Review Article

International Journal of Advanced Computer Research, Vol 13(64)
ISSN (Print): 2249-7277 ISSN (Online): 2277-7970
http://dx.doi.org/10.19101/IJACR.2021.1152071

The barriers and prospects related to big data analytics implementation in public institutions: a systematic review analysis

Matendo Didas*
The University of Dodoma, 1 Benjamin Mkapa Road, 41218 Iyumbu, Dodoma, Tanzania

Received: 03-April-2023; Revised: 10-June-2023; Accepted: 13-June-2023

Abstract

Modern society has always relied on data, which is generated by individuals, businesses, and government entities. This data serves various citizen-centric purposes, including monitoring, weather forecasting, healthcare management, and disease prediction. Technological advancements have expanded the sources of data, allowing it to be produced from any device, anywhere, and in any format. However, the challenge lies in comprehending, managing, and effectively utilizing this vast data resource. Public organizations are known for generating significant amounts of data. The question arises: can this data be integrated with technology-generated data to create societal value? Yet, accessing and integrating data can be complex for public organizations and nonprofits, especially when crossing international borders, due to legal, cultural, and political considerations. Nevertheless, big data applications are making their way into public institutions, and their cumulative impact on big data analytics (BDA) has the potential to provide a competitive advantage for improved public service delivery. Despite the recent attention garnered by BDA, many BDA projects in public institutions fall short of expectations, primarily due to substantial capital investments that make their return on investment questionable. One key reason for these failures is a lack of understanding regarding the challenges and opportunities associated with BDA in the context of public institutions. This article aims to systematically review existing literature to provide comprehensive insights into the prospects and barriers of BDA in public institutions. This review paper employs a systematic literature review analysis (SLRA) to shed light on the application of state-of-the-art BDA barriers and prospects within public institutions. It draws upon existing works that provide perspectives and theoretical constructs while identifying barriers and prospects. The review underscores that BDA holds immense potential for supporting public institutions in harnessing big data for evidence-based public service delivery. While there are numerous potential benefits, including food security, knowledge management, and informed policy-making, among others, the review also highlights critical gaps that need attention to fully realize these merits. This study delves into the use of BDA systems in public institutions, addressing both opportunities and challenges in this context. Based on these findings, recommendations are offered for future directions.

Keywords

Big data analytics (BDA), Big data, Public institutions, Preferred report items for the systematic review and meta-analysis (PRISMA), SLRA.

1.Introduction

Undoubtedly, one of the primary drivers of digitization is big data analytics (BDA) [1]. Vast volumes of big data are frequently generated, often in real-time [2]. As Lampropoulos [3] explains, BDA plays a crucial role as the technological foundation for numerous cutting-edge advancements and technologies, including the Internet of Things (IoT), artificial intelligence (AI), autonomous driving (AD), and many others.

To enhance the citizen experience and boost institutional performance, public institutions are actively seeking suitable technologies to streamline their resources [4, 5]. One technology that has demonstrated a significant impact on the operations of public institutions, including decision support and healthcare improvement, is BDA [6, 7]. As new challenges emerge in the analysis, storage, sharing, transportation, and processing of massive datasets across public organizations, the concept of BDA is evolving. Factors such as digitalization, the widespread use of social media networks, and the proliferation of smart devices are contributing to the growth of these massive datasets (Pramanik and Bandyopadhyay [8], Supriyadi et al. [9]). However,

*Author for correspondence

the notion of BDA is not universally understood in the academic and business communities. There are varying definitions of big data and its associated analytics, and numerous attempts have been made to quantify big data, but academics remain divided.

Numerous publications, including those authored by Chen et al. [10], Fan et al. [11], and Xiang et al. [12], have explored the potential of BDA, particularly for private firms and online businesses. Therefore, it is essential to enhance our understanding of the role that BDA plays in public institutions and how these institutions can benefit from extensive datasets. Academics widely recognize the advantages of implementing BDA tools in public institutions. According to Vassakis et al. [13], active utilization of BDA can significantly boost the productivity of public institutions. As per Gamage [14], BDA has the potential to address numerous challenges faced by public institutions, including enhancing efficiency, transparency, and the well-being of the population, as noted by Moorthy et al. [15]. In fact, BDA is already assisting American public institutions in fulfilling their responsibilities in various domains, including health, security, agriculture, education, and law enforcement [16, 17]. The primary objective of this review paper is to identify common opportunities and associated challenges related to BDA in public institutions.

Organizations, both in established and emerging markets, can significantly enhance their competitive positions by collecting, storing, and analyzing all this data. Public institutions are no exception. Understanding governance in the digital era necessitates an exploration of the challenges and opportunities that public institutions must grapple with when adopting technologies to manage and analyze vast datasets. If public organizations can effectively implement BDA, there is reportedly a wealth of fantastic potential [18]. The repercussions of organizations not being able to utilize BDA effectively are bound to be profoundly negative [19].

Although the use of BDA in public institutions has gained more attention thanks to authors like Väyrynen et al. [20], Hossain et al. [21], and Shi [22], developing a comprehensive understanding of the common prospects and challenges in public institutions can further stimulate the utilization of the underutilized big data assets already present in these institutions. The primary objective of this paper was to provide a comprehensive understanding of what BDA can positively contribute to in public institutions, the obstacles hindering these potential benefits, and ultimately, to offer recommendations and a perspective for future actions.

To present the most thoroughly studied opportunities and barriers in the context of public institutions, this paper assessed both the quantity and quality of research efforts on BDA at both the international and national levels.

The current state of research on the examined topic was assessed through various criteria, including the distribution of publications by year, publishers, inclusion and exclusion criteria, and the methodologies employed in the reviewed studies. Notably, among the 132 publications analyzed here, which encompass the most recent literature on BDA and public institutions, 58 were published between 2015 and 2023, underscoring the relevance of the research question to this body of literature. The findings were further elaborated upon in Section 5, and a comprehensive framework for the utilization of BDA in institutions was established based on the insights and observations. Additionally, the study's limitations and conclusions were addressed.

The following review questions directed our review in accordance with the systematic literature review analysis (SLRA):
RQ1: What are the key barriers that BDA encounter in public institutions?
RQ2: What significant opportunities can BDA offer to public institutions?

The current study contributes in two key ways: it synthesizes scholarly works on BDA opportunities and barriers for public institutions and provides guidance for aspiring researchers by offering a framework for their own examination of the prospects and drawbacks of BDA in public institutions.

This paper is organized as follows for the remaining sections: In Section 2, a summary of the literature on big data, its characteristics, and related analytics, especially in public organizations, is provided. Section 3 contains a description of the study's methodology. Findings are presented in Section 4. A discussion of this study is presented in Section 5. It is concluded in Section 6.

2.Review of literature

Under this part, different related worked have been discussed.

The concept of big data

The term "Big Data" refers to vast, intricate, and challenging datasets that defy conventional processing techniques, such as relational database structures and data warehousing. The emergence of the "big data" concept is a response to the proliferation of data in our daily lives. The diverse characteristics of big data profoundly impact the efficacy of traditional warehousing when it comes to collecting, processing, monitoring, and analyzing it. Furthermore, as defined by Emmanuel and Stanier, the term "big data" denotes datasets that surpass the scale typically found in regular databases [23].

It is indisputable that big data plays a pivotal role in the digital sphere [24]. BDA is inherently linked to big data, hence the term "big data analytics." BDA focuses on the analysis of large datasets using technologies such as the Hadoop Distributed File System (HDFS) and MapReduce (MR), in addition to structured query language (NoSQL), Red Miner, Cassandra, among others. These BDA technologies and their potential in handling big data have recently garnered significant attention [25].

Through effective management and utilization of big data using BDA technologies, public institutions stand to benefit immensely, including the development of innovative indicators for economic, educational, infrastructure, health, social, and other domains [26–28]. Some authors provide a concise overview of what big data entails; according to Talaoui et al. [29], big data is a digital transformation brought about by the proliferation of numerous data sources that generate massive volumes of datasets at rapid or real-time speeds. Visco et al. [30] and Gul et al. [31] define big data as information volumes that are significantly larger than what conventional data management methods can handle. Factors driving the prevalence of big data include increased digitization levels, the shift from a data society to an information society, the evolution of social media networks, and the growing use of electronic devices [32, 33].

It is recommended that BDA should not only encompass big data but also receive significant attention from societies concerned with developing technologies to ensure the quality, security, and privacy of big data [34, 35].

While the potential benefits of BDA technologies are mentioned in industrial, online, and technological sectors [36, 37], scholars and academics concur that the prospects of BDA in public institutions are inevitable and cannot be overlooked. The practical applications, such as increased efficiency, enhanced transparency, and improved public welfare, can be achieved through the implementation of BDA in public institutions [38]. Key areas in the public sector that can potentially benefit from BDA applications include politics, health, education, disaster management, justice, national security, agriculture and food security, and poverty eradication programs [39].

Considering the above discussions, it is evident that there is a mixture of limited knowledge regarding the common understanding, prospects, and barriers associated with harnessing BDA in public institutions. This review conducts a SLRA to provide a comprehensive understanding of the critical aspects related to BDA barriers and prospects in public institutions. The SLRA technique was chosen for its suitability in conducting an exhaustive and careful review of the literature, adhering to predefined rules [40]. The SLRA also ensures that the selected literature is relevant to the study area. The findings of this paper indicate that, despite the significant applications of BDA in public institutions, there are critical barriers that must be addressed before, during, and after the implementation of BDA to realize the potential benefits.

