Article Type: Systematic Literature Review

Big data analytics in the public sector: A systematic literature review

Muhammad Choirul Huda* and Bambang Jatmiko

Abstract

Research aims: This study presents a systematic literature review of empirical research on Big Data Analytics (BDA) in the public sector. The purpose is to examine how BDA has been applied, the research settings, dominant themes, and the lessons learned across the literature.

Design/Methodology/Approach: Using a review method adapted from Hoque (2014), the study analyzed 64 articles published between 2015 and 2023 sourced from the Scopus database. Articles were selected based on relevance, journal quality, accessibility, and methodological rigor.

Research findings: The review identified six key thematic areas in public sector BDA research: accountability, energy efficiency, sustainability, innovation, analytics, and governance. The dominant theories found include Big Data Theory, Stakeholder Theory, and Agency Theory, while archival and survey methods were the most commonly employed research approaches. The United Kingdom, United States, and international multi-country studies contributed the most publications.

Theoretical contribution/ Originality: The novelty of this study lies in its exclusive focus on the public sector, its integrative thematic analysis using co-word mapping, and its implications for academic theory building, practical implementation, and policy formulation in public administration.

Practitioner/Policy implication: This study can be used these insights to develop effective BDA strategies, enhance performance, and foster trust in public institutions through more responsive and evidence-based decision-making.

Research limitation/Implication: Future research is expected to further study and research: (1) The Impact of Big Data on the Public Sector using journals or other references outside the Scopus database; (2) Future research can use keywords that are different from this research, (3) and can also access more journals to be reviewed.

Keywords: Big Data Analytics; Public Sector; Literature Review

JAI Website:

Introduction

In recent decades, the development of information technology has significantly changed the landscape of the public sector. One of the most important developments to emerge in this context is the Big Data phenomenon. In recent years, the rapid advancement of information and communication technologies has transformed the landscape of public sector management.

One of the most disruptive developments in this domain is the emergence of Big Data, which represents the ability to process, analyze, and interpret massive volumes of structured and unstructured data beyond the capabilities of traditional systems (Zikopoulos et al., 2017).

Governments around the world increasingly utilize Big Data analytics to enhance decision-making processes, improve service delivery, and respond to the evolving needs of society (Storey et al., 2015). However, despite growing interest in this area, there remains a lack of comprehensive synthesis that systematically maps the scope, trends, and insights of Big Data applications specifically in the public sector context.

Previous literature reviews have generally focused on Big Data in private or mixed-sector settings, without adequately addressing the distinctive challenges, governance issues, and implications found in the public domain. Moreover, there is no clear consolidation of lessons learned from empirical studies that could guide practitioners, policymakers, or future researchers.

Therefore, this study addresses this gap by conducting a systematic literature review of studies published between 2015 and 2023 that examine Big Data applications in the public sector. This study seeks to identify research trends, theoretical underpinnings, regional and methodological patterns, and to distill key themes and insights derived from the literature. The research aims to answer two key questions:

RQ1: *What has existing research in the public sector done in relation to Big Data (e.g., country background, themes, and findings)?*

RQ2: *What lessons can be learned from previous empirical studies?*

The literature was selected based on its novelty and discussion of important topics related to big data, to fulfill our research objectives. The publication years range from 2015-2023, with most of the literature focusing on big data ranging from 2017-2023. This is due to big data being a topic of focus recently. In addition, this study mostly includes research from several journals and articles published by reputable publishers. Due to the lengthy journal review process, most of the journals that discuss big data analysis, its tools and methods, and its applications turned out to be journals and articles.

The public sector was chosen for study in understanding big data because it may indicate the level of usage or application in terms of big data itself. In addition, big data has increased progress in companies or organizations that apply big data well, especially in the public sector. Thus, it is hoped that this research can open new knowledge for companies or organizations engaged in the public sector, industry, and even stakeholders on how to understand big data analysis that may occur in the future. The contribution of this research is to provide an analysis of the available literature on big data analytics in the public sector. Therefore, several applicable big data tools, methods and technologies

are discussed, and their available applications and opportunities in several decision domains are described.

The sections of this paper are organized as follows. Section 1 comprises the theoretical background that has been previously discussed. The subsequent section (Section 2) delineates the methodology employed to evaluate the journals. The subsequent section, Section 3, presents the frequency distribution of articles by publication quality and year, journal, research setting, theory, method, and country. This section also includes a companion word analysis to identify the research themes studied in this literature review. For each identified theme, the current research findings are discussed and summarized. The benefits that can be drawn from Big Data in the Public Sector are then summarized in Section 4. Finally, the conclusions and limitations of this study are presented in Section 5.

Literature Review

Big data refers to datasets that are so vast, varied, and rapidly generated that traditional data management tools cannot efficiently store, analyze, or interpret them (Gandomi & Haider, 2015). The emergence of big data stems from the exponential growth of digital interactions, sensors, mobile devices, and online activities that produce large-scale information continuously. According to De Mauro et al. (2016), big data encompasses more than just large volumes of information—it also involves the ability to extract meaningful insights from complex and unstructured sources. This ability has transformed how data is perceived not merely as records, but as strategic assets in modern decision-making systems.

Big data is commonly characterized by the "5Vs": Volume (massive amount of data), Velocity (speed of data creation and processing), Variety (diverse data types), Veracity (uncertainty or reliability of data), and Value (potential benefits derived from data analysis) (Wamba et al., 2015). Recent advancements in data storage, cloud computing, artificial intelligence, and analytics tools have enabled governments and organizations to manage big data more effectively. The development of platforms such as Hadoop, Spark, and other real-time analytics solutions further supports the shift toward data-centric planning and operations, not only in business but also in public administration.

In the public sector, big data plays an increasingly vital role in promoting evidence-based policy-making, enhancing service delivery, improving disaster response, and managing urban infrastructure (Janssen et al., 2017; Kalema & Mokgadi, 2017). Governments use big data analytics to monitor population trends, predict public health risks, detect fraudulent activities, and optimize resource allocation. For example, real-time traffic and weather data are now commonly integrated into smart city systems to improve transportation and safety. This analytical capability allows for more responsive, transparent, and citizen-centered governance, particularly when combined with open data initiatives.

In Indonesia, the government has introduced several digital transformation initiatives, such as SPBE (Electronic-Based Government System) and the "Satu Data Indonesia" program, aimed at standardizing and centralizing public data systems. These initiatives reflect a growing awareness of the importance of big data in enhancing public accountability and policy coordination (Puspita et al., 2023). Despite these efforts, challenges remain—especially related to interagency data integration, digital infrastructure gaps in rural areas, and limited data literacy among public officials (Gunawan et al., 2023). Nevertheless, the momentum for adopting big data in Indonesian public administration continues to grow, particularly in urban planning, e-government services, and tax management.

