance mechanisms to prevent or mitigate identified risks (Janssen et al., 2020).

In addition, data governance and veracity issues may prompt a discussion for further study. One concern is how the data organization division manages the flow of data and information both within and outside the firm, especially in the context of open innovation. Achieving this requires collaboration and coordination among the parties concerned. Sharing data could be challenging due to differences in data formats, the existence of unreliable fake data, and the presence of unstructured data. All these factors compromise veracity (Del Vecchio et al., 2018).

Data from external sources may expose the company to legal or copyright issues, affecting the completeness of the retrieved data. Restricted access to external data may impact the accuracy and credibility of the gathered data. Additionally, unstructured massive data, such as online customer reviews or social media data, may contain false information that requires filtering before it is processed for business purposes. Another potential concern is the misalignment between organizational structure and data-flow logic in the internal data management flow, potentially resulting in a lack of control over missing data (Janssen et al., 2020).

### *Social Impact Perspective*

According to Cluster 3, there is abundant research on the role of data veracity in supporting sustainability performance. Still, most publications focus on the environmental perspective, including topics like the transparency of environmental audit reports or strategies to prevent greenwashing practices. This prompts a need for a more extensive exploration of sustainability perspectives. A

notable sustainability issue in the data veracity domain that remains unexplored is related to the social impact of organizations.

Studies related to companies' reports on their social impact and CSR performance revealed veracity issues that require attention. Some studies identified instances of bluewashing, in which firms deceive through their sustainability reports, offering mere lip service to their CSR performances without substantial efforts toward actual environmental or social improvement (Berliner & Prakash, 2015; Macellari et al., 2021). This misleading company disclosure potentially occurs due to the lack of transparency and reliability in the reports as perceived by regulators (Talbot & Boiral, 2015).

Another contributing factor is the company's motives for publishing reports. While sustainability performance should arise from the organization's internal commitment to social responsibility, external pressures like maintaining a positive image with stakeholders may compel the company to engage in CSR activities and, in turn, enhance its report (Besio & Pronzini, 2014). Consequently, the CSR performance and sustainability report might be established as a symbolic gesture to influence stakeholders' perceptions rather than reflecting the firm's actual performance.

There is a need for studies to address the veracity of any sustainability reports to improve the quality of information provided by business entities and minimize deception within these reports. Despite growing interest in addressing environmental veracity, more research should be directed toward evaluating the veracity of social responsibility reports. The lack of strict regulations and control is a significant factor contributing to the inadequate quality of social reports (Talbot & Boiral, 2015). It is, therefore, crucial to further identify appropriate policies for reports, the capabilities of regulators, and the establishment of systems to maintain data quality for assessment.

Another essential area for investigation is the development of methods to detect deception and process sustainability reports. One study proposed a counter-accounting approach as an alternative assessment method, comparing company disclosure with external and diversified information to evaluate the transparency and accuracy of sustainability reports (Macellari et al., 2021). This approach could also be adopted to develop a veracity system for sustainability reports, ensuring internal firm reports are validated using external sources like media, stakeholder testimony, non-governmental organizations, and government agencies.

## **RESEARCH CONTRIBUTIONS AND IMPLICATIONS**

### **Research Contributions**

The bibliometric and literature review identified dominant study areas and research gaps, highlighting potential future discussions on the topic. Given the limited identification of study development regarding data veracity to date, this study offers conceptual clarity and proposes future research directions to address these gaps.

By focusing on bibliometric analysis, this study has elucidated the current state of data veracity research, including dominant keywords, fields discussed, most-cited articles, and key contributing authors. The bibliometric analyses, particularly the co-occurrence analysis, form the basis for a comprehensive literature review of the included articles. Three dominant themes emerged: (1) data validity; (2) big data processing techniques and applications; and (3) data veracity in the context of sustainability. These themes provide a foundation for identifying potential future research directions that remain underexplored.

### **Theoretical Implications**

The authors' findings expand the understanding of data veracity theoretical developments, characteristics, and implementation. While previous literature associated data veracity with terms like data quality, accuracy, and reliability, this exploration includes authenticity, governance, and



ethical perspectives. This broader definition can serve as a foundation for future studies to deepen the understanding of data veracity.

Clusters 2 and 4 revealed that data veracity is primarily connected with one of the key data characteristics of big data, with current research linking data veracity to the big data context. This indicates a need for more research attempts to identify the role of various resources and management aspects that could foster data veracity, expanding the view of it beyond just one big data characteristic.

The study also revealed that many areas demand more investigation, as the number of dominant research tends to polarize into technical perspectives. One research gap highlighted in this study is the role of human judgment and managerial actions in data veracity efforts. While AI and systems can enhance accuracy, human judgment and ethical considerations are essential in the final decision-making process. Integrating human and automated tools is crucial for making sound decisions in complex situations. Future research can build on this framework, exploring the dynamics between human capabilities and automated systems in improving data veracity.

In terms of the managerial perspective, the authors emphasize the importance of data veracity governance, including the role of chief data officers and robust mechanisms to manage data veracity within business environments. This opens discussions on governance models and theories supporting the implementation of data veracity strategies.

These recommendations reflect the dominant research attempts that focus on the technical focus and automated tools that are included in Clusters 1, 2, and 4. Research into areas like human resources and management is essential to catch up with the rapid development of technical areas. Efforts may address the human capability needed to address and support data veracity strategies, integration between humans and systems to implement veracity strategies, or the impact of human judgment and experience in maintaining data quality.

From a managerial focus, more research is needed on the governance of data management, the resources and organizational capabilities required, and key aspects that support veracity, such as integrity, ethics, and company cultures. This area opens more research opportunities in the field of business and management, which has yet to catch up with computer science, where technical and system development remain the main focus of research.

## Practical Implications

From a practical perspective, the study's findings provide a basis for developing strategies related to data management, particularly for improving data quality through veracity efforts. Companies applying data governance should focus on data accuracy and authenticity by investing in data validation systems and enhancing human resource capabilities.

Fostering a data-driven culture and adhering to data management ethics are critical to supporting these strategies. The integration of system development with human capabilities and management contributions is essential for advancing data veracity strategies. This will open more discussions and management actions relevant to the business and management fields, implementing theories and models applicable to this integration to their data management strategies.

Another issue revealed is the consequence of faster data demand. Given the rapid demand for data usage, faster data collection and sharing have become essential aspects of data management. However, this also increases the risks of processing poor-quality data. This study, therefore, urges companies to balance these two aspects. Clear data governance, such as explicit policies and data processing standards supported by proper data validation systems, would help ensure both the velocity and quality of managed data. As human roles become a key research area requiring further identification, companies should also consider enhancing human resource capabilities as a central focus of data veracity strategies. Training and development programs that address digital literacy and emphasize ethical data management practices would improve employees' abilities to implement data veracity strategies, leading to more trustworthy and reliable data management operations.

## CONCLUSION

Data plays a pivotal role in supporting business operations across various sectors, particularly as digitalization and technological advancements reshape industries. Organizations are increasingly treating data as an asset for securing competitive advantages. However, maintaining data quality remains a significant challenge, as data is vulnerable to issues like misinformation, fake data, dark data, and overloaded datasets. Addressing these risks and improving the accuracy and reliability of data management has become a top priority for both practitioners and academics.

The concept of data veracity has emerged as a critical dimension of big data quality, emphasizing credibility, objectivity, and truthfulness in providing valuable information for decision-making processes. This focus on veracity is relevant in today's post-truth era. Misinformation and data manipulation have become widespread, making it increasingly difficult to discern fact from fiction.