Characteristics of big data

While the precise definition of the concept of big data appears to be lacking in the literature, many authors concur that big data represents a new and distinctive paradigm with a unique perspective on data characterized by three dimensions: volume, velocity, and variety [41]. From this delineation of big data, it is evident that the three Vs—volume, velocity, and variety—are mentioned in the definition of big data. However, despite the proliferation of other attributes associated with big data, such as value and veracity, among others, the valid Vs for scientifically defining the big data concept are volume, velocity, and variety [41, 42]. The other Vs are disregarded because they are found to be either non-compliant or partially unsuitable for scientifically defining the big data concept [41, 42].

Big data encompasses datasets produced in massive volumes with a wide range of characteristics. It assists analysts, investigators, and organizations in making informed decisions by employing various techniques such as statistics, predictive analytics, machine learning, data mining, deep machine learning analytics, and text analytics. BDA finds

applications in various domains, including financial crisis management, medical research, education, banking, natural language processing, and knowledge management. Due to its benefits, the academic community is increasingly utilizing BDA to address complex problems, such as identifying disease outbreaks. Eleven "V's" are used to summarize the characteristics of big data.

Volume (V) (size): The most prominent characteristic of big data is its sheer abundance, which is the essence of the "Big" in the big data concept [43]. Without this attribute, attention to BDA and associated use cases is unnecessary [44–51]. Furthermore, this dimension has led to the introduction of new units for measuring data. Data units like Petabytes, Exabytes, and Zettabytes are now used to describe the vast volumes of data (PB-EB-ZB, respectively). The total amount of data generated, collected, duplicated, and utilized globally was projected to reach 64.2 zettabytes by 2020. It is anticipated that the global data generation will surpass 180 zettabytes over the next five years, from now until 2025. The highest amount of data was generated and duplicated in 2020, driven by increased demand due to pandemic outbreaks, heightened use of home entertainment alternatives, and more people working and learning from home. Storage capacity is also on the rise. In 2021, only 2% of the data generated and consumed in 2020 was preserved, indicating that only a small portion of this newly created data is retained. With the rapid growth in data volume, it is estimated that installed storage capacity will increase proportionally, with a compound annual growth rate of 19.2% expected from 2020 to 2025. In 2020, installed storage capacity exceeded 6.7 zettabytes. The projected expansion of data in zettabytes [52] is illustrated in *Figure 1*.

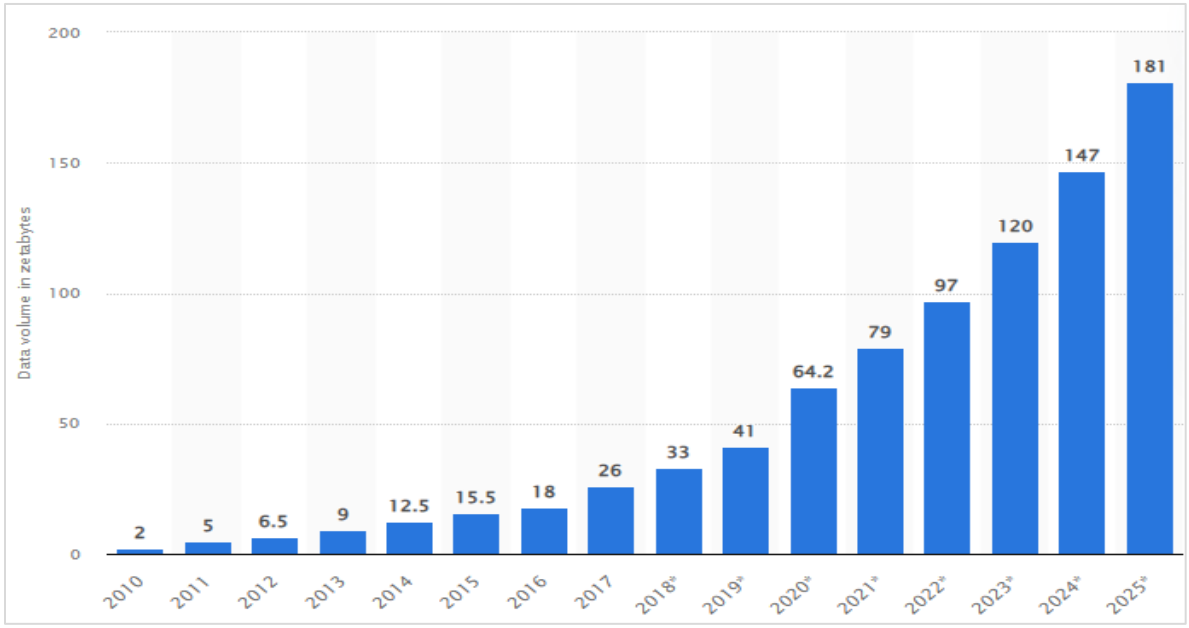

Figure 1 Volumes of data production forecast (zettabytes)

Velocity (v) (speed): BDA manages the substantial influx of data while simultaneously conducting real-time data analysis. This addresses the rapid generation, processing, and analysis of big data [46, 47]. The factors commonly contributing to this velocity (V) include, but are not limited to, car sensors, advancements in wearable technology, the proliferation of mobile records and applications, expanded capabilities of mobile technologies, and the increasing number and usage of social media networks, among others. These factors, among others, necessitate the development of technologies capable of handling the vast datasets generated in near real-time instances.

Variety (v) (complexity):

Text, images, audio, video, and sensor data represent only a handful of the numerous diverse types of organized, semi-structured, and unstructured data that comprise big data. This constitutes the third (3rd) vital "V" of big data, focusing on the significant increase in the multitude of information sources from which substantial volumes of datasets originate. This

increase in data variety is also attributed to the proliferation of pathways like those contributing to the velocity factor mentioned earlier [47, 48].

Veracity (V) (quality): This dimension emphasizes the need for data accuracy when dealing with the comprehensive and detailed information collected by big data.

Value (V) (knowledge): This pertains to the extensive provision of knowledge by the vast volumes of data regarding a subject under discussion, thereby highlighting the importance of ensuring the usefulness of the data.

Variability (V) (flexibility): This dimension involves the dynamic and ever-evolving nature of data sources. Big data accommodates this dynamic nature through extensibility (the addition of new data fields) and scalability (expanding in size).

Validity (V): This dimension focuses on managing extensive data, governance, and data quality.

Venue (V): This dimension centers on distributed heterogeneous data from multiple platforms.

Vocabulary (V): This dimension addresses data models and data structure semantics.

Vagueness (V): This pertains to the ambiguity in the usage of big data and the unclear meanings of tools.

Valence (V) (connectedness): This dimension concerns the utilization of big data to link various datasets by connecting common fields (*Figure 2*).

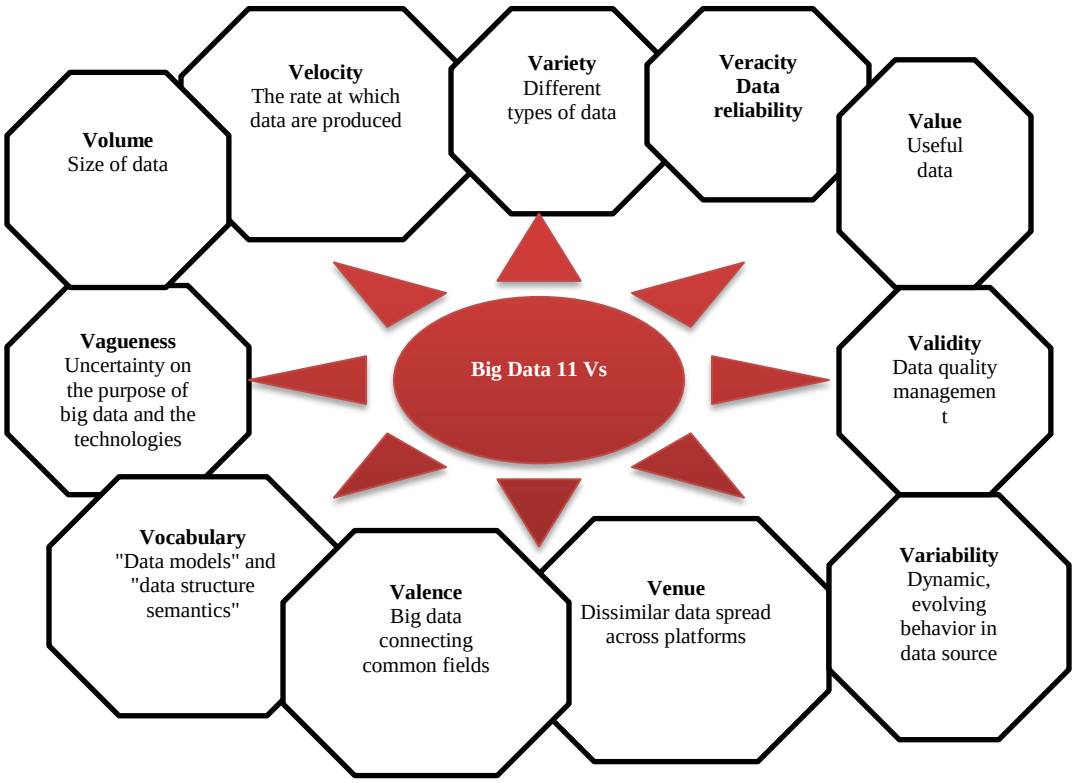

Figure 2 The graphical representation of the 11 Vs of big data

Big data analytics in public institutions

The concept of BDA encompasses various data-oriented methodologies for consolidating and analyzing massive datasets [49, 50]. The term "big data" is commonly used to refer to the vast amount of data items flowing across multiple systems [51, 52]. Despite being labeled as "big" data, its significance lies in its ability to capture even the smallest pieces of information relevant to the research topic [53]. The transformative decade of the 2000s paved the way for the emergence of "Big" data, which can be categorized into graphical, forecasting, prescriptive, social networking, and statistical analysis, primarily due to the amplification of the "3Vs" (volume, velocity, and variety) of information. In contrast, the task of applying BDA technologies to analyze and explore big data to uncover hidden meanings, insights, patterns, correlations, and new market trends, among other things, is simply referred to as BDA [54, 55].