The integration of big data in public sector governance is closely linked to core values such as transparency, accountability, responsiveness, and efficiency. By utilizing big data analytics, public institutions can proactively detect inefficiencies, reduce bureaucratic corruption, and enhance public trust (Kettunen et al., 2022). Studies also suggest that big data can bridge the gap between citizens and the state by enabling more inclusive decision-making and feedback mechanisms (Zhou & Yang, 2021). Therefore, understanding the role and impact of big data in the public sector is essential not only for technological advancement, but also for improving the integrity and sustainability of governance systems.

Research Method

This methodology section explains the literature review and search strategy used by the researcher. This research adopts the procedures of Hoque (2014) with some modifications, including the application of shared word analysis in mapping themes (Poje & Zaman Groff, 2021). There are at least two reasons for adopting this procedure. First, the research question of this study implies the use of a descriptive review methodology, which has three definite stages: extraction, analysis and synthesis, among others (for details, see Xiao and Watson, (2019). Second, it suggests that a systematic literature review should identify those motivating research topics and questions, apply screening criteria, balance between breadth and depth (relevance criteria in this study), focus on concepts, provide synthesis and interpretation of analysis and follow an organized structure (Fisch & Block, 2018).

According to Hoque (2014) the procedure corresponds to a three-stage descriptive review methodology and six recommendations for a good systematic literature review. The application of Hoque's (2014) procedure in this study (Figure 1) is detailed as follows. The first step is to determine the research topic, namely Big Data in the Public Sector.

As indicated in the previous section, this study worked on two research questions, what research has been found on this topic and what can be gained from the reported results. The next step was to collect the journals drawn from this research using the Scopus database. The relevant journals to be reviewed were identified based on a research

strategy involving search terms, journal category, journal type, accessibility, language, publication quality and relevance.
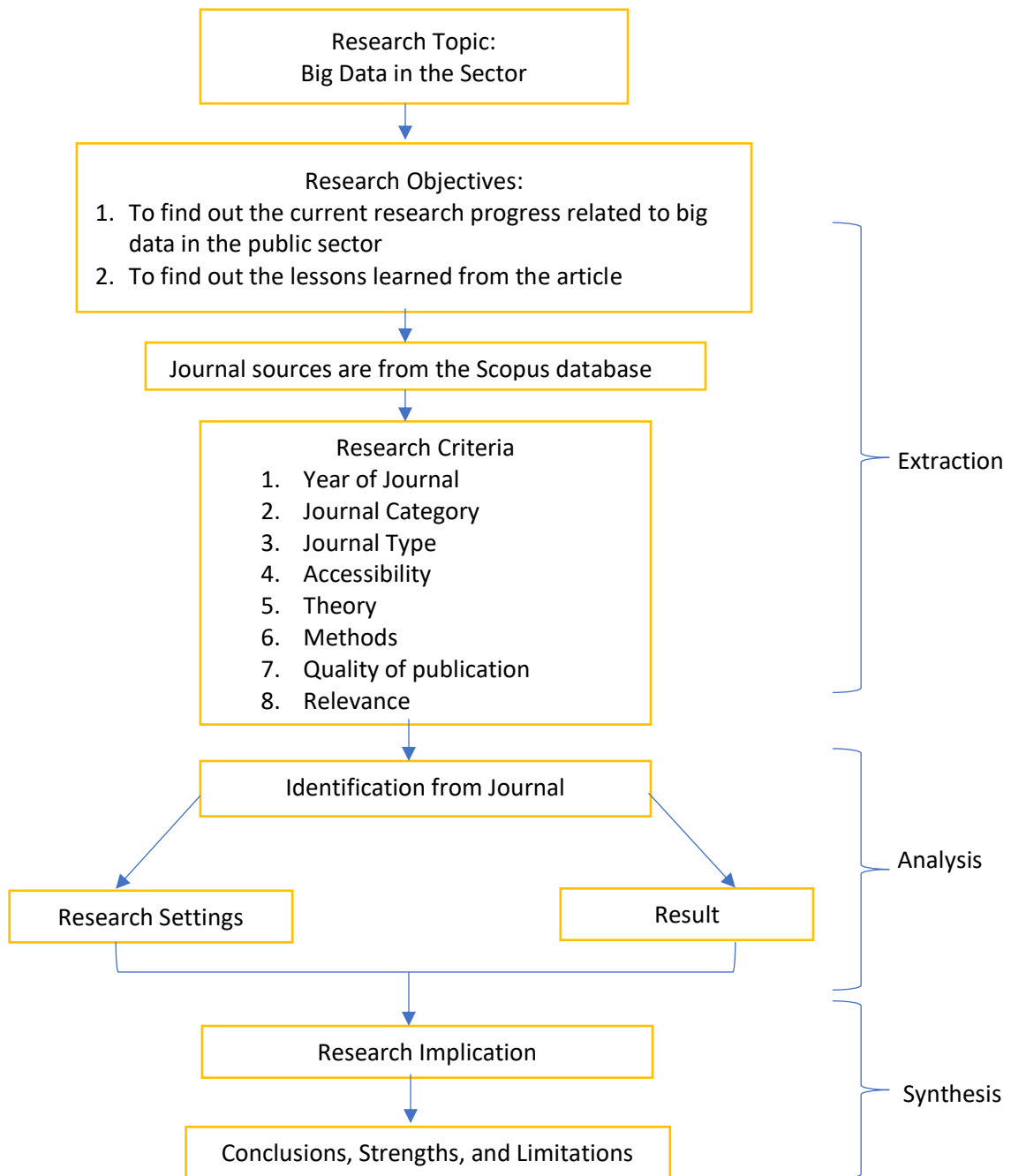

Figure 1 Literature review process adopted from Hoque (2014) with modifications.

The review process started by formulating the research questions and defining the topic: Big Data in the Public Sector. The literature was then identified using Scopus as the primary database, applying the keywords "Big" AND "Data" AND "Sector" AND "Public".

Only empirical studies published between 2015 and 2023 were included. This time range was chosen because Big Data analytics in public administration started gaining global traction in mid-2010s, and the years after 2015 represent a rapid increase in interest and technological development in this area. Researchers chose journals that are open access or can be downloaded. The journals were also written in English. The Scimago journal ranking was used to determine the quality of publications, where unindexed journals were excluded. Journals that were not directly relevant to the topic under investigation were also removed. By applying all the criteria, the final number of journals to be analyzed was 71 out of the initial 415 journals, then of the 71 journals that were successfully downloaded pdf files were 64 journals.

All selected journals were read and manually tagged or coded. Some of the information to be identified from each journal is the country background and findings. Since the reported results were discussed based on theme clusters, co-word analysis was used to classify the journals based on the keyword similarity of the authors. Co-word analysis was performed using VOSviewer software. Therefore, lessons learned from this study can be carried out.