The current study presents a comprehensive taxonomy of data veracity theories in the business and management field through a bibliographic analysis and literature review of publications from 2002 to 2023. By examining key works and identifying gaps in the existing research, this review offers valuable insights for future research on data veracity. The findings underscore the importance of developing robust theories and best practices for managing data quality, particularly in a post-truth business environment. As concerns about data credibility intensify, these contributions will be instrumental in helping organizations mitigate risks associated with misinformation and maintain trust with stakeholders.

While the study achieved its primary objectives, two limitations were identified. First, the reliance on a single database, Scopus, may have excluded relevant publications available in other databases. Expanding the scope could provide additional insights. Second, the co-citation and bibliographic coupling analyses revealed a relatively weak connection between clusters. This suggests that research on data veracity is still in its early stages within the business and management field.

Despite these limitations, the findings highlight the need for more research to address data veracity issues, particularly in the context of big data and the growing prevalence of misinformation in the post-truth era. The results of this study can serve as a foundation for future research, which will be crucial in shaping how businesses handle data veracity challenges and thrive in an increasingly complex data-driven world.

## CONFLICTS OF INTEREST

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

## FUNDING STATEMENT

No funding was received for this work.

## PROCESS DATES

Received: May 10, 2024, Revision: September 20, 2024; Accepted: September 20, 2024

## CORRESPONDING AUTHOR

Correspondence should be addressed to Putra Endi Catyanadika; putra.endi.catyanadika@rmit.edu.au

## REFERENCES

- Abbasi, A., Sarker, S., Chiang, R. H. L., & Lindner, C. H. (2016). Big data research in information systems: Toward an inclusive research agenda. *Journal of the Association for Information Systems*, 17(2), I–XXXII. DOI: 10.17705/1jais.00423
- Ackers, B. (2017). The evolution of corporate social responsibility assurance—A longitudinal study. *Social and Environmental Accountability Journal*, 37(2), 97–117. DOI: 10.1080/0969160X.2017.1294097
- Ackoff, R. (1989). From data to wisdom. *Journal of Applied Systems Analysis*, 16(1), 3–9. DOI: 10.5840/du2005155/629
- Akter, S., Dwivedi, Y. K., Sajib, S., Biswas, K., Bandara, R. J., & Michael, K. (2022). Algorithmic bias in machine learning-based marketing models. *Journal of Business Research*, 144, 201–216. DOI: 10.1016/j.jbusres.2022.01.083
- Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics*, 182, 113–131. DOI: 10.1016/j.ijpe.2016.08.018
- Akturan, U. (2018). How does greenwashing affect green branding equity and purchase intention? An empirical research. *Marketing Intelligence & Planning*, 36(7), 809–824. DOI: 10.1108/MIP-12-2017-0339
- Ali, S., Zhou, Y., & Patterson, M. (2022). Efficient analysis of COVID-19 clinical data using machine learning models. *Medical & Biological Engineering & Computing*, 60(7), 1881–1896. DOI: 10.1007/s11517-022-02570-8 PMID: 35507111
- Alshwabkeh, R. O. K., Al-Awamleh, H. K., Alkhawaldeh, M. I. G., Kanaan, R. K., Al-Hawary, S. I. S., Mohammad, A. A. S., & Alkhawaldeh, R. A. (2022). The mediating role of supply chain management on the relationship between big data and supply chain performance using SCOR model. *Uncertain Supply Chain Management*, 10(3), 729–736. DOI: 10.5267/j.uscm.2022.5.002
- Amrollahi, A. (2021). A conceptual tool to eliminate filter bubbles in social networks. *AJIS. Australasian Journal of Information Systems*, 25, 25. DOI: 10.3127/ajis.v25i0.2867
- Apuke, O. D., & Omar, B. (2021). Fake news and COVID-19: Modelling the predictors of fake news sharing among social media users. *Telematics and Informatics*, 56, 101475. Advance online publication. DOI: 10.1016/j.tele.2020.101475 PMID: 34887612
- Ardito, L., Scuotto, V., Del Giudice, M., & Petruzzelli, A. M. (2019). A bibliometric analysis of research on big data analytics for business and management. *Management Decision*, 57(8), 1993–2009. DOI: 10.1108/MD-07-2018-0754
- Assiri, F. (2020). Methods for assessing, predicting, and improving data veracity: A survey. *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal*, 9(4), 5–30. DOI: 10.14201/ADCAIJ202094530
- Azimi, S., Chan, K., & Krasnikov, A. (2022). How fakes make it through: The role of review features versus consumer characteristics. *Journal of Consumer Marketing*, 39(5), 523–537. DOI: 10.1108/JCM-04-2021-4597
- Banerjee, S., Bhattacharyya, S., & Bose, I. (2017). Whose online reviews to trust? Understanding reviewer trustworthiness and its impact on business. *Decision Support Systems*, 96, 17–26. DOI: 10.1016/j.dss.2017.01.006
- Barba-Montoya, J., dos Reis, M., Schneider, H., Donoghue, P. C. J., & Yang, Z. (2018). Constraining uncertainty in the timescale of angiosperm evolution and the veracity of a cretaceous terrestrial revolution. *The New Phytologist*, 218(2), 819–834. DOI: 10.1111/nph.15011 PMID: 29399804
- Baškarada, S., & Koronios, A. (2013). A semiotic theoretical and empirical exploration of the hierarchy and its quality dimension. *AJIS. Australasian Journal of Information Systems*, 18(1). Advance online publication. DOI: 10.3127/ajis.v18i1.748
- Batistič, S., & van der Laken, P. (2019). History, evolution and future of big data and analytics: A bibliometric analysis of its relationship to performance in organizations. *British Journal of Management*, 30(2), 229–251. DOI: 10.1111/1467-8551.12340