There are various forms of BDA that public institutions can employ to enhance the delivery of public services. These analytics include, but are not limited to, text analytics, predictive analytics,

prescriptive analytics, diagnostic analytics, language analytics, descriptive analytics, and sentiment analytics, among others [55].

Despite the long history of public organizations preserving and managing data for administrative, legal, and practical purposes, this history primarily focuses on traditional survey and secondary data [56]. However, the datasets generated from contemporary platforms such as the web, email texts, opinions, online social networks, mobile devices, sensors, blogs, wikis, satellites, machine-generated data, among others, have shifted the focus of public institutions from traditional data management approaches to modern BDA solutions.

Public institutions must tap into the vast volumes of data generated from various sources at high speeds if they are to remain competitive and relevant to their citizens [57]. For those unable to harness the advantages of BDA, the consequences can be quite severe [57]. However, leveraging and harnessing emerging big data and associated analytics technologies extend beyond the capabilities of conventional data management approaches in public institutions [56, 57]. This is because traditional data management approaches are ill-equipped to handle the high-volume data generated rapidly from multiple sources. As mentioned in the introduction, one of the public sector domains with significant potential benefits from BDA is agriculture and food security. According to Mishra et al. [57], when appropriately applied, BDA can contribute to addressing the world's food shortage, a solution they refer to as "BDA to feed the world." Furthermore, Mishra et al. [57] connect this solution to other remedies, such as reducing both natural and man-made disasters, which can include landslides, earthquakes, and terrorism, among others.

In conclusion, public institutions today stand to gain significantly from successfully implementing BDA in their routine business operations. As the volume of big data continues to grow in various contexts, including healthcare, one of the public sectors [58, 59], the prospects of BDA also expand. Heterogeneous and multispectral observations in the healthcare sector, such as patient demographics, treatment histories [60, 61], and diagnostic reports [62], generate substantial big data. According to Singh et al. [63], this data can be either unstructured (e.g., clinical notes, prescriptions, or medical imaging) or structured (e.g., genotype, phenotype, or genomics data). To make effective use of healthcare data, real-time, high-quality data must be consistently generated and collected [64]. Informed decisions can be made by healthcare organization decision-makers based on insightful conclusions drawn from big data [65]. Healthcare businesses leverage technology to adapt to the evolving nature of big data [66]. BDA in healthcare can also facilitate the integration of various disciplines to comprehensively analyze medical conditions [67].

Current challenges of BDA in public institutions:

To summarize, all the attributes of big data mentioned earlier can be observed within the context of public institutions. However, there are various barriers (as listed in *Table 1*) associated with implementing BDA in public institutions, which can hinder the successful realization of BDA's potential benefits for these institutions [68, 69].

Despite the potential benefits of BDA in public institutions, BDA projects continue to struggle with performance issues [72]. Approximately 85% of BDA projects fall short of achieving their intended transformation [72, 73]. Additionally, market research and advisory firm Gartner projected that between 2016 and 2018, approximately 60% of BDA projects may either stall after the pilot and experimentation stage or be abandoned. According to Gartner, faulty BDA deployments could contribute to up to 50% of ethical business violations by 2018. Many public institutions have not even started working on BDA projects that progress beyond the conceptual phase.

A study by the University of Munich in 2015, which examined the development of big data initiatives in twenty-five organizations, revealed that only one in five organizations had BDA projects that had reached the deployment stage. Four out of twenty-five were still in the planning stage, while the remaining organizations had not advanced to any other stage [73]. Given this background, this paper aims to contribute to the improvement of BDA project performance in public institutions by offering a comprehensive understanding of the challenges and prospects associated with successful BDA implementations in public institutions. It will also provide actionable takeaways and recommendations for prospects.

Table 1 Common challenges related to BDA

Challenge	Description
Starting capital	For institutions providing public services, the initial costs associated with establishing the required infrastructure to harness the benefits of big data are exceedingly high [70].
Information quality	The organization's efforts to collect big data may be hindered by staff members who lack adequate training and institutional habits that resist change. The potential consequence of poor-quality heterogeneous data is that it can lead to inaccurate interpretations and recommendations [70, 71].
Security and confidentiality	The exposure of patients' personal information to unauthorized data access during inter-system exchanges raises privacy and security concerns, especially in the healthcare sector. In the broader context, this information is typically referred to as personal identification information (PII).

3.Research methodology

Existing approaches

There are five primary types of literature reviews, each serving a distinct purpose:

Traditional or narrative literature review aims to examine and summarize a body of literature. It involves a comprehensive review of relevant literature to identify new areas of inquiry, knowledge gaps, or discrepancies [74, 75]. In contrast to traditional reviews, systematic reviews take a more rigorous approach. They are often employed for well-structured and specific research objectives, following strict guidelines for review [74, 75]. Meta-analysis literature review employs statistical methods to analyze conclusions drawn from selected literature. Meta-synthesis literature review is a non-statistical method that evaluates and analyzes findings from qualitative studies, aiming to advance previous conceptualizations and interpretations [74, 75]. Scoping literature review is a preliminary assessment of the scope and size of research literature, often involving ongoing research. It seeks to identify the type and extent of research evidence. The preferred reporting items for systematic reviews and meta-analysis (PRISMA) statement can guide scoping reviews [74, 75].

For this study, a SLRA approach was chosen [76]. The study adhered to the PRISMA statement and the SALSA framework (Search, Appraisal, Synthesis, and Analysis) to conduct the review. PRISMA is a widely recognized guideline for reporting evidence in systematic reviews and meta-analyses, ensuring methodological accuracy and reproducibility [77]. SALSA helped determine the search standards for the SLRA, ensuring methodological accuracy, systematization, thoroughness, and reproducibility in the review process.

Table 2 presents the summarized findings, advantages, and limitations of the selected approach, supported by the analysis of twenty-six papers that have employed this chosen approach.

Table 2 Review of the papers to identify the approach's findings, advantages, and limitations

No.	Author	Method/ approach	Findings	Advantages of the method used	Limitations
1	Zainal et al. [78]	Systematic review	Issues and difficulties with government-identified big data initiatives	Thorough screening of pertinent literature using a predetermined set of standards.	The SLRA's constrained scope and mandated procedures prevent full coverage.
2	Al-Sai and Abdullah [79]	Systematic review	Identified general BDA impacts and challenges	Is regarded as original work because it was carried out using a strict, organized procedure.	Not as appropriate for students or new academics who are short on time and resources.
3	Kamilaris, et al. [80]	Systematic review & meta-analysis	Identified practices of big data analysis in agriculture	Incorporate a summary or statistical breakdown of the findings from each individual study; be executed in a way that can be replicated, yielding identical results with identical information.	Every step of the review, including the search, must be thoroughly documented to ensure reproducibility.
4	Desai [81]	Systematic review and descriptive analysis	Stated big data applications and challenges	It is explicit and systematic.	It is highly structured.
5	Mureddu et al.	Systematic	Research issues for using big data in	Using explicit techniques to	Time and resource intensive