Result and Discussion

Frequency distribution of articles by publication quality and year

The frequency distribution of journal publication quality and year is shown in Figures 2 and Table 1. Out of the 64 selected articles, a majority were published in high-ranking journals (Q1: 46.9%), reflecting strong academic interest in the topic. The peak in publications occurred in 2023, accounting for 18.8% of total studies. This trend confirms the growing urgency and relevance of Big Data analytics in public sector discourse, driven by global digital transformation and increasing emphasis on data-driven governance.
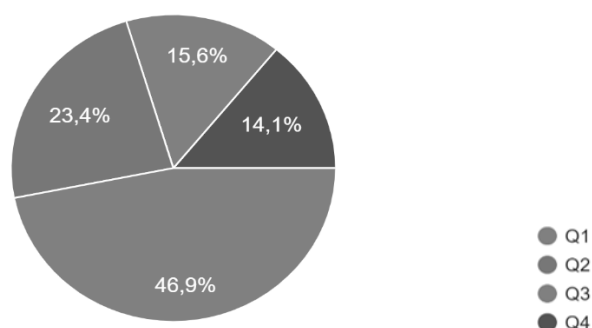

Figure 2 Frequency distribution of articles based on publication quality.

For all quality publications, except Q4, there are far more journals published in 2023 compared to other years. This is quite reasonable considering that in previous years there were not too many researchers who studied or analyzed big data in the public sector. Based on 64 journals that have been reviewed, 6 journals (9.4%) were published in 2015, 2 journals (3.1%) were published in 2016, 8 journals (12.5%) were published in 2017, 5

journals (7.8%) were published in 2018, 6 journals (9.4%) were published in 2019, 9 journals (14.1%) were published in 2020, 6 journals (9.4%) were published in 2021, 10 journals (15.6%) were published in 2022, and 12 journals (18.8%) were published in 2023. In addition to showing an increase in publication quality, this also indicates an increase in the attention given to and importance of this topic, given that high journal quality comes with a high scale of impact.

Table 1. Frequency Distribution of articles by Year

Year	Number	*Frequency %*
2015	6	9.4
2016	2	3.1
2017	8	12.5
2018	5	7.8
2019	6	9.4
2020	9	14.1
2021	6	9.4
2022	10	15.6
2023	12	18.8

Frequency distribution of articles by journal

Figure 3 presents the distribution of journals, by journal, for the top 10 journals identified in this study. The top 10 journals were selected from 19 journals based on frequency, which implies that these journals have the most journals on Public Sector and Big Data topics. Journal of Business Research is the journal that publishes the most journals, with 10 out of 64 journals (15.63%). This may be due to the journal's focus on research exploring current and emerging issues in the field of Big Data. Since international has become a popular place to study this topic, Accounting, Auditing and Accountability Journal is in second place with eight journals (12.5%). This journal is indeed specialized in accounting, auditing, and accountability issues especially in the Asian region including Central Asia, East Asia, South Asia, Southeast Asia, and the Middle East. Thus, many journals on financial studies related to big data are interested in investigating this phenomenon.
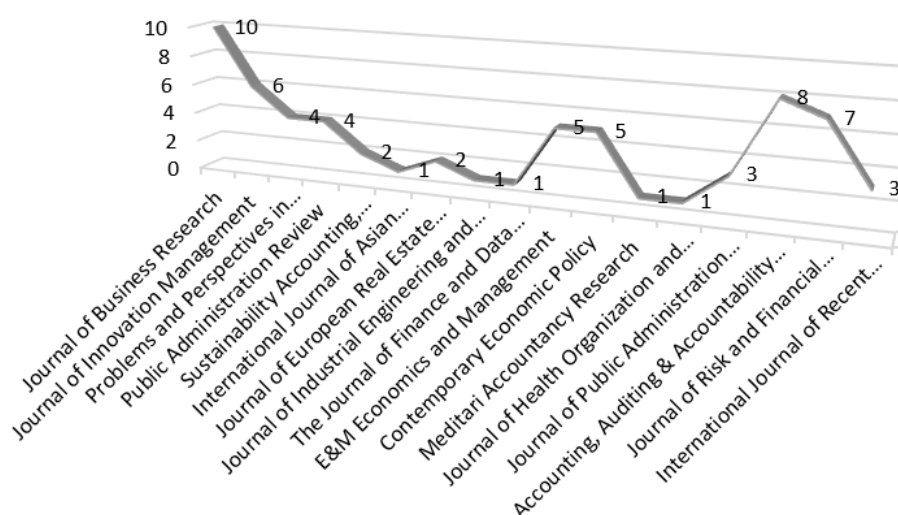
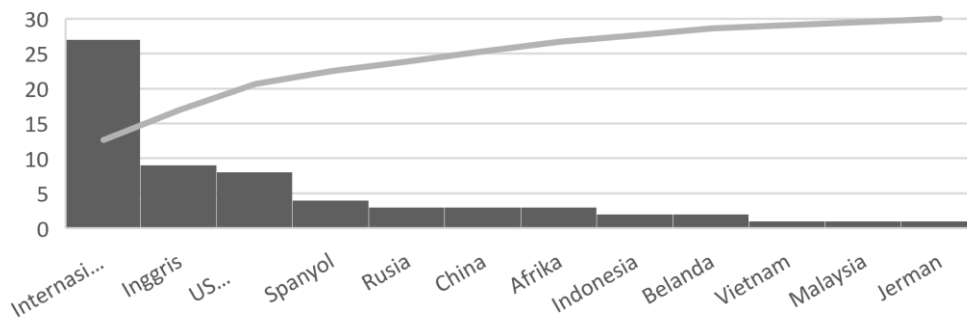**Figure 3** Frequency distribution of articles by journal

Meanwhile, the journal that is in third place is published in the Journal of Risk and Financial Management with seven journals (10.9%). In fourth place is the Journal of Innovation Management, which is six journals (9.4%), followed by the sixth position in the Contemporary Economic Policy and Economics and Management journals with 5 journals each (7.8%). Then in the next position there are 4 (6.25%), 3 (4.7%), and 2 (3.1%) journals respectively. And the other five journals total 7.8%, each of which includes one journal.

Frequency distribution of articles based on the country of research

This study classifies the country of the reviewed journals into several categories, as shown in Figure 4, in the following manner. First, journals that use the background of one or more countries from the same region are classified and labeled using the name of the region. Second, journals that use the background of many countries from different regions but still on the same continent are grouped by continent. Third, journals that display samples from many countries from various regions and continents are named "International".

International is the most frequently used background by research on this topic, which is 27 journals (42.2%). This shows that various studies on big data have been conducted in various countries in the world. The next most frequently used country backgrounds in this study are the United Kingdom and the United States with 9 journals (14.1%) and 8 journals (12.5%) respectively. Then Spain with 4 journals (6.25%), then Russia, China, and Africa each have 3 journals (4.7%), and Indonesia and the Netherlands both have 2 journals (3.1%), and finally Vietnam, Malaysia, and Germany each have 1 journal (1.6%). The United Kingdom and United States were among the top contributors, likely due to advanced digital infrastructure and active research communities in public governance. Indonesia, though less frequent, reflects a growing interest in integrating data analytics into public policy, which is also evident in regional digital transformation initiatives.