- Bayou, M. E., Reinstein, A., & Williams, P. F. (2011). To tell the truth: A discussion of issues concerning truth and ethics in accounting. *Accounting, Organizations and Society*, 36(2), 109–124. DOI: 10.1016/j.aos.2011.02.001
- Bell, D., Lycett, M., Marshan, A., & Monaghan, A. (2021). Exploring future challenges for big data in the humanitarian domain. *Journal of Business Research*, 131, 453–468. DOI: 10.1016/j.jbusres.2020.09.035
- Berliner, D., & Prakash, A. (2015). “Bluewashing” the firm? Voluntary regulations, program design, and member compliance with the United Nations global compact. *Policy Studies Journal: the Journal of the Policy Studies Organization*, 43(1), 115–138. DOI: 10.1111/psj.12085
- Besio, C., & Pronzini, A. (2014). Morality, ethics, and values outside and inside organizations: An example of the discourse on climate change. *Journal of Business Ethics*, 119(3), 287–300. DOI: 10.1007/s10551-013-1641-2
- Björkdahl, J., & Holmén, M. (2019). Exploiting the control revolution by means of digitalization: Value creation, value capture, and downstream movements. *Industrial and Corporate Change*, 28(3), 423–436. DOI: 10.1093/icc/dty022
- Bond, C. F.Jr, & Depaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, 10(3), 214–234. DOI: 10.1207/s15327957pspr1003\_2 PMID: 16859438
- Bond, C. F.Jr, Omar, A., Mahmoud, A., & Bonser, R. N. (1990). Lie detection across cultures. *Journal of Nonverbal Behavior*, 14(3), 189–204. DOI: 10.1007/BF00996226
- Brynjolfsson, E., & McAfee, A. (2012). Big data: The management review. *Harvard Business Review*, 1–12. <https://tarjomefa.com/wp-content/uploads/2017/04/6539-English-TarjomeFa-1.pdf>
- Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory*, 6(3), 203–242. Advance online publication. DOI: 10.1111/j.1468-2885.1996.tb00127.x
- Cappa, F., Oriani, R., Peruffo, E., & McCarthy, I. (2021). Big data for creating and capturing value in the digitalized environment: Unpacking the effects of volume, variety, and veracity on firm performance\*. *Journal of Product Innovation Management*, 38(1), 49–67. DOI: 10.1111/jpim.12545
- Castka, P., Searcy, C., & Mohr, J. (2020). Technology-enhanced auditing: Improving veracity and timeliness in social and environmental audits of supply chains. *Journal of Cleaner Production*, 258, 120773. Advance online publication. DOI: 10.1016/j.jclepro.2020.120773
- Chae, B., Yang, C., Olson, D., & Sheu, C. (2014). The impact of advanced analytics and data accuracy on operational performance: A contingent resource based theory (RBT) perspective. *Decision Support Systems*, 59(1), 119–126. DOI: 10.1016/j.dss.2013.10.012
- Chatterjee, S. (2020). Drivers of helpfulness of online hotel reviews: A sentiment and emotion mining approach. *International Journal of Hospitality Management*, 85, 102356. Advance online publication. DOI: 10.1016/j.ijhm.2019.102356
- Chatterjee, S., Rakshit, K., & Mandal, P. (2023). Exploring the antecedents for hospitality reviewers’ trustworthiness and its impact on business patronage. *International Journal of Hospitality Management*, 110, 103448. Advance online publication. DOI: 10.1016/j.ijhm.2023.103448
- Chen, H., Chiang, R. H. L., Storey, V. C., Lindner, C. H., & Robinson, J. M. (2012). Business intelligence and analytics: From big data to big impact. *Management Information Systems Quarterly*, 36(4), 1165–1188. DOI: 10.2307/41703503
- Cheng, S., Zhang, Q., & Qin, Q. (2016). Big data analytics with swarm intelligence. *Industrial Management & Data Systems*, 116(4), 646–666. DOI: 10.1108/IMDS-06-2015-0222
- Chernyavskiy, A., Ilvovsky, D., & Nakov, P. (2021). WhatTheWikiFact: Fact-checking claims against Wikipedia. *International Conference on Information and Knowledge Management, Proceedings*, 4690–4695. DOI: 10.1145/3459637.3481987
- Chesbrough, H., Lettl, C., & Ritter, T. (2018). Value creation and value capture in open innovation. *Journal of Product Innovation Management*, 35(6), 930–938. DOI: 10.1111/jpim.12471
- Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales—Online book reviews. *JMR, Journal of Marketing Research*, 43(3), 345–354. DOI: 10.1509/jmkr.43.3.345

- Chi, H., & Liao, B. (2022). A quantitative argumentation-based Automated eXplainable Decision System for fake news detection on social media. *Knowledge-Based Systems*, 242, 108378. Advance online publication. DOI: 10.1016/j.knosys.2022.108378
- Cohen, M. (2017). Fake news and manipulated data, the new GDPR, and the future of information. *Business Information Review*, 34(2), 81–85. DOI: 10.1177/0266382117711328
- Connolly, J. M., Uscinski, J. E., Klofstad, C. A., & West, J. P. (2019). Communicating to the public in the era of conspiracy theory. *Public Integrity*, 21(5), 469–476. DOI: 10.1080/10999922.2019.1603045
- Conroy, N. J., Rubin, V. L., & Chen, Y. (2015). Automatic deception detection: Methods for finding fake news. *Proceedings of the Association for Information Science and Technology*, 52(1), 1–4. DOI: 10.1002/ptra2.2015.145052010082
- Constantiou, I. D., & Kallinikos, J. (2015). New games, new rules: Big data and the changing context of strategy. *Journal of Information Technology*, 30(1), 44–57. DOI: 10.1057/jit.2014.17
- 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. Advance online publication. DOI: 10.1016/j.im.2019.01.003
- Dai, H. N., Wang, H., Xu, G., Wan, J., & Imran, M. (2020). Big data analytics for manufacturing internet of things: Opportunities, challenges and enabling technologies. *Enterprise Information Systems*, 14(9–10), 1279–1303. DOI: 10.1080/17517575.2019.1633689
- De Haes, S., Van Grembergen, W., & Debreceny, R. S. (2013). COBIT 5 and enterprise governance of information technology: Building blocks and research opportunities. *Journal of Information Systems*, 27(1), 307–324. DOI: 10.2308/isis-50422
- de Maeyer, P. (2012). Impact of online consumer reviews on sales and price strategies: A review and directions for future research. *Journal of Product and Brand Management*, 21(2), 132–139. DOI: 10.1108/10610421211215599
- De Mauro, A., Greco, M., Grimaldi, M., & Ritala, P. (2018). Human resources for big data professions: A systematic classification of job roles and required skill sets. *Information Processing & Management*, 54(5), 807–817. DOI: 10.1016/j.ipm.2017.05.004
- Del Vecchio, P., Di Minin, A., Petruzzelli, A. M., Panniello, U., & Pirri, S. (2018). Big data for open innovation in SMEs and large corporations: Trends, opportunities, and challenges. *Creativity and Innovation Management*, 27(1), 6–22. DOI: 10.1111/caim.12224
- Demir, K. A., Döven, G., & Sezen, B. (2019). Industry 5.0 and human-robot co-working. *Procedia Computer Science*, 158, 688–695. DOI: 10.1016/j.procs.2019.09.104
- Deng, D., Leung, C. K., Zhao, C., Wen, Y., & Zheng, H. (2021). Spatial-temporal data science of COVID-19 data. *Proceedings - 2021 IEEE 15th International Conference on Big Data Science and Engineering, BigDataSE 2021*, 7–14. DOI: 10.1109/BigDataSE53435.2021.00011
- Despeisse, M., Baumers, M., Brown, P., Charnley, F., Ford, S. J., Garmulewicz, A., Knowles, S., Minshall, T. H. W., Mortara, L., Reed-Tsochas, F. P., & Rowley, J. (2017). Unlocking value for a circular economy through 3D printing: A research agenda. *Technological Forecasting and Social Change*, 115, 75–84. DOI: 10.1016/j.techfore.2016.09.021
- Dubey, R., Gunasekaran, A., Childe, S. J., Blome, C., & Papadopoulos, T. (2019). Big data and predictive analytics and manufacturing performance: Integrating institutional theory, resource-based view and big data culture. *British Journal of Management*, 30(2), 341–361. DOI: 10.1111/1467-8551.12355
- Eachempati, P., Srivastava, P. R., Kumar, A., Tan, K. H., & Gupta, S. (2021). Validating the impact of accounting disclosures on stock market: A deep neural network approach. *Technological Forecasting and Social Change*, 170, 120903. Advance online publication. DOI: 10.1016/j.techfore.2021.120903
- Eembijamil, N. C., Ishak, I., & Sidi, F. (2017). Deception detection approach for data veracity in online digital news: Headlines vs contents. *AIP Conference Proceedings*, 1891. Advance online publication. DOI: 10.1063/1.5005369