No.	Author	Method/ approach	Findings	Advantages of the method used	Limitations
	[82]	review	policy-making have been identified.	systematically explore, evaluate, and synthesize literature on a certain topic.	requiring experience.
6	Praharaj [83]	Systematic examination	Difficulties with massive data command and control centers for smart cities.	• Eliminate reviewer bias by using objective, repeatable criteria to choose pertinent individual publications and evaluate their validity.	Broader coverage is exchanged for more explicit approaches when compared to a standard narrative evaluation.
7	Cannataci et al. [84]	Systematic review	Examined Legal challenges of big data	Provide a statistical breakdown or meta-analysis of the findings from each individual study.	Time and resource intensive requiring experience.
8	Tse et al.[85]	Literature review and content analysis	Identifies challenges of big data governance in healthcare	Provide a statistical breakdown or meta-analysis of the findings from each individual study.	Time and resource intensive requiring experience.
9	Pradhan and Shakya [86]	Systematic review	They discover the possibilities of digitalizing e-government services through creative problem-solving ways by modifying the big data application.	When outcomes from multiple research studies are combined, an average is produced that is more accurate than the findings from each individual study. As a result, the SLR uses explicit, transparent methodologies that are presented explicitly and are repeatable by others.	Time and resource intensive requiring experience.
10	Albqowr et al. [87]	Systematic review	Gives an overview of the advantages that BDA may have for public supply chain management.	Concentrates on promoting research knowledge.	All review procedures, including the search, must be documented to ensure reproducibility.
11	Shastri and Deshpande [88]	Systematic review	This article provides a brief description of big data and its applications in both the public and healthcare sectors.	Rigorous search designed to locate all eligible studies.	N/A
12	Younas [89]	Systematic review	Gives an overview of a couple of the research issues related to BDA in the public sector.	Transparent, reproducible methods.	For reproducibility, each stage of the review, including the search, must be documented.
13	Kondraganti, et al. [90]	Systematic review	Explains BDA tics in detail in the context of disaster relief and humanitarian help.	Requires clear aims with predetermined eligibility and relevance criteria for studies.	Time and resource intensive requiring experience.
14	Eachempati, and Srivastava [91]	Systematic review	Identifies some emerging sectors that apply big-data analytics including public sector	Accurate documentation of the method implies.	Demands accurate efforts.
15	Candra et al. [92]	Systematic review	Identifies BDA challenges in cloud computing	It combines narrative, quantitative, or qualitative synthesis.	More effort is required for studying analyses.
16	Pappas et al.[93]	Systematic review	Outlines techniques by which BDA can be used to improve performance in the digital age and the development of ecologically sound societies.	Conclusions are drawn based on evidence.	A broader coverage is traded for more explicit approaches when compared to a standard narrative review.
17	Kolajo et al.[94]	Systematic review	Explore the latest recent breakthroughs in big data streaming tools and technologies, as well as the international techniques and approaches used to analyze enormous data streams.	A detailed and comprehensive search strategy is developed.	A trade-off occurs, exchanging broader coverage for more explicit approaches when compared to a standard narrative evaluation.
18	Mikalef et al. [95]	Systematic review	Highlights strategies and procedures by which big data might improve the bottom line for public firms.	The literature is specifically related to the subject being studied.	It is time and resource-intensive, requiring expertise.
19	Vanani and Majidian [96]	Systematic review	Identifies BDA methods in institutions	Results can be reproduced, and reviews can be periodically updated to	To ensure reproducibility, every stage of the review, including the search

No.	Author	Method/ approach	Findings	Advantages of the method used	Limitations
				incorporate new evidence.	process, must be thoroughly documented.
20	Aggarwal [97]	Descriptive and Literature review	Explains the benefits and drawbacks of adopting big data in the public sector.	The material picked is solely pertinent to the subject under consideration.	N/A
21	Morabito and Morabito [98]	Case studies	The utilization of (a) new data sources, such as crowd funding and the IoT, (b) public talent participation, (c) institutionalizing private-public partnerships and (d) additional ways were explored.	Often fast, focused, scientific and relatable	Because the results depend on numerical responses, the insights into beliefs, motivations, and factors guiding the investigation may be limited.
22	Batko and Ślęzak [99]	Systematic review and questionnaire	Identifies benefits of BDA to public medical facilities	Thorough screening of pertinent literature using a predetermined set of standards.	Time and resource intensive requiring experience.
23	Verma [100]	Literature review	Looked into the use of sentiment evaluation to increase innovation, accountability, citizen participation, and public service efficiency.	Provide researchers with new evidence to support psychological theories	Broader coverage is exchanged for more explicit approaches when compared to a standard narrative evaluation.
24	Rafiq et al. [101]	Qualitative and quantitative research and case study	Identified privacy and security concerns in big data in institutions	Often fast, focused, scientific and relatable	Limited coverage and results can be difficult to replicate
25	Lazarevska et al.[102]	Systematic review	A comprehensive deployment of BDA as one of the most inventive technological advances being employed in public sector assessments.	Often fast, focused, scientific and relatable Cover author's expertise	Preconceived notions or biases and may overestimate the value of some studies.
26	Jin et al. [103]	Using literature and the research situation as sources, the positivist research paradigm,	Investigated how BDA affected the efficiency of government agencies.	Handle a wide range of situations in just a brief amount of time.	Requires researcher's expertise.

We conducted a meticulous examination of the relationships among big data, BDA, and associated concerns and opportunities within public institutions. Our focus was on scholarly literature published from 2015 to 2023. This systematic review aimed to uncover common themes and conclusions in the discussions surrounding the advantages and disadvantages of BDA in the mentioned institutional contexts.

The process began with the formulation of research questions and the rational selection of keywords and databases. We conducted both manual and internet searches to identify pertinent papers. Subsequently, we developed criteria for inclusion and exclusion. The third phase involved assessing the quality of the identified literature and its relevance to our investigation. The fourth step entailed a comprehensive review of the substance of the selected studies, followed by the analysis and formulation of conclusions.

Selecting the appropriate databases and keywords constituted a crucial aspect of our systematic literature review. We engaged in discussions and brainstorming sessions with subject matter experts to ensure that our chosen terms effectively covered both the core domain of BDA and public institutions. BDA and public institutions emerged as the two primary spheres of interest for our review, and we combined these keywords to conduct our search for relevant publications.

To obtain the most pertinent articles for our study, we meticulously followed pre-established exclusion and inclusion criteria (as detailed in *Table 5*). We conducted the review globally, without geographical restrictions, to encompass the widest possible range of literature. We also searched multiple scholarly

databases, as reflected in the search result count in *Table 3*, to ensure comprehensive coverage.

Our systematic literature review process adhered to the four detailed steps of PRISMA as the paper selection strategy. These steps encompassed article identification, screening, eligibility determination, and inclusion, ultimately resulting in the selection of pertinent papers for our analysis.

The list of various scholarly databases was carefully chosen, and their search results are documented in *Table 3*. Our SLRA process adhered to the four distinct steps outlined by PRISMA in a comprehensive manner, serving as our paper selection strategy.

These are as under:

Step 1: Article identification - In this initial stage of the process, we outline the plans and procedures used to identify relevant databases and knowledge repositories. We describe the key terms and search tactics employed to locate these resources. For this study, this stage marks the beginning of our SLRA, involving comprehensive searches of article abstracts and citations in specific scholarly databases. Keywords and relevant scholarly databases were logically defined and justified. The selection of keywords and keyword combinations, as well as the choice of databases, resulted from consultations with experts from both academic and industrial backgrounds, extensive literature review, and critical thinking. A total of 1,176 articles were discovered for screening through the quest for identified keywords/keyword combinations (*Table 4*). Following the completion of the search, all retrieved records were thoroughly checked for duplicates.

Step 2: Article screening - During this stage, each article's title, abstract, or summary, along with its keywords, were carefully reviewed. The objective was to determine whether each selected article contained content suitable for the study subject. We aimed to establish exclusion and inclusion criteria based on references found on the web and through manual searches, and to categorize and label suitable publications accordingly. Articles that fell under any of the following criteria were excluded from the subsequent eligibility step for content analysis and evaluation: non-English papers, editorial letters, legal conventions, interviews, non-academic texts, book reviews, and summaries of meetings or conferences. Articles eligible for content analysis and evaluation had to meet one of the following criteria: conference papers written in English, books, academic journal articles, empirical and case reports, editorials, subjects related to information technology and science, humanities, theoretical studies, and BDA interventions in fields such as agriculture science, technical science, and health science.

Step 3: Determining eligibility of the article - In this step, we rigorously coded the content of the reviewed papers to facilitate the summary and comparison of results. After thorough examination, articles that did not meet the eligibility criteria were excluded from further consideration. Additionally, titles, abstracts, and summaries were screened for relevance to the selected keywords. The primary focus of this step was to identify articles of high eligibility, suitability, and quality for addressing the study's objectives. Quality assessment of each article was integral to this step, which involved determining its relevance to the study.

Step 4: Included - This final stage encompasses the key findings, both numerical and descriptive, that were frequently present in the reviewed papers. We estimated the extent of challenges and promise associated with BDA in public organizations based on the shared conclusions presented by researchers whose work was evaluated. Following the eligibility determination, which included the exclusion of irrelevant articles, we defined the number of articles eligible for inclusion in the result review for content analysis. A total of seventy-two articles were removed and set aside following the exclusion rules, as represented in *Figure 3*. Additionally, two articles were excluded during the content analysis phase. The final selection of articles for the systematic review comprised eight papers obtained from various conference papers.

Indeed, as previously described, the four stages of the systematic review process - identification, screening, eligibility, and inclusion - were executed. In the PRISMA-style representation of the paper selection technique illustrated in *Figure 3*, the results of each of these stages are summarized. Initially, a total of 1,176 research publications were identified from the selected databases. Subsequently, 1,044 papers underwent a basic screening process, which involved applying filters to limit the year range and implementing predetermined exclusion criteria to eliminate certain papers. As a result, 72 papers were excluded, leaving the researcher with 60 papers to review. This more comprehensive evaluation included an assessment of each paper's abstract to confirm compliance with the research's eligibility criteria. Finally, 58 papers met all eligibility requirements and were included in the final

evaluation after the full texts of each paper were reviewed once more. Consequently, the systematic review's final sample selection consisted of 58 papers, as depicted in *Figure 3*.