Frequency Distribution of articles by Research Country Setting

Figure 4 Frequency distribution of articles based on the country of research

Frequency distribution of articles based on theory

The frequency distribution of articles based on theory is shown in table 2. Of the 64 journals reviewed, the average theory that is widely used by researchers is Big Data Theory (56.3%), followed by Stakeholder Theory and Agency Theory. This suggests that while technical exploration is important, researchers are also concerned with the implications for transparency, accountability, and relationships between public institutions and the communities they serve. The use of Agency Theory reflects tension between public agents (e.g., government institutions) and principals (e.g., citizens) in data governance.

Then, the next theory that is widely used by previous research is Accounting Theory, then Contingency Theory, and up to the theory that is rarely used, namely Scientific Theories and Social Trust Theory.

Table 2 Frequency Distribution of articles based on Theory

Theory	Number	Frequency %
Big Data Theory	36	56,3%
Stakeholder Theory	23	35,9%
Theory of Innovation	4	6,3%
Agency Theory	20	31,3%
Accounting Theory	14	21,9%
Contingency Theory	11	17,2%
Institutional Theory	8	12,8%
Theory of Modern Banking	2	3,1%
Ownership theory	3	4,7%
Theory of innovation	2	3,1%
Scientific theories	1	1,6%
Social trust theory	1	1,6%

Frequency distribution of articles based on method

The frequency distribution of journal methods is shown in Figure 5. Based on 64 journals that have been reviewed, the method most often used by previous researchers is Archival

(secondary) at 31.3%. According to researchers, this method is widely used because it is relatively easy to obtain data, because researchers only need to access the website or request data from certain agencies to retrieve the required data. The second method that is often used is survey (primary) at 26.6%, which according to researchers is a method that is often used because it is also relatively easy to obtain the desired data, this method can be done by coming directly and giving questionnaires to respondents or can be online sending the questionnaire.

Then for the third method that is most widely used is Literature Review at 18.8% which is done by collecting various journals according to the topics desired by the researcher and summarizing them into one discussion finding, followed by the Interview method at 15.6% which is done by interviewing parties or respondents who are trusted by researchers to provide good information, and the last is Mix Method at 7.8% which is a combination of 2 or more research methods.
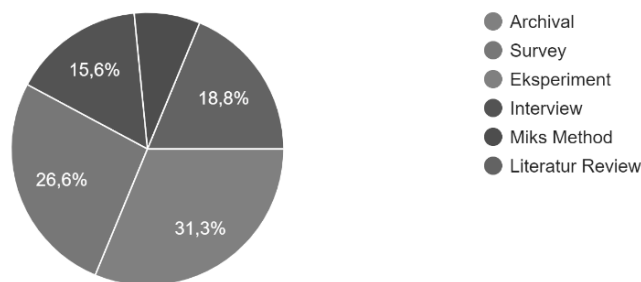

Figure 5 Frequency distribution of articles by method.

Co-Word Analysis

Based on Figure 6 below shows the results of co-word analysis using author keywords for the 64 journals reviewed. The frequency of a word in a journal keyword is indicated by the size of the circle in which the word is located. Meanwhile, the color indicates the grouping that shows how often the words appear together in the journal keywords. The closer and thicker the lines connecting the words, the closer the relationship. As the topic studied in this review, the most frequently occurring words are Big Data and Public Sector, which are symbolized by the two largest circles in the middle. There are six groups identified by VOSviewer related to Big Data in the Public Sector, as implied by the number of different colors. Some keywords appear in the wrong groups because the classification is based on the frequency of occurrence of the words. For example, "governance" is not included in the same group as "regulation" and "personality traits". Another example is "big data analytics", which is not grouped with "industry 4.0" and "innovation". Therefore, this study made adjustments and reductions to clarify the grouping, leaving six groups of themes.
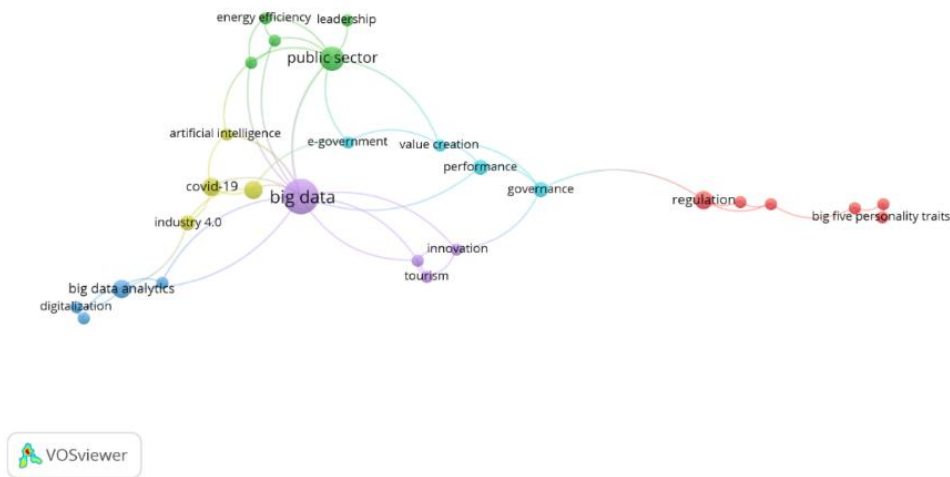
Figure 6 Analysis of co-occurrence of words in VOSviewer

Figure 6 above illustrates the six clusters that will be discussed further in this journal. The six clusters are: accountability, energy efficiency, sustainability, innovation, big data analytics and governance.

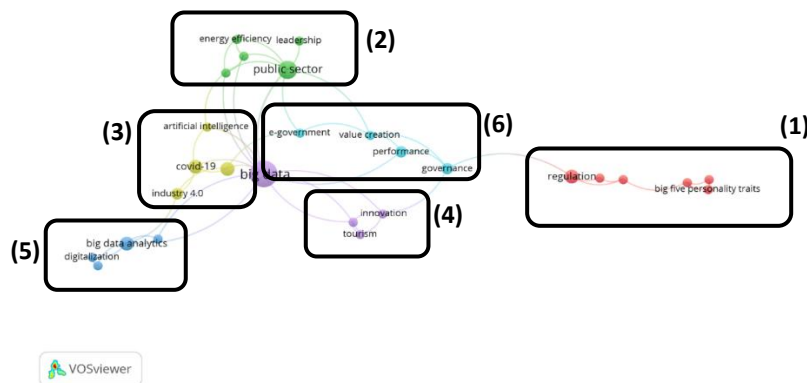**Existing research findings based on theme**

Figure 7 Grouped themes identified for further discussion

Figure 7 shows the themes from the clusters that have been grouped into six sections. Each theme offers a unique lens to understand the role of Big Data in transforming public sector functions. Interestingly, these themes are interconnected: 1) Accountability and Governance emphasize ethical and institutional responsibilities; 2) Innovation and Big Data Analytics focus on technological capacity and service improvement; 3) Energy Efficiency and Sustainability tie Big Data to environmental and economic goals.