- Eismann, K. (2021). Diffusion and persistence of false rumors in social media networks: Implications of searchability on rumor self-correction on Twitter. *Journal of Business Economics*, 91(9), 1299–1329. DOI: 10.1007/s11573-020-01022-9
- Emeseh, E., & Songi, O. (2014). CSR, human rights abuse and sustainability report accountability. *International Journal of Law and Management*, 56(2), 136–151. DOI: 10.1108/IJLMA-01-2013-0001
- Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big data consumer analytics and the transformation of marketing. *Journal of Business Research*, 69(2), 897–904. DOI: 10.1016/j.jbusres.2015.07.001
- Evans, O., Cotton-Barratt, O., Finnveden, L., Bales, A., Balwit, A., Wills, P., Righetti, L., & Saunders, W. (2021). *Truthful AI: Developing and governing AI that does not lie*. <http://arxiv.org/abs/2110.06674>
- Fell, M. J., Pagel, L., & Chen, C. fei, Goldberg, M. H., Herberz, M., Huebner, G. M., Sareen, S., & Hahnel, U. J. J. (2020). Validity of energy social research during and after COVID-19: Challenges, considerations, and responses. In *Energy Research and Social Science* (volume 68). Elsevier Ltd. DOI: 10.1016/j.erss.2020.101646
- Figueira, Á., & Oliveira, L. (2017). The current state of fake news: Challenges and opportunities. *Procedia Computer Science*, 121, 817–825. DOI: 10.1016/j.procs.2017.11.106
- Fuller, C. M., Biros, D. P., & Wilson, R. L. (2009). Decision support for determining veracity via linguistic-based cues. *Decision Support Systems*, 46(3), 695–703. DOI: 10.1016/j.dss.2008.11.001
- Gal, G., & Akisik, O. (2020). The impact of internal control, external assurance, and integrated reports on market value. *Corporate Social Responsibility and Environmental Management*, 27(3), 1227–1240. DOI: 10.1002/csr.1878
- Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137–144. DOI: 10.1016/j.ijinfomgt.2014.10.007
- Gao, J., Xie, C., & Tao, C. (2016). Big data validation and quality assurance - Issues, challenges, and needs. *Proceedings - 2016 IEEE Symposium on Service-Oriented System Engineering, SOSE 2016*, 433–441. DOI: 10.1109/SOSE.2016.63
- García, S. A., García, G. G., Prieto, M. S., Guerrero, A. J. M., & Jiménez, C. R. (2020). The impact of term fake news on the scientific community. Scientific performance and mapping in Web of science. *Social Sciences (Basel, Switzerland)*, 9(5), 73. Advance online publication. DOI: 10.3390/socsci9050073
- García Lozano, M., Brynielsson, J., Franke, U., Rosell, M., Tjörnhammar, E., Varga, S., & Vlassov, V. (2020). Veracity assessment of online data. *Decision Support Systems*, 129, 113132. DOI: 10.1016/j.dss.2019.113132
- Gärtner, B., & Hiebl, M. R. W. (2018). *Issues with Big Data*.
- Geerts, G. L., & O'Leary, D. E. (2022). V-Matrix: A wave theory of value creation for big data. *International Journal of Accounting Information Systems*, 47, 100575. Advance online publication. DOI: 10.1016/j.accinf.2022.100575
- Gelman, I. A. (2010). Setting priorities for data accuracy improvements in satisficing decision-making scenarios: A guiding theory. *Decision Support Systems*, 48(4), 507–520. DOI: 10.1016/j.dss.2009.11.001
- Genevès, P., Calmant, T., Layaïda, N., Lepelley, M., Artemova, S., & Bosson, J. L. (2018). Scalable machine learning for predicting at-risk profiles upon hospital admission. *Big Data Research*, 12, 23–34. DOI: 10.1016/j.bdr.2018.02.004
- George, J., Gerhart, N., & Torres, R. (2021). Uncovering the truth about fake news: A research model grounded in multi-disciplinary literature. *Journal of Management Information Systems*, 38(4), 1067–1094. DOI: 10.1080/07421222.2021.1990608
- Ghasemaghaei, M. (2021). Understanding the impact of big data on firm performance: The necessity of conceptually differentiating among big data characteristics. *International Journal of Information Management*, 57, 102055. Advance online publication. DOI: 10.1016/j.ijinfomgt.2019.102055
- Ghasemaghaei, M., & Calic, G. (2019). Can big data improve firm decision quality? The role of data quality and data diagnosticity. *Decision Support Systems*, 120, 38–49. DOI: 10.1016/j.dss.2019.03.008

- Gupta, S., Kar, A. K., Baabdullah, A., & Al-Khowaiter, W. A. A. (2018a). Big data with cognitive computing: A review for the future. *International Journal of Information Management*, 42, 78–89. DOI: 10.1016/j.ijinfomgt.2018.06.005
- Gupta, S., Kar, A. K., Baabdullah, A., & Al-Khowaiter, W. A. A. (2018b). Big data with cognitive computing: A review for the future. *International Journal of Information Management*, 42, 78–89. DOI: 10.1016/j.ijinfomgt.2018.06.005
- Hamlen, K. W., & Thuraisingham, B. (2013). Data security services, solutions and standards for outsourcing. *Computer Standards & Interfaces*, 35(1), 1–5. DOI: 10.1016/j.csi.2012.02.001
- Harding, N., & Trotman, K. T. (2017). The effect of partner communications of fraud likelihood and skeptical orientation on auditors' professional skepticism. *Auditing*, 36(2), 111–131. DOI: 10.2308/ajpt-51576
- Hariri, R. H., Fredericks, E. M., & Bowers, K. M. (2019). Uncertainty in big data analytics: Survey, opportunities, and challenges. *Journal of Big Data*, 6(1), 44. Advance online publication. DOI: 10.1186/s40537-019-0206-3
- Hassan, N., Li, C., & Tremayne, M. (2015). Detecting check-worthy factual claims in presidential debates. *International Conference on Information and Knowledge Management, Proceedings* (pp. 1835–1838). <https://doi.org/10.1145/2806416.2806652>
- Hazen, B. T., Boone, C. A., Ezell, J. D., & Jones-Farmer, L. A. (2014). Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications. *International Journal of Production Economics*, 154, 72–80. DOI: 10.1016/j.ijpe.2014.04.018
- Herschel, R., & Miori, V. M. (2017). Ethics & big data. *Technology in Society*, 49, 31–36. DOI: 10.1016/j.techsoc.2017.03.003
- Himanen, L., Geurts, A., Foster, A. S., & Rinke, P. (2019). Data-driven materials science: Status, challenges, and perspectives. *Advancement of Science*, 6(21), 1900808. Advance online publication. DOI: 10.1002/adv.201900808 PMID: 31728276
- Jabbour, C. J. C., Jabbour, A. B. L. de S., Sarkis, J., & Filho, M. G. (2019). Unlocking the circular economy through new business models based on large-scale data: An integrative framework and research agenda. *Technological Forecasting and Social Change*, 144, 546–552. DOI: 10.1016/j.techfore.2017.09.010
- Jamil, N. B. C. E., Ishak, I., Sidi, F., Affendey, L. S., & Mamat, A. (2015). A systematic review on the profiling of digital news portal for big data veracity. *Procedia Computer Science*, 72, 390–397. DOI: 10.1016/j.procs.2015.12.154
- Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., & Janowski, T. (2020). Data governance: Organizing data for trustworthy artificial intelligence. *Government Information Quarterly*, 37(3), 101493. Advance online publication. DOI: 10.1016/j.giq.2020.101493
- Janssen, M., van der Voort, H., & Wahyudi, A. (2017). Factors influencing big data decision-making quality. *Journal of Business Research*, 70, 338–345. DOI: 10.1016/j.jbusres.2016.08.007
- Jensen, M. L., Lowry, P. B., Burgoon, J. K., & Nunamaker, J. F. (2010). Technology dominance in complex decision making: The case of aided credibility assessment. *Journal of Management Information Systems*, 27(1), 175–202. DOI: 10.2753/MIS0742-1222270108
- Johar, G. V. (2022). Untangling the web of misinformation and false beliefs. *Journal of Consumer Psychology*, 32(2), 374–383. DOI: 10.1002/jcpsy.1288
- Katoch, R. (2022). IoT research in supply chain management and logistics: A bibliometric analysis using vosviewer software. *Materials Today: Proceedings*, 56, 2505–2515. DOI: 10.1016/j.matpr.2021.08.272
- Kaul, M., Storey, V. C., & Woo, C. (2017). A framework for managing complexity in information systems. *Journal of Database Management*, 28(1), 31–42. DOI: 10.4018/JDM.2017010103
- Keenan, P. B., & Jankowski, P. (2019). Spatial decision support systems: Three decades on. *Decision Support Systems*, 116, 64–76. DOI: 10.1016/j.dss.2018.10.010
- Khaldarova, I., & Pantti, M. (2016). Fake News: The narrative battle over the Ukrainian conflict. *Journalism Practice*, 10(7), 891–901. DOI: 10.1080/17512786.2016.1163237