Table 3 The paper distribution from various publishers, indicating the number of search results

Database	Search results Frequency (%)	Percentage (%)
EBSCO Host	104	8.84
Web of Science	110	9.35
Springer Link	240	20.40
ACM digital library	147	12.51
IEEE Explore	247	21.00
Scopus	140	11.90
Science Direct	188	16.00
Total	**1176**	**100**

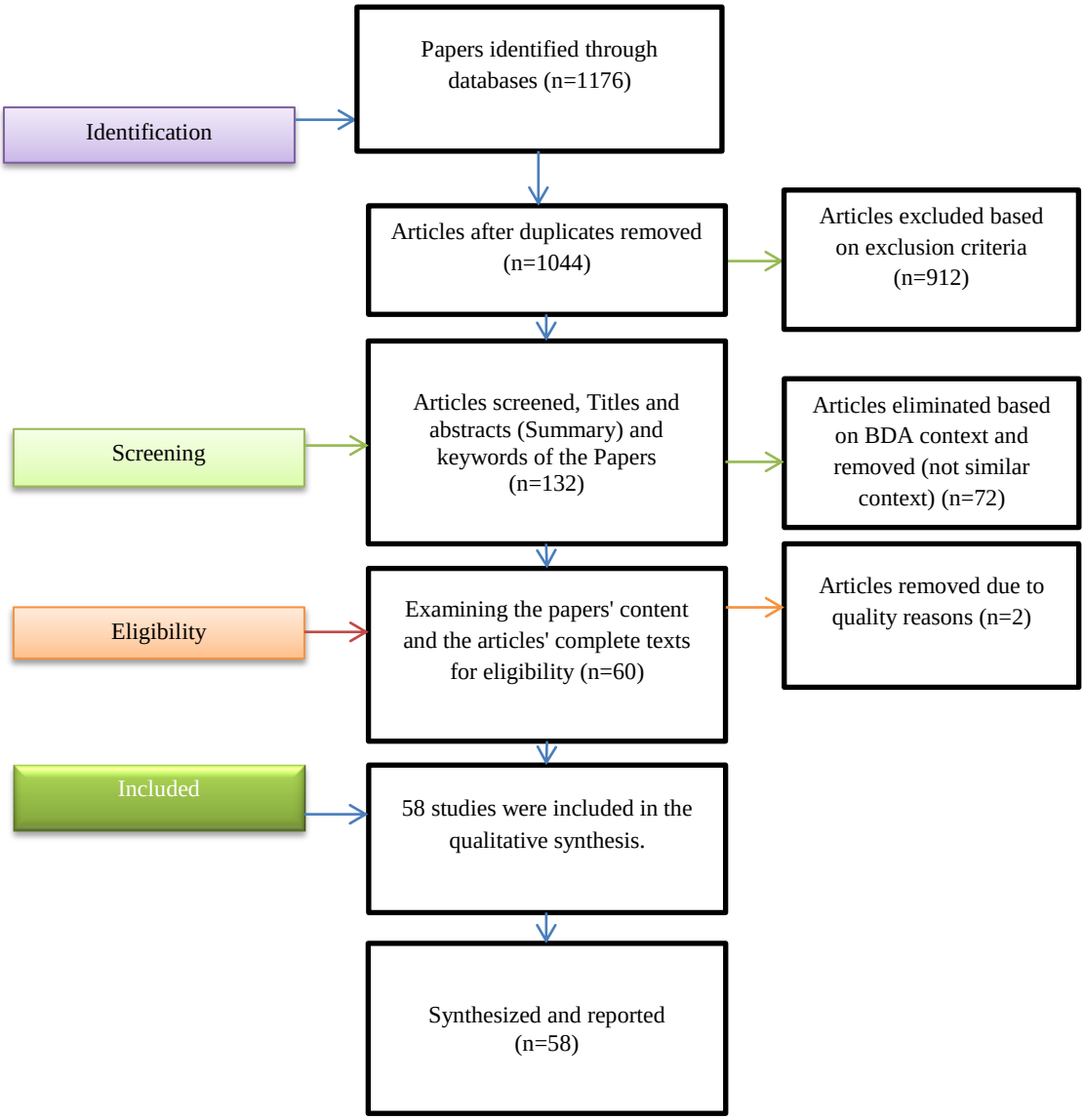

Figure 3 The PRISMA style representation for showing the paper selection strategy

Table 4 The keywords/combination of keywords applied for search

BDA
BDA + Public Institutions
Big data management
Big data

Table 5 The study inclusion and exclusion criteria

Inclusion	**Exclusion**
Conference papers, English language articles, journals	Non-academic text
Practical Studies, Books	Interviews
Fields information science and technology, humanities and Theoretical studies	Book reviews
Case studies and Editorials	Non-English
BDA papers related to various domains in public sector like health	

Year-wise distribution of the selected papers: *Figure 4* illustrates a particularly noticeable trend, with the number of publications in the target subject increasing since 2015. In relation to the two questions, the outcomes of the examination of the content of the fifty-eight chosen articles are discussed in the following sections.

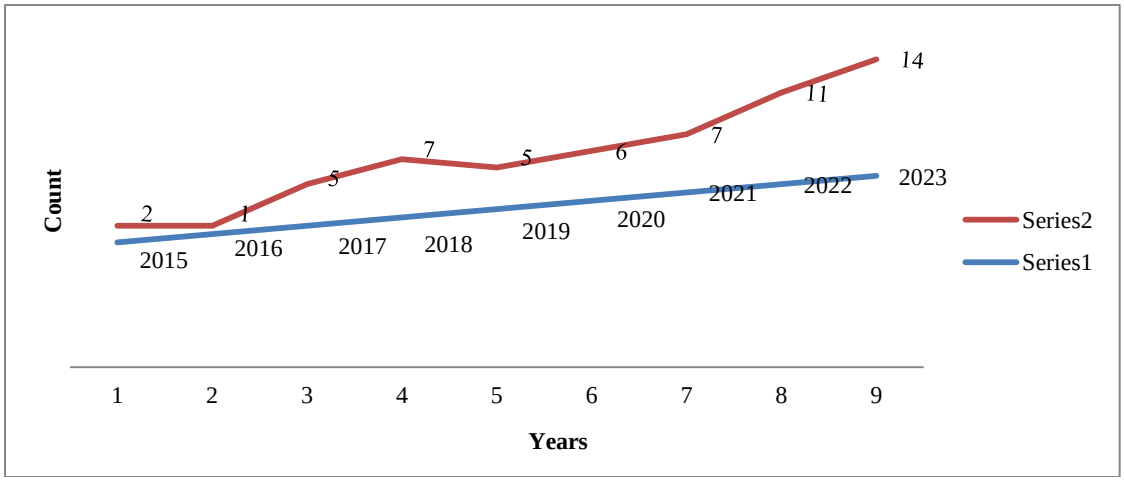

Figure 4 The year-wise distribution of the selected papers

Year-wise count of papers along with their categorization: *Figure 5* reflects the year-wise count of selected papers categorized into journals and books from 2015 to 2023.

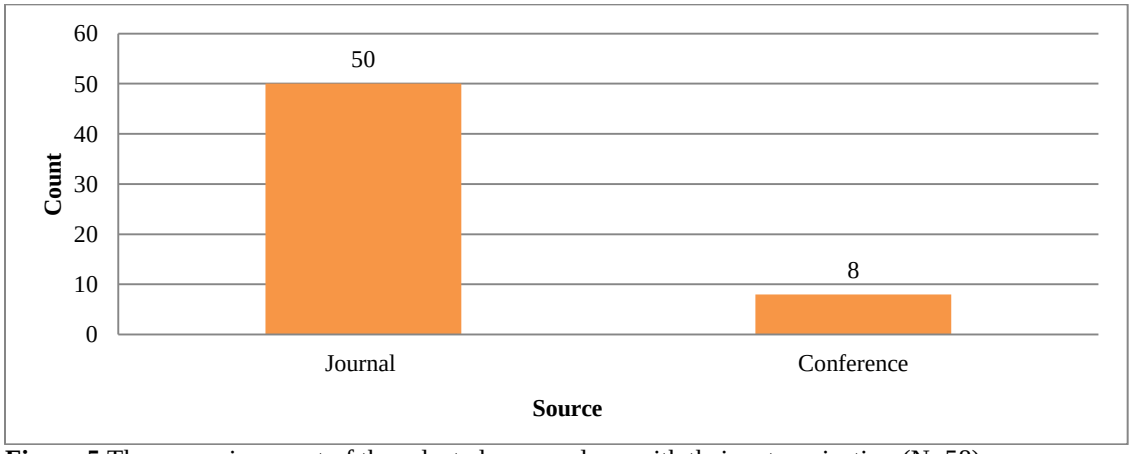

Figure 5 The year-wise count of the selected papers along with their categorization (N=58)

4. Results

The review's findings are reported in this section. This investigation led to the selection of 58 articles, including papers from conferences and journal articles. Out of the pool of articles, 8.84% were from EBSCO Host, 9.35% from Web of Science, 20.40% from Springer Link, 12.51% from ACM digital library, 21.00% from IEEE Explore, 11.90% from Scopus, and 16.00% from Science Direct. Among these, there were 50 journal articles and 8 conference proceedings papers. All of the selected articles were published between 2015 and 2023. After removing duplicates, 132 publications were evaluated for eligibility using title, abstract, and full-text criteria, and 58 research studies were ultimately included in the review. The complete research selection process is depicted in the PRISMA flow diagram shown in *Figure 3*. The study's findings are explained as follows:

Big data analytics barriers in public institutions

The most cited challenges encountered during the study are summarized in *Table 1*. These challenges were identified through the application of our methodology and by reading the adapted literature articles. The explanations provided here are within the context of public institutions, as indicated by the titles of the papers in our domain scope. Therefore, the explanations of the barriers are tailored to the setting of public institutions. However, it's important to note that the same challenges may also be encountered in other environments beyond the domain scope of this study. The challenges are categorized, and examples of specific challenges are provided under each respective category.