This clustering reinforces that Big Data initiatives are not siloed they must be understood in an integrated, cross-functional framework. For instance, enhancing governance through data-driven innovation requires accountability mechanisms and sustainable infrastructure.

Accountability

According to Power (2014) accountability is a key principle in public sector management that emphasizes responsibility and transparency in decision-making and the use of public resources. In the context of Big Data, the relationship between accountability and Big Data becomes increasingly important, because Big Data has the potential to strengthen accountability mechanisms in various ways. According to Popoola and Mohd Jamel (2020) accountability refers to the obligation of a person or organization to be responsible for the actions, decisions, and results of activities carried out. This concept includes transparency, integrity, and the obligation to provide accountability for actions or decisions taken. In Malaysia, it was found that big data resulted in government accountability having a significant effect and being positively related to public acceptance of the implementation of SST (Sales and Service Tax) in Malaysia. This further strengthens government accountability is very important for public acceptance of the implementation of SST (Sales and Service Tax) because it has economic implications (Popoola & Mohd Jamel, 2020).

Similar to Malaysia, the impact of big data is also seen in accountability in Indonesia (Khoirul Aswar, 2021). The results obtained in the form of effective managerial accountability influence the administration of accrual accounting implementation. This study implies that the implementation of effective accountability depends on attention related to humans and culture. In addition, according to Lavertu (2016) shows that big data also has a positive impact on accountability in public administration which is the latest form of ongoing efforts to base public decision-making on measurable quantities and to increase transparency and responsibility.

Furthermore, research from Pollock and Williams (2015) found that accountability through the development of big data also increases the use of information technology in the industrial market. The positive impact of the implementation of big data was also found in other studies, namely Shah et al. (2017) which shows that accountability has a good or positive influence on organizations or companies operating in the public sector. According to Aoki et al. (2023) accountability has a significant impact on economic development in a country, with the implementation of good big data it will certainly encourage development in various sectors, especially the economic sector. The use of big data and accountability in the tourism sector is an increasing trend. Big data creates hope for a better understanding of tourism demand and adjustment of supply by tourism companies to meet tourist demand and profitable tourism business activities (Belias et al., 2021).

Based on most of the explanations or findings from researchers, it is concluded that accountability can have a significant and positive influence related to various aspects in companies or organizations, especially those engaged in the public sector. The use of big data in supporting accountability is also highly recommended because it can increase responsibility and transparency for all parties. Thus, accountability supported by the application of big data analysis shows a good influence on companies, organizations, or governments in the public sector.

Energy Efficiency

According to Moerkerken et al. (2021) energy efficiency refers to the extent to which a system or process is able to produce the desired output using less energy. In the context of energy, efficiency is a measure of how well energy is used and not wasted. The main goal of energy efficiency is to reduce energy consumption without reducing the quality or quantity of output produced (Moerkerken et al., 2021).

Energy efficiency in relation to big data in the public sector involves the use of sophisticated data analysis technologies to optimize energy use in various government operations (Deepica et al., 2022). Big data can be used to collect and analyze information about energy consumption in various government facilities, such as offices, schools, and public facilities (Carrillo Sánchez et al., 2020). This analysis provides in-depth insights into energy consumption patterns, allowing governments to identify areas where efficiency improvements are needed.

According to Wilbert et al. (2018) big data can be used to monitor and manage energy infrastructure, such as lighting, heating, cooling, and public transportation systems. By utilizing sensors and connected smart devices, governments can collect real-time data on the performance of these infrastructures. Big data analysis can then provide recommendations for improvements and adjustments to increase energy efficiency, reduce waste, and lower operational costs (Moerkerken et al., 2021).

Big data can also be used to design smarter and more targeted energy policies. Big data analysis allows governments to understand the impact of existing energy policies, identify opportunities for innovation, and develop new, more efficient strategies (Pratama, 2019). This includes using big data to predict energy demand, manage the electricity grid, and support the transition to renewable energy sources (George et al., 2019).

According to Radicic and Petković (2023) big data can support energy awareness and public education programs. By analyzing energy consumption data at the household or community level, governments can design more precise and personalized information campaigns. This analysis can help raise public awareness of energy-saving practices and encourage behavioral changes that support energy efficiency.

Based on research from Moerkerken et al. (2021) shows that big data can be used to monitor and evaluate the impact of sustainable energy projects. By analyzing the performance data of these projects, governments can assess the effectiveness of their investments in reducing energy consumption, measure the reduction in greenhouse gas emissions, and evaluate the social and economic benefits of implemented energy efficiency measures. This helps governments make more informed decisions for the future (Sait & Ali, 2022).

Based on the explanation of previous studies, it is concluded that energy efficiency integrated with big data in the public sector not only supports operational efficiency, but also creates a more sustainable environment, reduces environmental impacts, and

provides significant economic benefits. This combination has the potential to realize positive transformations in the provision of public services and achieve sustainable development goals as a whole.

Sustainability

Sustainability according to Huy and Phuc (2023) is a concept that refers to the ability to maintain or maintain a system or process for a long period of time without sacrificing economic, social, and environmental balance. Big Data enables the public sector to optimize the use of resources more efficiently. With in-depth data analysis, the government can identify areas where efficiency can be improved, leading to more sustainable resource management. Big Data can be used to support the development of sustainable policies and provide in-depth insights into trends and policy impacts on various aspects of sustainability, such as economic, environmental, and social (Huy & Phuc, 2023).

According to Moerkerken et al. (2021) sustainability in the context of big data and the public sector refers to how the use of big data can have a long-term positive impact on society, the environment, and the government. Big data can support sustainability in the public sector by increasing efficiency and effectiveness in service delivery. Big data analysis can help governments identify ways to optimize resource use, reduce waste, and improve overall service quality (Bortoló et al., 2023).

Research from Tilba (2022) explains that sustainability in big data can be linked to environmental protection. The big data analysis process can help governments identify ways to reduce environmental impacts, monitor pollution, and design policies that support sustainability principles. In addition, the use of big data in managing transportation, energy, and waste can help create smart cities that are more sustainable and environmentally friendly (Shah et al., 2017). In the public sector, big data sustainability also includes ethical and privacy aspects. Governments need to ensure that data collection, storage, and analysis are carried out in accordance with privacy norms and individual rights. Strong and transparent policies on data management can provide a basis for public trust in government big data initiatives (Moerkerken et al., 2021).