- Khatri, V., & Brown, C. V. (2010). Designing data governance. *Communications of the ACM*, 53(1), 148–152. DOI: 10.1145/1629175.1629210
- Kimble, C., & Milolidakis, G. (2015). Big data and business intelligence: Debunking the myths. *Global Business and Organizational Excellence*, 35(1), 23–34. DOI: 10.1002/joe.21642
- Kitchin, R., & McArdle, G. (2016). What makes big data, big data? Exploring the ontological characteristics of 26 datasets. *Big Data & Society*, 3(1), 2053951716631130. Advance online publication. DOI: 10.1177/2053951716631130
- Kollmer, T., Eckhardt, A., & Reibenspiess, V. (2022). Explaining consumer suspicion: Insights of a vignette study on online product reviews. *Electronic Markets*, 32(3), 1221–1238. DOI: 10.1007/s12525-022-00549-9 PMID: 35692740
- Korherr, P., & Kanbach, D. (2021). Human-related capabilities in big data analytics: A taxonomy of human factors with impact on firm performance. *Review of Managerial Science*. Advance online publication. DOI: 10.1007/s11846-021-00506-4
- Kuruppu, S., & Milne, M. J. (2010). Dolphin deaths, organizational legitimacy and potential employees' reactions to assured environmental disclosures. *Accounting Forum*, 34(1), 1–19. DOI: 10.1016/j.accfor.2009.12.001
- Kushwaha, A. K., Kar, A. K., & Dwivedi, Y. K. (2021). Applications of big data in emerging management disciplines: A literature review using text mining. *International Journal of Information Management Data Insights*, 1(2), 100017. Advance online publication. DOI: 10.1016/j.ijime.2021.100017
- 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(3), 387–394. DOI: 10.1016/j.ijinfomgt.2014.02.002
- La Torre, M., Sabelfeld, S., Blomkvist, M., & Dumay, J. (2020). Rebuilding trust: Sustainability and non-financial reporting and the European Union regulation. *Meditari Accountancy Research*, 28(5), 701–725. DOI: 10.1108/MEDAR-06-2020-0914
- Lam, S. K., Sleep, S., Hennig-Thurau, T., Sridhar, S., & Saboo, A. R. (2017). Leveraging frontline employees' small data and firm-level big data in frontline management: An absorptive capacity perspective. *Journal of Service Research*, 20(1), 12–28. DOI: 10.1177/1094670516679271
- Landrum, N. E., & Ohsowski, B. (2017). Content trends in sustainable business education: An analysis of introductory courses in the USA. *International Journal of Sustainability in Higher Education*, 18(3), 385–414. DOI: 10.1108/IJSHE-07-2016-0135
- Laney, D. (2001). *3D data management: Controlling data volume, velocity and variety*.
- LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big data, analytics and the path from insights to value. *MIT Sloan Management Review*, 20–32.
- Lee, C. C., Welker, R. B., & Wang, T. W. (2013). An experimental investigation of professional skepticism in audit interviews. *International Journal of Auditing*, 17(2), 213–226. DOI: 10.1111/ijau.12001
- Lee, Y. C., Shen, B. C., & Sia, C. L. (2022). Bringing credibility through portals: Examining the multiple information sources on information quality and credibility. *Journal of Database Management*, 33(1), 1–26. Advance online publication. DOI: 10.4018/JDM.313968
- Leung, C. K., Chen, Y., Hoi, C. S. H., Shang, S., & Cuzzocrea, A. (2020). Machine learning and OLAP on big COVID-19 data. *Proceedings - 2020 IEEE International Conference on Big Data. Big Data, 2020*, 5118–5127. DOI: 10.1109/BigData50022.2020.9378407
- Li, M., Zhao, L., & Srinivas, S. (2023). It is about inclusion! Mining online reviews to understand the needs of adaptive clothing customers. *International Journal of Consumer Studies*, 47(3), 1157–1172. Advance online publication. DOI: 10.1111/ijcs.12895
- Lin, C., Kunnathur, A. S., & Li, L. (2020). The cultural foundation of information security behavior: Developing a cultural fit framework for information security behavior control. *Journal of Database Management*, 31(2), 21–41. DOI: 10.4018/JDM.2020040102