Infrastructure and management barrier category

The systematic literature review found numerous infrastructure and management challenges of BDA in public institutions. These included the following:

Lack of innovative and business-driven approach for big data implementation: One of the critical challenges hindering the realization of BDA potentials in public institutions is the absence of innovative and business-driven approaches to implementing BDA technologies [103]. A significant failure rate of BDA projects in public institutions can be attributed to the lack of a proper implementation approach [104, 105]. Many government agencies have yet to initiate BDA initiatives that can progress beyond the Proof-of-Concept Stage, causing them to miss out on benefits such as cost reduction in public service delivery and efficient resource utilization in sectors like healthcare [105, 106]. The current methodologies used in these initiatives, such as the National Institute of Science and Technology and the Nirmala and Raj References Framework, among others, are ill-suited for the unique business processes of public institutions [106, 107]. While these approaches have gained acceptance in technical, online, and industrial corporations due to their technology-focused nature, they still suffer from critical weaknesses when applied to public institutions, such as a lack of business data requirements based on business processes, an excessive focus on technology infrastructure before addressing business process requirements, resulting in capital-intensive projects that are challenging to generalize because they are product-dependent.

Organizational alignment: The concept of organizational alignment emerged as a key concern in the literature review, questioning whether the use of big data aligns with the structure, primary functions, and overall strategy of the respective public institutions. It emphasizes the need for projects involving big data to align with the organization in which they are implemented.

Poor information and communication technology (ICT) infrastructure: Issues such as slow network connectivity are associated with this barrier [107–109]. This encompasses various aspects, including computer hardware platforms, operating systems, business software, data storage, networking and telecommunications, internet connectivity, consulting services for system integration, and more [110, 111]. Information and communication technology infrastructure was assessed to determine if it met the characteristics required for the specific BDA application types. Evaluating the alignment between the present information technology system in public institutions and the IT network necessary for the chosen big data service category was deemed essential. It aimed to gauge the agreement between the current situation and a potential BDA-based future scenario.

Inadequate scalability of legacy storage systems: Scalable database systems are essential for designing, developing, implementing, modifying, and reforming public institutions through BDA. Public organizations have grappled with the limited scalability of outdated database systems. Big data introduces changing data volumes with dynamic data formats from diverse sources at varying velocities, posing challenges to the effective application of big data and associated analyses [110, 111]. A consequence of inadequate scalability in legacy data storage systems is an added burden on the data visualization and presentation process, a crucial step in the data management pipeline. Visualization and presentation tools must be configured to ensure that

data processing and analysis results can be easily visualized and presented to the target audience. Otherwise, public institutions may accumulate, process, and analyze vast amounts of data without effectively communicating the outcomes to their intended audience. The literature review emphasized that neglecting scalability issues can result in difficulties in deducing and comprehending data processing and analytics results.

Inaccurate interpretation of processing and analytics results: The accurate interpretation of data analysis results is crucial for deriving meaningful knowledge and effectively utilizing it [112].

BDA for data processing challenge: The review also identified that the existing infrastructure and management approach is inadequate for supporting and handling big data, particularly during the third phase of the data management pipeline, which involves "data processing and analytics" [113, 114]. Modern analytics systems, such as BDA, capable of managing the entire spectrum and dynamics of generated data are essential for processing vast volumes of rapidly incoming data from various sources [114–116]. While there are proposed solutions for some of the challenges posed by BDA, such as emerging innovations in the fourth industrial revolution (4thIR), these technologies are advancing faster than their adoption [117, 118].

Big data storage and retrieval: Infrastructure and management constraints are also associated with data storage and retrieval issues, as indicated by the research review [119, 120]. The challenges arise from the diverse sources and high volumes of big data, which can hinder the generality of conclusions, whether predictive or prescriptive, and create bottlenecks in infrastructure and management within public institutions [120]. To effectively handle the increasing frequency of big data, public institutions must adapt, implement, or develop analytics technologies capable of managing the growing data volumes and sources in the 4thIR era [121].

Integration of new big data technology: Finding an approach that combines new big data technology with existing traditional information technology platforms is an example of how the infrastructure and management barrier category intersects with information technology itself. It is also related to the challenge of integrating disparate data sets. The three valid Vs of big data—quantity, speed, and variety—are causally linked to the continuum of infrastructural and management challenges in BDA within public institutions, as evident from the category of infrastructural and management issues [122, 123]. It is strongly recommended that a business-driven approach, along with a thorough understanding of the barriers and opportunities related to BDA in public institutions, is essential to improving the delivery of public services over time [122, 123].

Data quality challenge: Another category of challenges for BDA in public institutions relates to data quality. This category encompasses issues such as missing data, inconsistent data, data inaccuracies, and duplications, resulting in unnecessary redundancy and inefficient resource utilization [124–126]. Data quality is crucial for obtaining accurate and clean data suitable for the intended purpose and evidence-driven decision-making and policy formulation. As highlighted in the introduction, one of the public domains benefiting from BDA is healthcare, where human lives are at stake. Poor-quality datasets used to make decisions regarding patient treatment can lead to catastrophic loss of life.

Organizational challenge: During the study, organizational barriers were identified as features of the organizational framework and culture that were incompatible with modern technologies. These barriers included issues with the flow of authority, challenges in aligning goals and expectations between IT and business departments, and resistance to change rooted in the organization's historical operations. Typical symptoms included an inability to recognize the strategic advantages of investing in BDA and a lack of organizational cohesion and collaboration due to organizational fragmentation [127].

Human Barrier Category: These challenges can be attributed to employees lacking data analysis abilities [128]. The human challenge category was also identified as one of the barriers related to BDA in public institutions, encompassing two aspects: a lack of skills and a shortage of training opportunities for dealing with BDA in public institutions [128, 129]. The adoption of BDA and the management of big data demand highly specialized data analytics skills and ongoing training in BDA [129]. End-user proficiency in evaluating and interpreting analytical results from BDA should not be overlooked [130, 131].

E-Ethics, security, and privacy barrier category: Finally, the ethics, security, and privacy barrier category were found to be a significant hurdle if public institutions are to successfully implement BDA in their business processes. This category encompasses atomic barriers such as data sharing, the digital divide in big data technology, ownership, privacy protection, and access controls. Ethics, security, and privacy issues were found to be closely related to the volume and multiple sources of big data

[132, 133]. The literature review analysis recommended that key stakeholders in public institutions should strive to enact approaches and frameworks focused on streamlining data exchange, enhancing access control, bridging the big data digital divide, and ensuring privacy protection [134]. Without public institutions openly sharing data with the public, there can be no trust and confidence in public institutions' business processes and procedures, resulting in a lack of transparency and accountability in government operations as perceived by citizens [134–136]. This category of obstacles can also pertain to privacy and the security of individual data both inside and outside of organizations.

In conclusion, the overall barriers to BDA implementation in public institutions can be summarized as follows: the most current and critical barriers are associated with infrastructure and management approaches aligned with the business processes of public institutions. While ethics, security, and privacy issues are also important considerations for public institutions when implementing BDA, a significant portion of these issues is predominantly linked to human factors [136]. Research should focus particularly on the area of infrastructure and management hurdles, as well as on the ethical, security, and privacy issues related to the use of BDA in public organizations.

BDA presents a range of critical ethical challenges in addition to technological ones. Given the importance of data analysis in governing society, ethical awareness is crucial. Therefore, the rapid advancement of information technologies underscores the need for further research in the field of ethics and information technology within public institutions. Future research should address privacy and security vulnerabilities associated with big data, including those related to international legal and regulatory disparities. *Table 6* provides a summary of categories and atomic barriers to BDA implementation in public institutions.