According to Carrillo Sánchez et al. (2020) big data can help governments face sustainability challenges, including climate change, health crises, and natural disasters. Big data analytics can provide deep insights into trends and changes that may impact sustainability, enabling governments to respond quickly and effectively. To achieve sustainability in big data, collaboration between the public, private and civil society sectors is crucial. Joint initiatives can help address complex challenges, including ensuring equitable access to the benefits of big data, supporting sustainable innovation and building capacity to manage and understand the implications of big data. With this holistic approach, big data can be an effective tool in supporting sustainable development at multiple levels.

Innovation

According to Roy's research (2016), innovation in relation to big data in the public sector involves the use of sophisticated data analysis technology to generate deeper understanding and valuable information from large and complex data. Big data allows the government to identify patterns, trends, and insights that cannot be obtained through traditional methods. Radicic and Petković (2023) provide an explanation that the implementation of Big Data involves the use of the latest technologies such as advanced data analysis, artificial intelligence, and machine learning. This creates opportunities to adopt innovative solutions in decision making and public sector operations. Big Data can be used to improve the quality and innovation of public services. In-depth data analysis can provide insights to create new services or improve existing ones, according to the needs of the community.

The application of big data allows the government to collect, store, and analyze large amounts of data from various sources (Moro et al., 2018). Innovation lies in the ability to explore this data using sophisticated technologies such as machine learning and predictive analysis. Thus, the government can gain deep insights and understand the complexity of the problems faced by the community. Innovations in big data analytics enable governments to make more informed and responsive decisions. Using intelligent algorithms, governments can identify relevant patterns and trends from big data, enabling more informed decision-making. This carries the potential to improve operational efficiency, allocate resources more effectively, and design policies that better suit the needs of the community.

According to Moro et al. (2018) the adoption of big data in the public sector expands the government's capacity to plan long-term policies and address potential problems. Predictive analytics enables governments to forecast future developments and identify potential risks or opportunities. With this, big data innovation can be a strategic tool in sustainable development planning, risk management, and governance sustainability. Innovation is seen in the ability of big data to improve response to crises and emergencies. By analyzing real-time data, governments can monitor and respond to critical situations more quickly and effectively.

Big data innovation also includes efforts to increase public participation in decision-making (Radicic & Petković, 2023). Governments can use big data to design participatory platforms that allow citizens to provide feedback, voice their preferences, and participate in the decision-making process. This creates a more transparent, inclusive and responsive government to the interests of the people (Roy, 2016). However, with all this innovative potential, it is important to ensure strong privacy protection policies to keep people's personal data safe.

Big Data Analytics

Big Data refers to large volumes of complex, high-velocity, and high-variety data that are generated continuously from various sources such as sensors, social media, transaction

records, and digital interactions. The defining characteristics of Big Data are often summarized by the "5 Vs": Volume, Variety, Velocity, Veracity, and Value (Gandomi & Haider, 2015). Big Data is not simply about having a large dataset, but also about the challenges in storing, managing, and processing it.

In contrast, Big Data Analytics (BDA) is the process and set of technologies used to extract meaningful insights from Big Data. BDA involves the application of advanced analytical techniques, such as machine learning, statistical modeling, data mining, and predictive analytics, to uncover patterns, trends, correlations, and decision-support outputs (Wamba et al., 2015). While Big Data is the raw material, Big Data Analytics is the toolset that turns data into knowledge.

Thus, Big Data is the "what", while Big Data Analytics is the "how"—Big Data describes the phenomenon, and BDA describes the methods used to harness its value for decision-making.

Big data analytics in relation to the public sector has become an invaluable tool for improving the efficiency, responsiveness, and quality of government services. Big data analytics enables governments to collect and integrate data from multiple sources, including administrative records, sensors, social media, and more (Kalema & Mokgadi, 2017). This creates a very large and diverse data set, which can be processed to gain deeper insights into the needs and behaviors of the community.

According to Dekimpe (2020) explains that big data analytics enables governments to make more informed and measurable decisions. By using advanced data analytics techniques, such as machine learning and predictive analytics, governments can identify patterns and trends that are not easily visible through traditional methods. This helps in designing more effective policies, intelligent resource allocation, and improving operational efficiency in various areas, including health, education, and transportation.

Big data analytics can be used to understand and monitor community satisfaction (Radicic & Petković, 2023). By analyzing feedback data from citizens through various channels, governments can gain better insights into citizens' perceptions of public services. This allows governments to tailor policies and programs to better suit the needs of the community. In the context of crisis prevention and response, big data analytics helps governments to detect potential risks earlier (Huy & Phuc, 2023). Real-time data monitoring enables rapid response to emergency events, such as natural disasters or the spread of diseases. Big data analytics can also be used for risk planning and management, helping governments anticipate and address potential threats more effectively (Sait & Ali, 2022).

According to Radicic and Petković (2023), big data analytics can strengthen participatory and transparent aspects of government. By leveraging big data to form participatory platforms, governments can involve citizens in decision-making, gain direct input from citizens, and build public trust. Thus, big data analytics not only provides operational and strategic benefits but also creates a more open and inclusive government.

Big Data Analytics involves the use of sophisticated analytical techniques to extract insights from the large volumes of data generated by the public sector (Antony & Sony, 2020). By using advanced data analysis algorithms, governments can gain a deeper understanding of trends, patterns, and relationships in data. Big data analytics enables the public sector to make more informed and timely decisions (Lu et al., 2016). By analyzing data in real-time, governments can respond quickly to changing conditions and optimize their policies and programs.

Governance

Articles in this theme cluster discuss governance in relation to the application of big data in the public sector. Governance in relation to big data in the public sector includes how the government manages, secures, and uses big data to achieve public service goals and better decision-making (Lu et al., 2016). Data governance includes policies and procedures implemented by the government to ensure that the collection, storage, and use of big data is carried out ethically and in accordance with privacy regulations (Sait & Ali, 2022).

According to Carrillo Sánchez et al. (2020) governance in big data in the public sector involves aspects of transparency and accountability. The government must provide clear information to the public about how their data is used. Accountability mechanisms are also needed to ensure that data is used in accordance with public objectives and is not misused. This transparency and accountability build the foundation of public trust in government big data initiatives (Carrillo Sánchez et al., 2020).

Governance includes data risk management. Governments need to identify and manage risks related to data security, privacy breaches, and potential negative impacts of big data use. This involves implementing robust information security measures, risk mitigation policies, and a deep understanding of the ethical and social implications of big data analytics (Moerkerken et al., 2021).

Big data governance should also include building capacity and expertise within the government. The government needs to have personnel trained and skilled in data analytics, information technology management, and data security. This creates an environment where the government can effectively manage and use data for better policy purposes (Popoola & Mohd Jamel, 2020).

According to Weerakkody et al. (2017) big data in government in the public sector involves public engagement. The government must effectively communicate big data policies and practices to citizens. Moreover, the government needs to seek feedback and input from the public on how data use can be improved or improved. Public engagement creates a government that is more open, inclusive, and responsive to the needs and concerns of its citizens in the era of big data (Radicic & Petković, 2023).