- Lin, R., Xie, Z., Hao, Y., & Wang, J. (2020). Improving high-tech enterprise innovation in big data environment: A combinative view of internal and external governance. *International Journal of Information Management*, 50, 575–585. DOI: 10.1016/j.ijinfomgt.2018.11.009
- Liu, X., Li, Q., Nourbakhsh, A., Fang, R., Thomas, M., Anderson, K., Kociuba, R., Vedder, M., Pomerville, S., Wudali, R., Martin, R., Duprey, J., Vachher, A., Keenan, W., & Shah, S. (2016). Reuters tracer: A large scale system of detecting & verifying real-time news events from twitter. *International Conference on Information and Knowledge Management, Proceedings* (pp. 207–216). DOI: 10.1145/2983323.2983363
- Lloyd, C. D., & Lilley, K. D. (2009). Cartographic veracity in medieval mapping: Analyzing geographical variation in the gough map of Great Britain. *Annals of the Association of American Geographers*, 99(1), 27–48. DOI: 10.1080/00045600802224638
- Luca, M., & Zervas, G. (2016). Fake it till you make it: Reputation, competition, and yelp review fraud. *Management Science*, 62(12), 3412–3427. DOI: 10.1287/mnsc.2015.2304
- Lugmayr, A., Stockleben, B., Scheib, C., & Mailaparampil, M. A. (2017). Cognitive big data: Survey and review on big data research and its implications. What is really “new” in big data? *Journal of Knowledge Management*, 21(1), 197–212. DOI: 10.1108/JKM-07-2016-0307
- Lukasik, M., Bontcheva, K., Cohn, T., Zubiaga, A., Liakata, M., & Procter, R. (2019). Gaussian processes for rumour stance classification in social media. *ACM Transactions on Information Systems*, 37(2), 1–24. Advance online publication. DOI: 10.1145/3295823
- Lyu, T., Eidson, A., Jun, J., Zhou, X., Cui, X., & Liang, C. (2020). Data veracity of patients and health consumers reported adverse drug reactions on Twitter: Key linguistic features, Twitter variables, and association rules. *MedRxiv*. DOI: 10.1101/2020.11.03.20225532
- Macellari, M., Yuriev, A., Testa, F., & Boiral, O. (2021). Exploring bluewashing practices of alleged sustainability leaders through a counter-accounting analysis. *Environmental Impact Assessment Review*, 86, 106489. Advance online publication. DOI: 10.1016/j.eiar.2020.106489
- Manoharan, R. (2019). Big data and health care system using mLearning. *International Journal of Recent Technology and Engineering*. <https://www.researchgate.net/publication/340807666>
- Maroun, W. (2018). Modifying assurance practices to meet the needs of integrated reporting: The case for “interpretive assurance.”. *Accounting, Auditing & Accountability Journal*, 31(2), 400–427. DOI: 10.1108/AAAJ-10-2016-2732
- Martín, A., Huertas-Tato, J., Huertas-García, Á., Villar-Rodríguez, G., & Camacho, D. (2022). FacTeR-Check: Semi-automated fact-checking through semantic similarity and natural language inference. *Knowledge-Based Systems*, 251, 109265. Advance online publication. DOI: 10.1016/j.knsys.2022.109265
- Martin-Fuentes, E., Mateu, C., & Fernandez, C. (2018). Does verifying uses influence rankings? Analyzing booking.com and tripadvisor. *Tourism Analysis*, 23(1), 1–15. DOI: 10.3727/108354218X15143857349459
- Mateo-Márquez, A. J., González-González, J. M., & Zamora-Ramírez, C. (2022). An international empirical study of greenwashing and voluntary carbon disclosure. *Journal of Cleaner Production*, 363, 132567. Advance online publication. DOI: 10.1016/j.jclepro.2022.132567
- Mayzlin, D., Dover, Y., & Chevalier, J. A. (2012). Promotional review: Empirical investigation of online review manipulation. *The American Economic Review*, 104(8), 2421–2455. <https://www.nber.org/papers/w18340>. DOI: 10.1257/aer.104.8.2421
- Meel, P., & Vishwakarma, D. K. (2020). Fake news, rumor, information pollution in social media and web: A contemporary survey of state-of-the-arts, challenges and opportunities. *Expert Systems with Applications*, 153, 112986. DOI: 10.1016/j.eswa.2019.112986
- Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics and firm performance: Findings from a mixed-method approach. *Journal of Business Research*, 98, 261–276. DOI: 10.1016/j.jbusres.2019.01.044
- Mikalef, P., Krogstie, J., Pappas, I. O., & Pavlou, P. (2020). Exploring the relationship between big data analytics capability and competitive performance: The mediating roles of dynamic and operational capabilities. *Information & Management*, 57(2), 103169. Advance online publication. DOI: 10.1016/j.im.2019.05.004

- Mirbabaie, M., Amojó, I., & Stieglitz, S. (2021). Affording twitter in emergency situations: The occurrence of rumor sense-making. *Journal of Database Management*, 32(2), 50–66. DOI: 10.4018/JDM.2021040104
- Mithas, S., Lee, M. R., Earley, S., Associates, E. & Murugesan, S., & Djavanshir, R. (2013). *Leveraging Big Data and Business Analytics*. www.infochimps.com/press/
- Moke, K. C., Low, T. J., & Khan, D. (2021). IoT blockchain data veracity with data loss tolerance. *Applied Sciences (Basel, Switzerland)*, 11(21), 9978. Advance online publication. DOI: 10.3390/app11219978
- Mondal, S., & Samaddar, K. (2021). Reinforcing the significance of human factor in achieving quality performance in data-driven supply chain management. *The TQM Journal*. Advance online publication. DOI: 10.1108/TQM-12-2020-0303
- Mønsted, T. (2019). Achieving veracity: A study of the development and use of an information system for data analysis in preventive healthcare. *Health Informatics Journal*, 25(3), 491–499. DOI: 10.1177/1460458218796665 PMID: 30198372
- Müller, O., Fay, M., & vom Brocke, J. (2018). The effect of big data and analytics on firm performance: An econometric analysis considering industry characteristics. *Journal of Management Information Systems*, 35(2), 488–509. DOI: 10.1080/07421222.2018.1451955
- Mullon, P. A., & Ngoepe, M. (2019). An integrated framework to elevate information governance to a national level in South Africa. *Records Management Journal*, 29(1–2), 103–116. DOI: 10.1108/RMJ-09-2018-0030
- Muninger, M. I., Hammedi, W., & Mahr, D. (2019). The value of social media for innovation: A capability perspective. *Journal of Business Research*, 95, 116–127. DOI: 10.1016/j.jbusres.2018.10.012
- Myers, K. R., Tham, W. Y., Yin, Y., Cohodes, N., Thursby, J. G., Thursby, M. C., Schiffer, P., Walsh, J. T., Lakhani, K. R., & Wang, D. (2020). Unequal effects of the COVID-19 pandemic on scientists. *Nature Human Behaviour*, 4(9), 880–883. DOI: 10.1038/s41562-020-0921-y PMID: 32669671
- Nadeem, A., Marjanovic, O., & Abedin, B. (2022). Gender bias in AI-based decision-making systems: A systematic literature review. *AJIS. Australasian Journal of Information Systems*, 26, 1–34. DOI: 10.3127/ajis.v26i0.3835
- Nahavandi, S. (2019). Industry 5.0 - A human-centric solution. *Sustainability (Basel)*, 11(16), 4371. Advance online publication. DOI: 10.3390/su11164371
- Nakov, P., Corney, D., Hasanain, M., Alam, F., Elsayed, T., Barrón-Cedeño, A., Papotti, P., Shaar, S., & Martino, G. D. S. (2021). Automated fact-checking for assisting human fact-checkers. *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence* (pp. 4551–4558). <http://arxiv.org/abs/2103.07769>
- Nakov, P., & Da San Martino, G. (2021). Fake news, disinformation, propaganda, and media bias. *International Conference on Information and Knowledge Management, Proceedings* (pp. 4862–4865). DOI: 10.1145/3459637.3482026
- Nguyen, V. H., Sugiyama, K., Nakov, P., & Kan, M. Y. (2020). FANG: Leveraging social context for fake news detection using graph representation. *International Conference on Information and Knowledge Management, Proceedings* (pp. 1165–1174). DOI: 10.1145/3340531.3412046
- Ngwakwe, C. C., & Netswera, F. G. (2013). On veracity in corporate sustainability claims: Why society should be cautious. *Corporate Ownership & Control*, 11(1).
- Pandey, S., Singh, R. K., Gunasekaran, A., & Kaushik, A. (2020). Cyber security risks in globalized supply chains: Conceptual framework. *Journal of Global Operations and Strategic Sourcing*, 13(1), 103–128. DOI: 10.1108/JGOSS-05-2019-0042
- Parviainen, P., Tihinen, M., Kääriäinen, J., & Teppola, S. (2017). Tackling the digitalization challenge: How to benefit from digitalization in practice. *International Journal of Information Systems and Project Management*, 5(1), 63–77. DOI: 10.12821/ijispm050104
- Pauna, V. H., Picone, F., Le Guyader, G., Buonocore, E., & Franzese, P. P. (2018). The scientific research on ecosystem services: A bibliometric analysis. *Ecological Questions*, 29(3), 53–62. DOI: 10.12775/EQ.2018.022