Table 6 The summary of categories and atomic barriers of BDA implementation in public institutions

	Barrier Category	Atomic Barrier
A	**Infrastructure and Management**	Lack of innovative and business-driven approach for big data implementation
		Poor information communication and technology infrastructure
		Inadequate scalability of legacy storage systems
		Data visualization and accessibility
		Results of incorrect data analytics as well as interpretation
		Large-scale data processing puzzle solved by BDA
		Big data storage and retrieval
B	**Data Quality**	Missing data
		Inconsistent data
		Inaccurate data
		Duplicate & redundancy of data
C	**Organizational**	Authority flows
		Communication
		Goal and expectation alignment between IT and business departments is challenging
		Resistance
D	**Human**	Lack of Skills
		Lack of training avenues for BDA
E	**Ethics, Security, and Privacy**	Data Sharing
		Privacy Protection
		Ownership
		Accessibility
		Big data digital divide

Prospects for BDA in public institutions

The benefits that BDA can offer to public institutions have become evident after examining the findings of the research under discussion. According to the potential highlighted by the systematic literature review, BDA has much to offer when applied and adapted appropriately to the business operations of public organizations. The analysis revealed that the benefits that can be derived from BDA in public institutions can be categorized into seven themes (*Figure 6*), as follows:

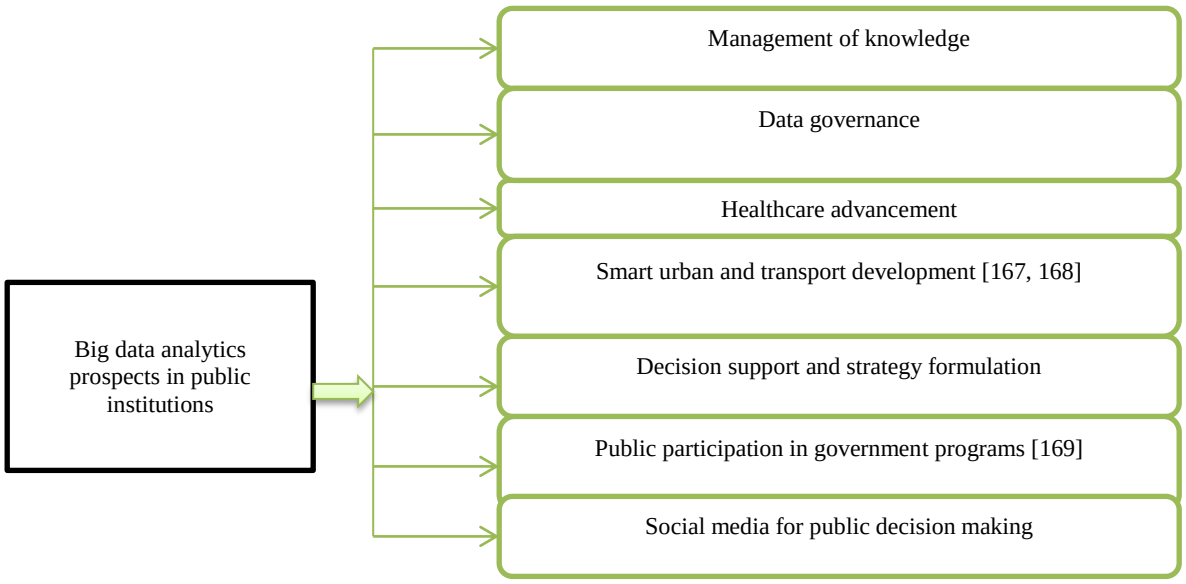

Figure 6 The reviewed prospects of BDA in public institutions

First, social media for public decision-making: BDA can enable public institutions to make informed decisions based on interactions on social media networks by citizens [136, 137]. The vast amount of information people shares and communicate through their social media accounts can provide valuable insights for sentiment analysis of citizens [137–140]. According to [138, 141], by applying data mining techniques to social media data, challenges ranging from poor resource management to financial misuse can be addressed quickly [141, 142].

Secondly, knowledge management: BDA technologies can serve as a foundation for knowledge management, encompassing knowledge preservation and dissemination in any form. Public institutions are known for enacting policies and frameworks that promote knowledge creation [142, 143]. However, without advanced data analytics tools for preserving and disseminating this knowledge, the efforts put into its creation, along with the enacted policies and frameworks, can go to waste. With effective knowledge dissemination strategies using BDA techniques, public institutions can anticipate the need for appropriate policies before citizens demand them [143]. It's important to note that government datasets should be made available to assist public institutions in aligning with citizens' interests. It is the responsibility of public institutions to collect data from citizens safely and utilize it for the betterment of society, ranging from economic and health aspects to education and overall safety.

Thirdly, public participation in government programs: BDA can promote citizen participation in government planning and policy-making activities. Citizens are the focus and beneficiaries of public institutions' existence [144, 145]. Overlooking citizen participation and their roles in government planning and policy-making can render discussions about leveraging and harnessing BDA in public institutions futile, along with citizens' non-compliance with government programs [145]. Citizen participation in government activities is crucial for building trust and confidence in public institutions' processes.

Fourth, healthcare advancement: BDA technologies are known for advancing the healthcare sector in terms of both service quality and throughput. Big data supports subsystems such as life support systems, electronic health records, recommendation advisory systems, and surveillance support systems for tracking epidemic outbreaks via the internet, among others [146–153]. To fully harness the promising potential of BDA in public institutions, Schulte and Bohnet-Joschko [153] recommend continuous research in this area. BDA enables governments to gain insights into their spending and identify inconsistencies.

Fifth, smart urban and transport development: BDA can facilitate smarter urban planning and development by enabling the creation of smart cities through innovations in the internet and data science, both of which rely heavily on big data [154, 155]. To enhance the well-being of citizens in urban

environments, smart cities can use large-scale data for improved decision-making, air quality monitoring, and technology innovations in transportation and traffic management [155, 156]. Technology innovations in urban environments, such as transport and traffic management, heavily rely on large-scale data for training machine learning models and conducting simulation experiments to ensure the accuracy of these innovations [157, 158]. This can lead to reduced greenhouse gas emissions, improved security and emergency services, and better control of wastewater pollution, among other benefits.

Sixth, data governance: This aspect involves managing trade-offs between the effectiveness of public services and concerns related to privacy, ethics, and legal considerations surrounding data use and security in public institutions, including access rights management [159], personal data security, and [160, 161].

Seventh, decision support and strategy formulation: This theme highlights the role of BDA in facilitating informed decision-making in public institutions [162–164]. It also demonstrates how improved decision-making can lead to better national public policies [163–165]. BDA can assist public organizations in developing long-term business plans, including resource allocation, capacity development, and profit maximization strategies [163–166].

5. Discussion

The goal of RQ1 was to compile the most recent studies on BDA challenges in the realm of public institutions. To address this query, *Figure 4* illustrates a rising trend in the number of publications, indicating a growing interest in the research topic in academia. RQ2 aimed to explore the potential applications of BDA in the field of public institutions. These questions have been addressed by examining the context of the research papers that were analyzed. The purpose of this review was to assess the opportunities and constraints associated with the concept, which served as the framework for evaluating the selected papers, in order to gain a deeper understanding of how BDA is perceived in public institutions. BDA is being utilized in public institutions worldwide, but the results of the content analysis also reveal challenges that may impact their development and implementation in these institutions in the future. Since the study themes primarily focus on real-world usage and issues related to the use of BDA in public institutions, the challenges are addressed from a practical standpoint. The research limitations and conclusions, which also include insights into the future, indicate the potential for beneficial BDA developments and applications in public institutions.

BDA is theoretically defined as the analysis of a substantial volume of data with specific characteristics, generating new insights while also posing constraints for existing technology. These outcomes result from the methods employed. While academics and industry experts typically describe big data in terms of its features and associated analytics, definitions and these characteristics are essentially unrelated. Definitions are rarely provided in empirical research papers. The findings make it evident that there is no consensus on how big data should be defined. The study by [48], which also identifies a lack of consensus on a definition, is supported by the fact that other investigations have not identified a clear definition of big data either. This suggests that the absence of recommended definitions may not be solely attributable to the research methods applied. Since empirical studies primarily focus on the future of BDA, we turn to literature review assessments to gain insight into an appropriate definition. Nevertheless, one can be derived by applying the concept of information in public institutions, which is defined as "information, including all information products and services in any format, generated, created, collected, processed, preserved, maintained, funded by, or for public entities." "Information products in any format" could also be interpreted as "data in any form," which is a characteristic of big data, aside from the proliferation and explosion of information on a global scale. This type of "information product" could be generated in significantly larger quantities. Therefore, big data essentially resides in the data that government agencies possess. Public agencies utilize big data and related analytics, including public data analytics, which they and other organizations make available to promote transparency and stimulate economic growth. While there is no single, universally accepted definition or concept of open data, there does appear to be consensus on what qualifies as open data [170]. In a democratic setting, all public information is regarded as public data. In other words, BDA can employ color to represent the type, quantity, and rate of growth of this information.

One of the primary driving factors behind the development of BDA is the advancement of communication and information technology. Public entities have historically collected a wide range of

data types. Techniques like data mining are opening up new opportunities for uncovering insights, leading to the increasing prevalence of BDA in both corporate and government institutions. According to the selected papers, a significant component of this data-driven progress is focused on public decision-making. Consequently, the use of big data for evidence-based decision-making may become crucial in nearly every aspect of public institutions, offering new possibilities and benefits. The future of big data and its analytics could be further propelled by the growing popularity of social media and open data in public organizations.