Lessons Learned

This section discusses several lessons that can be learned from the analysis of journal findings related to big data analysis in the public sector. This includes lessons for organizations, governments, and companies, to help them deal with increasingly rapid technological changes or similar situations that may arise in the future. Rapid technological developments have encouraged companies or organizations, especially those in the public sector, to follow and implement them in their operational processes.

The application of big data analysis in the public sector brings a number of lessons that have the potential to change the way the government interacts with the public and carries out public service functions. The findings of this study hold significant implications for various stakeholders in the public sector, including policymakers, practitioners, and researchers.

For policymakers, the themes identified in this review such as accountability, governance, and sustainability highlight the need to implement strong regulatory frameworks for data ethics, privacy, and responsible data use. These themes emphasize the importance of transparent governance systems supported by real-time data monitoring and performance evaluation.

For practitioners in public sector organizations, the insights regarding innovation and energy efficiency point to concrete applications of Big Data for service delivery enhancement. Government institutions can utilize Big Data analytics not only to optimize operations and resource allocation, but also to predict public needs and tailor services accordingly.

For researchers, this review uncovers clear gaps in theory application (e.g., low usage of contingency and social trust theories), underexplored regions (e.g., Africa and Southeast Asia), and emerging areas such as predictive governance, digital trust, and data literacy in the public sector. These present opportunities for future empirical research and theory development.

Conclusion

This study aims to conduct a systematic review of empirical literature discussing the application of Big Data Analytics (BDA) in the public sector. Using a literature review approach that adapts the method from Hoque (2014), this study analyzes 64 articles from the Scopus database published between 2015 and 2023. The main focus of the research is to identify trends, country backgrounds, dominant themes, methodological approaches, as well as theoretical and practical contributions from previous studies related to BDA in public sector governance. The results of the review show that BDA research in the public sector is dominated by six main themes: accountability, energy efficiency, sustainability, innovation, analytics, and governance. The most widely used theory is Big Data Theory, followed by Stakeholder Theory and Agency Theory. Most of

the research was conducted in North America, Europe, and multinational studies, with archival and survey methods being the most frequently used.

Methodologically, this research contributes or implies by presenting a thematic mapping based on co-word analysis, resulting in a more structured and integrative topic classification. Theoretically, this study reinforces the position of BDA as an important approach in building a transparent, adaptive, and accountable data-driven governance system. Practically, these findings can serve as a reference for policymakers and public sector practitioners in developing more effective big data implementation strategies, particularly for improving public service quality and evidence-based decision-making.

However, this study has several limitations. First, the review only utilized journals from the Scopus database, thus not covering references from other databases such as Web of Science or national publications. Second, the selection of keywords is still limited, so the potential for other relevant literature has not been fully captured. Therefore, further research is expected to expand the scope of sources, use more varied keywords, and explore more deeply the impact of BDA implementation on public sector performance in various contexts of developing countries, including Indonesia.

References

Antony, J., & Sony, M. (2020). An empirical study into the limitations and emerging trends of Six Sigma in manufacturing and service organisations. *International Journal of Quality & Reliability Management*, 37(3), 470-493. https://doi.org/10.1108/IJQRM-07-2019-0230

Aoki, G., Ataka, K., Doi, T., & Tsubouchi, K. (2023). Data-driven estimation of economic indicators with search big data in discontinuous situation. *The Journal of Finance and Data Science*, 9, 100-106. https://doi.org/10.1016/j.jfds.2023.100106

Belias, D., Malik, S., Rossidis, I., & Mantas, C. (2021). The use of big data in tourism: current trends and directions for future research. *Academic Journal of Interdisciplinary Studies*, 10(5), 357-364. https://doi.org/10.36941/ajis-2021-0144

Bortoló, G. M., Valdés, J. Á., & Nicolas-Sans, R. (2023). Sustainable, technological, and innovative challenges post Covid-19 in health, economy, and education sectors. *Technological Forecasting and Social Change*, 190, 122424. https://doi.org/10.1016/j.techfore.2023.122424

Carrillo Sánchez, L., García Calderón, C., Cuevas Contreras, T., & Ruiz León, A. (2020). Big data in the Mexico City brand in the context of audience fragmentation. *Investigaciones Turísticas* (20), 124-142. https://doi.org/10.14198/INTURI2020.20.06

Chen, H., Chiang, R. H., & Storey, V. C. (2015). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*, 36(4), 1165-1188. https://doi.org/10.2307/41703503

De Mauro, A., Greco, M., & Grimaldi, M. (2016). A formal definition of Big Data based on its essential features. *Library Review*, 65(3), 122–135. https://doi.org/10.1108/LR-06-2015-0061

Deepica, S., Kalavathi, S., Angelin Blessy, J., & Maria Manuel Vianny, D. (2022). Machine Learning Based Approach for Energy Management in the Smart City Revolution.

Hybrid Intelligent Approaches for Smart Energy: Practical Applications, 161-179. https://doi.org/10.1002/9781119821878.ch9

Dekimpe, M. G. (2020). Retailing and retailing research in the age of big data analytics. *International Journal of Research in Marketing*, 37(1), 3-14. https://doi.org/10.1016/j.ijresmar.2019.09.001

Fisch, C. and Block, J. (2018), "Six tips for your (systematic) literature review in business and management research", *Management Review Quarterly*, Vol. 68 No. 2, pp. 103-106. https://doi.org/10.1007/s11301-018-0142-x

Gandomi, A., & Haider, M. (2015). *Beyond the hype: Big data concepts, methods, and analytics*. *International Journal of Information Management*, 35(2), 137–144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007

George, B., Van de Walle, S., & Hammerschmid, G. (2019). Institutions or contingencies? A cross‐country analysis of management tool use by public sector executives. *Public Administration Review*, 79(3), 330-342. https://doi.org/10.1111/puar.13018

Gunawan, H., Wulandari, D., & Putri, A. (2023). Big data in Indonesian public service transformation: Opportunities and barriers. *Jurnal Administrasi Publik Digital*, 2(1), 55–68.

Hoque, Z. (2014), "20 Years of studies on the balanced scorecard: trends, accomplishments, gaps and opportunities for future research", *The British Accounting Review*, Vol. 46 No. 1, pp. 33-59. https://doi.org/10.1016/j.bar.2013.10.003

Huy, P. Q., & Phuc, V. K. (2023). Acquiring Sustainability in The Prospect of Digital Transformation Through Global Brain-Reflective Accounting Practices Application. *International Journal of Economics & Management*, 17(2). https://doi.org/10.47836/ijeam.17.2.06

Janssen, M., et al. (2017). Big data for policy making: Challenges and opportunities. *Government Information Quarterly*, 34(1), 1–7. https://doi.org/10.1016/j.giq.2016.02.003

Kalema, B. M., & Mokgadi, M. (2017). Developing countries organizations' readiness for Big Data analytics. *Problems and perspectives in management* (15, Iss. 1 (cont.)), 260-270. https://doi.org/10.21511/ppm.15(1-1).2017.13

Kettunen, P., Kallio, J., & Ropponen, J. (2022). Big data and accountability in digital government. Information Polity, 27(1), 3–15.