- Peng, G., & Bhaskar, R. (2023). Artificial intelligence and machine learning for job automation. *Journal of Database Management*, 34(1), 1–12. DOI: 10.4018/JDM.318455
- Porter, M. E., & Heppelmann, J. E. (2015). *How smart connected products are transforming companies*.
- Ramachandramurthy, S., Subramaniam, S., & Ramasamy, C. (2015). Distilling big data: Refining quality information in the era of yottabytes. *TheScientificWorldJournal*, 2015(1), 453597. Advance online publication. DOI: 10.1155/2015/453597 PMID: 26495424
- Reimer, A. P., & Madigan, E. A. (2019). Veracity in big data: How good is good enough. *Health Informatics Journal*, 25(4), 1290–1298. DOI: 10.1177/1460458217744369 PMID: 29388495
- Revathy, R., & Canessane, A. (2018). IoT based decision making system to improve veracity of big data. *IACSIT International Journal of Engineering and Technology*, 7(1). www.sciencepubco.com/index.php/IJET
- Rialti, R., Marzi, G., Ciappei, C., & Busso, D. (2019). Big data and dynamic capabilities: A bibliometric analysis and systematic literature review. *Management Decision*, 57(8), 2052–2068. DOI: 10.1108/MD-07-2018-0821
- Riasetiawan, M., Anggara, F., Ashari, A., Winardi, S., & Prastowo, B. N. (2021). Data model and analysis for big data mapping and management in the energy data platform. *2021 International Conference on Data Science, Artificial Intelligence, and Business Analytics, DATABIA 2021 – Proceedings* (pp. 1–6). DOI: 10.1109/DATABIA53375.2021.9650117
- Richey, R. G. Jr, Morgan, T. R., Lindsey-Hall, K., & Adams, F. G. (2016). A global exploration of big data in the supply chain: Global exploration of big data. *International Journal of Physical Distribution & Logistics Management*, 46(8), 710–739. DOI: 10.1108/IJPDLM-05-2016-0134
- Ristevski, B., & Chen, M. (2018). Big data analytics in medicine and healthcare. *Journal of Integrative Bioinformatics*, 15(3), 20170030. Advance online publication. DOI: 10.1515/jib-2017-0030 PMID: 29746254
- Rocha, Y. M., de Moura, G. A., Desidério, G. A., de Oliveira, C. H., Lourenço, F. D., & de Figueiredo Nicolete, L. D. (2021). The impact of fake news on social media and its influence on health during the COVID-19 pandemic: A systematic review. *Journal of Public Health (Germany)*. <https://doi.org/10.1007/s10389-021-01658-z>
- Rowley, J. (2007). The wisdom hierarchy: Representations of the DIKW hierarchy. *Journal of Information Science*, 33(2), 163–180. DOI: 10.1177/0165551506070706
- Rubin, V. L. (2022). Philosophies of truth. In *Misinformation and disinformation* (pp. 95–120). Springer International Publishing., DOI: 10.1007/978-3-030-95656-1\_4
- Rubin, V. L., & Lukoianova, T. (2014). Veracity roadmap: Is big data objective, truthful and credible? *Advances in Classification Research Online*, 24(1), 4–15. DOI: 10.7152/acro.v24i1.14671
- Saha, B., & Srivastava, D. (2014). Data quality: The other face of big data. *Proceedings - International Conference on Data Engineering* (pp. 1294–1297). DOI: 10.1109/ICDE.2014.6816764
- Salehan, M., & Kim, D. J. (2016). Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics. *Decision Support Systems*, 81, 30–40. DOI: 10.1016/j.dss.2015.10.006
- Sampson, J., Morstatter, F., Wu, L., & Liu, H. (2016). Leveraging the implicit structure within social media for emergent rumor detection. *International Conference on Information and Knowledge Management, Proceedings* (pp. 2377–2382). DOI: 10.1145/2983323.2983697
- Sanney, K. J., Trautman, L. J., Yordy, E. D., Cowart, T. W., & Sewell, D. J. L. (2020). The importance of truth telling and trust. *Journal of Legal Studies Education*, 37(1), 7–36. DOI: 10.1111/jlse.12099
- Sawadogo, P., & Darmont, J. (2021). On data lake architectures and metadata management. *Journal of Intelligent Information Systems*, 56(1), 97–120. DOI: 10.1007/s10844-020-00608-7
- Schroeck, M., Shockley, R., Smart, J., Romero- Morales, D., & Tufano, P. (2012). *Analytics: The real-world use of big data. How innovative enterprises extract value from uncertain data*.
- Schubert, A., & Schubert, G. (2020). Internationality at university level. *Scientometrics*, 123(3), 1341–1364. DOI: 10.1007/s11192-020-03443-3

- Segers, S. (2022). Robot technology for the elderly and the value of veracity: Disruptive technology or reinvigorating entrenched principles? *Science and Engineering Ethics*, 28(6), 64. Advance online publication. DOI: 10.1007/s11948-022-00420-2 PMID: 36469167
- Sestino, A., Prete, M. I., Piper, L., & Guido, G. (2020). Internet of things and big data as enablers for business digitalization strategies. *Technovation*, 98, 102173. Advance online publication. DOI: 10.1016/j.technovation.2020.102173
- Shaar, S., Martino, G. D. S., Babulkov, N., & Nakov, P. (2020). That is a known lie: Detecting previously fact-checked claims. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 3607–3618). <http://arxiv.org/abs/2005.06058> DOI: 10.18653/v1/2020.acl-main.332
- Shaar, S., Nikolov, A., Babulkov, N., Alam, F., Barrón-Cedeño, A., Hasanain, M., Suwaileh, R., Haouari, F., Da, G., Martino, S., & Nakov, P. (2020). Overview of CheckThat! 2020 English: Automatic identification and verification of claims in social media. *International Conference of the CLEF Association*. <https://sites.google.com/view/clef2020-checkthat/>
- Shafieizadeh, K., & Tao, C. W. (2020). How does a menu's information about local food affect restaurant selection? The roles of corporate social responsibility, transparency, and trust. *Journal of Hospitality and Tourism Management*, 43, 232–240. DOI: 10.1016/j.jhtm.2020.04.007
- Shams, S. M. R., & Solima, L. (2019). Big data management: Implications of dynamic capabilities and data incubator. *Management Decision*, 57(8), 2113–2123. DOI: 10.1108/MD-07-2018-0846
- Shi, B., & Weninger, T. (2016). Discriminative predicate path mining for fact checking in knowledge graphs. *Knowledge-Based Systems*, 104, 123–133. DOI: 10.1016/j.knosys.2016.04.015
- Shrivastava, M., & Kumar, S. (2021). A pragmatic and intelligent model for sarcasm detection in social media text. *Technology in Society*, 64, 101489. Advance online publication. DOI: 10.1016/j.techsoc.2020.101489
- Shukla, A. K., Yadav, M., Kumar, S., & Muhuri, P. K. (2020). Veracity handling and instance reduction in big data using interval type-2 fuzzy sets. *Engineering Applications of Artificial Intelligence*, 88, 103315. Advance online publication. DOI: 10.1016/j.engappai.2019.103315
- Siau, K., & Wang, W. (2020). Artificial intelligence (AI) ethics: Ethics of AI and ethical AI. *Journal of Database Management*, 31(2), 74–87. DOI: 10.4018/JDM.2020040105
- Sivarajah, U., Irani, Z., Gupta, S., & Mahroof, K. (2020). Role of big data and social media analytics for business to business sustainability: A participatory web context. *Industrial Marketing Management*, 86, 163–179. DOI: 10.1016/j.indmarman.2019.04.005
- Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of big data challenges and analytical methods. *Journal of Business Research*, 70, 263–286. DOI: 10.1016/j.jbusres.2016.08.001
- Statista. (2023, February 6). *Average spending per employee on storage worldwide from 2016 to 2028(in U.S. dollars)*. Statista Research Department.
- Stone, J. D., Ulman, H. K., Tran, K., Thompson, A. G., Halter, M. D., Ramadan, J. H., Stephenson, M., Finomore, V. S.Jr, Galster, S. M., Rezai, A. R., & Hagen, J. A. (2021). Assessing the accuracy of popular commercial technologies that measure resting heart rate and heart rate variability. *Frontiers in Sports and Active Living*, 3, 585870. Advance online publication. DOI: 10.3389/fspor.2021.585870 PMID: 33733234
- Suiter, J. (2016). Post-truth Politics. *Political Insight*, 7(3), 25–27. DOI: 10.1177/2041905816680417
- Suthanthiradevi, P., & Karthika, S. (2019). A novel design framework for rumour analysis in Twitter. *International Journal of Recent Technology and Engineering*.
- Talbot, D., & Boiral, O. (2015). Strategies for climate change and impression management: A case study among Canada's large industrial emitters. *Journal of Business Ethics*, 132(2), 329–346. DOI: 10.1007/s10551-014-2322-5
- Tallon, P. P. (2013). *Corporate governance of big data: Perspectives on value, risk, and cost*.
- Tang, J., Ma, T., & Luo, Q. (2020). Trends prediction of big data: A case study based on fusion data. *Procedia Computer Science*, 174, 181–190. DOI: 10.1016/j.procs.2020.06.073

- Tao, H., Bhuiyan, M. Z. A., Rahman, M. A., Wang, G., Wang, T., Ahmed, M. M., & Li, J. (2019). Economic perspective analysis of protecting big data security and privacy. *Future Generation Computer Systems*, 98, 660–671. DOI: 10.1016/j.future.2019.03.042
- Trabucchi, D., & Buganza, T. (2019). Data-driven innovation: Switching the perspective on Big Data. *European Journal of Innovation Management*, 22(1), 23–40. DOI: 10.1108/EJIM-01-2018-0017
- Troisi, O., D'Arco, M., Loia, F., & Maione, G. (2018). Big data management: The case of Mulino Bianco's engagement platform for value co-creation. *International Journal of Engineering Business Management*, 10, 1847979018767776. Advance online publication. DOI: 10.1177/1847979018767776
- Trotman, K. T., & Wright, W. F. (2012). Triangulation of audit evidence in fraud risk assessments. *Accounting, Organizations and Society*, 37(1), 41–53. DOI: 10.1016/j.aos.2011.11.003
- Turner, J. R., & Baker, R. (2020). Collaborative research: Techniques for conducting collaborative research from the Science of Team Science (SciTS). *Advances in Developing Human Resources*, 22(1), 72–86. DOI: 10.1177/1523422319886300
- Tversky, A., & Kahneman, D. (2005). Judgement under uncertainty: Heuristics and biases. *Knowledge: Critical Concepts*, 371.
- Twyman, N., Elkins, A., Burgoon, J., & Nunamaker, J. (2014). A rigidity detection system for automated credibility assessment. *Journal of Management Information Systems*, 31(1), 173–202. DOI: 10.2753/MIS0742-1222310108
- van Eck, N. J., & Waltman, L. (2014). Visualizing bibliometric networks. In *Measuring scholarly impact* (pp. 285–320). Springer International Publishing., DOI: 10.1007/978-3-319-10377-8\_13
- Verganti, R., Vendraminelli, L., & Iansiti, M. (2020). Innovation and design in the age of artificial intelligence. *Journal of Product Innovation Management*, 37(3), 212–227. DOI: 10.1111/jpim.12523
- Viet, N. Q., Behdani, B., Bloemhof, J., & Hoberg, K. (2021). Value of data in multi-level supply chain decisions: A case study in the Dutch floriculture sector. *International Journal of Production Research*, 59(5), 1368–1385. DOI: 10.1080/00207543.2020.1821116
- Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, 70, 356–365. DOI: 10.1016/j.jbusres.2016.08.009
- Wang, Y. R., Ma, C. Q., & Ren, Y. S. (2022). A model for CBDC audits based on blockchain technology: Learning from the DCEP. *Research in International Business and Finance*, 63, 101781. Advance online publication. DOI: 10.1016/j.ribaf.2022.101781
- Wanner, J., & Janiesch, C. (2019). Big data analytics in sustainability reports: An analysis based on the perceived credibility of corporate published information. *Business Research*, 12(1), 143–173. DOI: 10.1007/s40685-019-0088-4
- Webb, H. W., & Webb, L. A. (2004). SiteQual: An integrated measure of Web site quality. *Journal of Enterprise Information Management*, 17(6), 430–440. DOI: 10.1108/17410390410566724
- Wichaisri, S., & Sopadang, A. (2018). Trends and future directions in sustainable development. *Sustainable Development (Bradford)*, 26(1), 1–17. DOI: 10.1002/sd.1687
- Wu, Y., Ngai, E. W. T., Wu, P., & Wu, C. (2022). Fake news on the internet: A literature review, synthesis and directions for future research. *Internet Research*, 32(5), 1662–1699. DOI: 10.1108/INTR-05-2021-0294
- Xia, H., & An, W. Zuopeng (Justin) Zhang. (2023). Credit risk models for financial fraud detection. *Journal of Database Management*, 34(1), 1–20. DOI: 10.4018/JDM.321739
- Xiao, Y., & Watson, M. (2019). Guidance on conducting a systematic literature review. *Journal of Planning Education and Research*, 39(1), 93–112. DOI: 10.1177/0739456X17723971
- Yang, L., Li, J., Elisa, N., Prickett, T., & Chao, F. (2019). Towards big data governance in cybersecurity. *Data-Enabled Discovery and Applications*, 3(1), 10. Advance online publication. DOI: 10.1007/s41688-019-0034-9

Yu, E. P., Luu, B. V., & Chen, C. H. (2020). Greenwashing in environmental, social and governance disclosures. *Research in International Business and Finance*, 52, 101192. Advance online publication. DOI: 10.1016/j.ribaf.2020.101192

Zhang, S., Zhao, B., & Ventrella, J. (2018). *Towards an archaeological-ethnographic approach to big data: Rethinking data veracity (Vol. 1)*. Ethnographic Praxis in Industry Conference Proceedings., DOI: 10.1111/1559-8918.2018.01197

Zhou, L., & Zhang, D. (2007). Typing or messaging? Modality effect on deception detection in computer-mediated communication. *Decision Support Systems*, 44(1), 188–201. DOI: 10.1016/j.dss.2007.03.012

Zhu, F., & Zhang, X. M. (2010). Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of Marketing*, 74(2), 133–148. DOI: 10.1509/jm.74.2.133

Zubiaga, A., Kochkina, E., Liakata, M., Procter, R., Lukasik, M., Bontcheva, K., Cohn, T., & Augenstein, I. (2018). Discourse-aware rumour stance classification in social media using sequential classifiers. *Information Processing & Management*, 54(2), 273–290. DOI: 10.1016/j.ipm.2017.11.009

Zvokuomba, K., & Batisai, K. (2020). Veracity of women's land ownership in the aftermath of land redistribution in Zimbabwe: The limits of Western feminism. *Agenda (Durban, South Africa)*, 34(1), 151–158. DOI: 10.1080/10130950.2019.1690310