Khoirul Aswar, E. a. W. J. (2021). Implementation of accrual accounting by the Indonesian central government: An investigation of social factors. *Public and Municipal Finance*, 10(1), 151-163. https://doi.org/10.21511/pmf.10(1).2021.12

Lavertu, S. (2016). We all need help:"Big data" and the mismeasure of public administration. *Public Administration Review*, 76(6), 864-872. https://doi.org/10.1111/puar.12436

Lu, W., Chen, X., Ho, D. C., & Wang, H. (2016). Analysis of the construction waste management performance in Hong Kong: the public and private sectors compared using big data. *Journal of Cleaner Production*, 112, 521-531. https://doi.org/10.1016/j.jclepro.2015.06.106

Moerkerken, A., Duijndam, S., Blasch, J., van Beukering, P., & Smit, A. (2021). Determinants of energy efficiency in the Dutch dairy sector: dilemmas for sustainability. *Journal of Cleaner Production*, 293, 126095. https://doi.org/10.1016/j.jclepro.2021.126095

Moro, M. A., McKnight, U. S., Smets, B. F., Min, Y., & Andersen, M. M. (2018). The industrial dynamics of water innovation: A comparison between China and Europe. *International Journal of Innovation Studies*, 2(1), 14-32. https://doi.org/10.1016/j.ijis.2018.03.001

Poje, T. and Zaman Groff, M. (2021), "Mapping ethics education in accounting research: a bibliometric analysis", *Journal of Business Ethics*, pp. 1-22. https://doi.org/10.1007/s10551-021-04846-9

Pollock, N., & Williams, R. (2015). Industry analysts–how to conceptualise the distinctive new forms of IT market expertise? *Accounting, Auditing & Accountability Journal*, 28(8), 1373-1399. https://doi.org/10.1108/AAAJ-03-2015-1989

Popoola, O. M. J., & Mohd Jamel, N. E. S. (2020). Examining public acceptance choice causes on sales and service tax implementation in Malaysia. *Problems and perspectives in management*, 18(4), 228-246. https://doi.org/10.21511/ppm.18(4).2020.20

Power, D. J. (2014). Using Big Data for Analytics and Decision Support. *Journal of Decision Systems*, 23(2), 222-228. https://doi.org/10.1080/12460125.2014.888848

Pratama, A. B. (2019). The landscape of public service innovation in Indonesia: A comprehensive analysis of its characteristic and trend. *Innovation & Management Review*, 17(1), 25-40. https://doi.org/10.1108/INMR-11-2018-0080

Puspita, A.F., Lilik, P., and Rusydi, M.K. (2023), "The Role of Big Data in Building Smart City for Regional Taxation Systems", *Journal of Research and Applications: Accounting and Management*, Vol. 6 No. 2. https://doi.org/10.33795/jraam.v6i2.004

Radicic, D., & Petković, S. (2023). Impact of digitalization on technological innovations in small and medium-sized enterprises (SMEs). *Technological Forecasting and Social Change*, 191, 122474. https://doi.org/10.1016/j.techfore.2023.122474

Rahmawati, E., Nazaruddin, I., Widiastuti, H., Sofyani, H., & Kholid, A. W. N. (2024). Environmental accounting in public sector: systematic literature review. *Journal of Accounting and Investment*, *25*(1), 75–90. https://doi.org/10.18196/jai.v25i1.21344

Roy, J. (2016). Data, dialogue, and innovation: Opportunities and challenges for "Open Government" in Canada. *Journal of Innovation Management*, 4(1), 22-38. https://doi.org/10.24840/2183-0606_004.001_0004

Sait, M. A., & Ali, M. A. (2022). Assessing Brunei Darussalam Public and Private Sector Readiness Towards Big Data Application. *International Journal of Asian Business and Information Management (IJABIM)*, 13(2), 1-22. https://doi.org/10.4018/IJABIM.20220701.oa7

Shah, N., Irani, Z., & Sharif, A. M. (2017). Big data in an HR context: Exploring organizational change readiness, employee attitudes and behaviors. *Journal of Business Research*, 70, 366-378. https://doi.org/10.1016/j.jbusres.2016.08.010

Tilba, A. (2022). Appearance or substance of Stewardship and ESG reporting? The challenges of translating 'commitment' into tangible outcomes. *Sustainability Accounting, Management and Policy Journal*, 13(5), 1015-1032. https://doi.org/10.1108/SAMPJ-03-2021-0091

Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). *How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study*. International Journal of Production Economics, 165, 234–246. https://doi.org/10.1016/j.ijpe.2014.12.031

Wang, H., & Wu, Z. (2016). Exploring the Use of Big Data Technology in Agriculture: A Systematic Review. *IEEE Access*, 2, 1497-1506.

Weerakkody, V., Kapoor, K., Balta, M. E., Irani, Z., & Dwivedi, Y. K. (2017). Factors influencing user acceptance of public sector big open data. *Production planning & control*, 28(11-12), 891-905. https://doi.org/10.1080/09537287.2017.1336802

Wilbert, J. K., Durst, S., Aisenberg Ferenhof, H., & Selig, P. M. (2018). Unlearning at the individual level: An exploratory case study in a high power distance country. *Journal of Innovation Management*, 6(3), 17-39. https://doi.org/10.24840/2183-0606_006-003_0003

Xiao, Y. and Watson, M. (2019), "Guidance on conducting a systematic literature review", *Journal of Planning Education and Research*, Vol. 39 No. 1, pp. 93-112. https://doi.org/10.1177/0739456X17723971

Zhou, Y., & Yang, Y. (2021). *Big data and citizen engagement in public service*. *Government Information Quarterly*, 38(3), 101604.

Zikopoulos, P., Eaton, C., deRoos, D., Deutsch, T., & Lapis, G. (2017). Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data. *McGraw-Hill Osborne Media*.

About the Authors

Muhammad Choirul Huda (MCH.) – is a graduate student in the Master of Accounting Program, Faculty of Graduate Studies, Universitas Muhammadiyah Yogyakarta, Bantul, Indonesia. Email: m.c.huda25@gmail.com

Bambang Jatmiko (B.J.) – is a Professor of Accounting in the Master of Accounting Program, Faculty of Economics and Business, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia. His areas of expertise are Management Accounting and Public Sector Accounting. Email: bambang.jatmiko@umy.ac.id

Author Contributions

Conceptualisation, MCH. and B.J.; Methodology, MCH.; Investigation, MCH.; Analysis, MCH.; Original draft preparation, MCH.; Review and editing, MCH.; Visualization, MCH.; Supervision, B.J.

Conflicts of Interest

The author declares no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.