Based on studies and discussions about the potential of BDA in public institutions, numerous examples exist of government entities practically utilizing BDA to benefit public authorities, organizations, society, and citizens. Many sectors, including urban management, healthcare, and citizen engagement, make use of BDA and vast datasets. For instance, BDA is employed to establish knowledge infrastructure in urban management, facilitating the management of pre- and post-travel conditions through an intelligence-based information system. Several cities, notably in Norway and Sweden, have explored the availability of such a system. Additionally, data from website searches, reservations, and customer reviews is frequently harnessed. There appears to be a consensus regarding the importance of data readiness, circulation, and their role in promoting openness, enhancing transparency, and increasing citizen engagement with government and its decisions. Concerns regarding the gap between the governing body and civil society in fostering an environment for freely accessible information are raised by [170]. Although experts continue to emphasize citizen interaction as a crucial element of the electronic government and data society era, this analysis demonstrates how much emphasis has been placed on how public institutions operate. The diverse range of BDA opportunities available to public institutions underscores the growing importance of big data for these organizations.

Future BDA implementations hold great promise for innovation, particularly in the public sector, healthcare, business, and industry. According to a recent study [171], BDA has the potential to assist healthcare professionals in maximizing their time with patients and enhancing the quality of care by utilizing accurate and easily accessible data to improve the precision and speed of patient information processing. Furthermore, BDA can facilitate electronic communication between patients and healthcare specialists, aiding in the generation of digital recommendations. Another possibility is the establishment of a centralized healthcare system, which may be achieved by conducting more extensive research on BDA and its utilization in constructing intricate medical networks connecting patients, healthcare providers, and pharmacists.

In addition to healthcare, new digital public institutions that enable people to vote and participate in civic life may emerge. While many countries are taking initiatives in this direction, their impact is limited because not everyone can engage in electronic government. One advantage of an electronic voting system is the potential for transparency in the voting process. Moreover, users should be able to access electronic services through a user-friendly graphical interface.

There are numerous obstacles and challenges associated with BDA, as well as several pressing concerns regarding its management. The collection, retention, retrieval, and overall processing and presentation of complex information are some of the challenges exacerbated by the sheer volume of datasets. Researchers argue that in order to cope with the exponential growth of data, public institutions and entities handling big data and related analytics must develop technology, management structures, and coding expertise. Additionally, as data volumes continue to expand, concerns about security and privacy risks associated with big data are likely to intensify. Therefore, each BDA program should assess potential privacy and security issues. Privacy may remain one of the most significant challenges in the context of big data, especially as some individuals are hesitant about having their data stored as electronic records. This poses a significant challenge when critical data is required to provide a digital service. Stringent data protection laws and regulations can also influence how users perceive data privacy. Moreover, having skilled and competent personnel is essential to address contemporary challenges effectively. Innovative management methods, processes, and policy frameworks are needed to tackle the challenges, transform BDA into actionable insights, and facilitate the comprehension of BDA outcomes. Other potential applications of BDA in public institutions that were examined include data governance, decision support, strategy formulation, and smart urban and transportation development.

A comprehensive framework was presented, consisting of six crucial components for BDA implementation in public institutions, based on the knowledge derived from the current systematic literature review and analysis, as illustrated in *Figure 7*. These components exhibit a degree of interconnectedness. The block diagram uses numbers (1, 2) to indicate that it first addresses the BDA components alphabetically (A-F), as indicated in *Figure 7* if public institutions are to harness and leverage BDA technologies in their daily business processes. However, the availability of both government public data and other types of data is necessary to ensure the feasibility of the beneficial outcomes while addressing the challenges. The overarching idea regarding the potential of BDA in public institutions is collectively summarized as the ability to make evidence-based decisions for the well-being of society, as depicted in *Figure 7*.

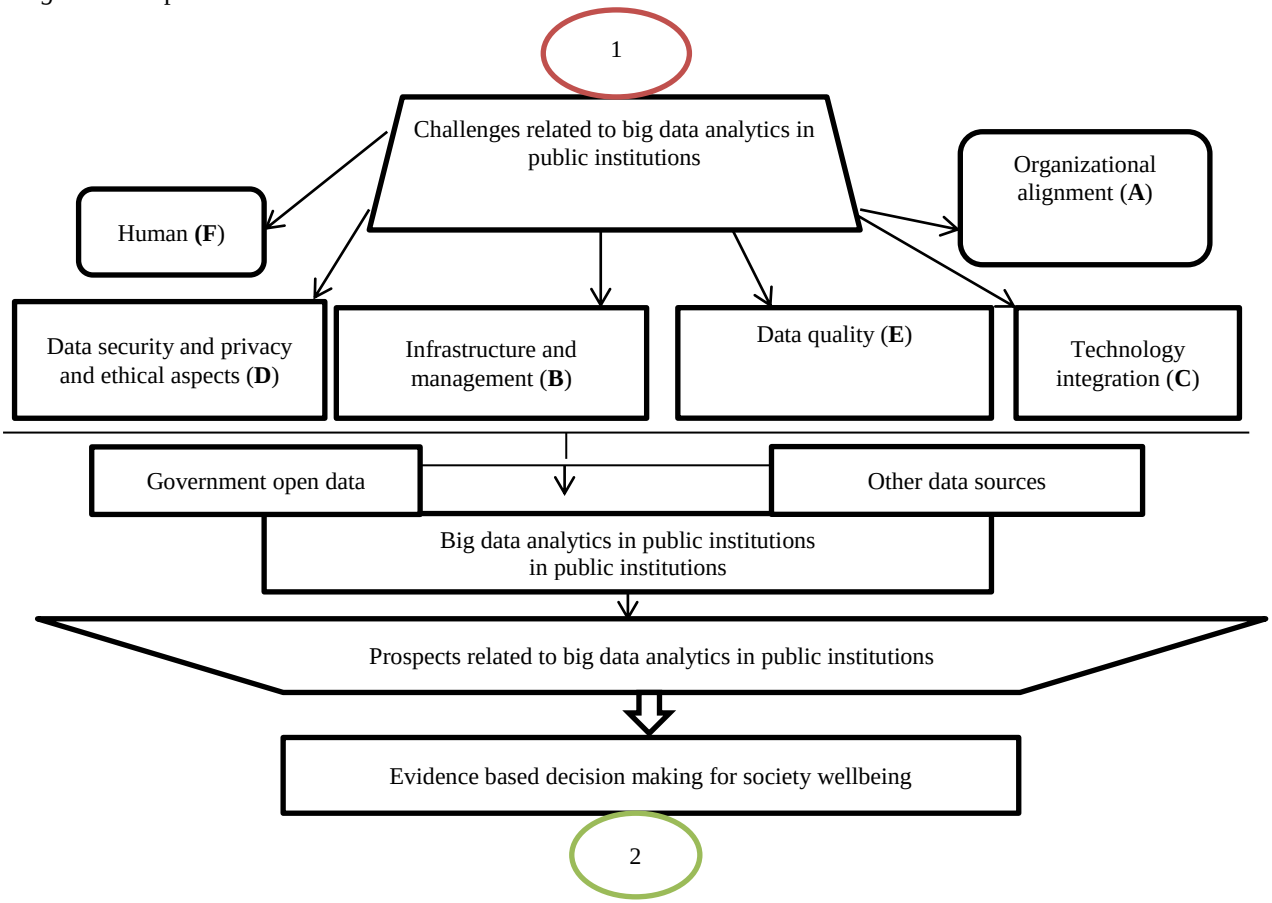

Figure 7 The comprehensive block diagram illustrating the complete approach of the study's observations about BDA in public institutions

6. Conclusion

Furthermore, two critical aspects, privacy, and security, which are essential for the development of any big data system, receive minimal attention. Current developments in data privacy require that BDA systems adhere to regional data privacy laws and meet stringent security standards. Despite potential concerns about security and privacy that could impede the growth of BDA, these issues are seldom addressed in studies, and clear guidelines are lacking. Additionally, information is collected cross-sectionally rather than longitudinally. The use of a longitudinal technique could enhance knowledge of the relationships and causation between different aspects.

Although various studies on BDA have been conducted, significant knowledge gaps remain regarding the public administration of big data, necessitating further research. Based on the findings of this literature review, recommendations are provided to identify potential areas for future research. Establishing a consistent definition of BDA within the context of public institutions, while

challenging, is essential to achieve a shared understanding. A clear definition could also benefit the coordinated use of big data in other industries.

This paper has provided a comprehensive understanding of various aspects related to BDA in public institutions. These aspects have been discussed in terms of barriers, prospects, and recommended actions for the future of BDA in the context of public institutions. While the literature review demonstrates that BDA holds significant potential to address various challenges faced by public institutions in domains such as health, education, agriculture, and food security, there are substantial barriers and gaps that must be addressed both before, during, and after the implementation of BDA in public institutions. Without overcoming these barriers, realizing the extensive benefits promised by BDA could remain elusive.

The design and development process of big data systems, while sharing similarities with other systems, differ in scale and processing nature. Some projections suggest that up to 90% of the data generated by devices may have limited value. Therefore, the initial step is to identify relevant data, which requires a deep understanding of functional requirements. Effective data management often involves handling a combination of structured and unstructured data. While there are various non-relational database structures available (such as key-value, key-document, and graph databases), processing such structures can be a complex task. Typically, processing in BDA relies on highly distributed processing frameworks, often utilizing architectures like Hadoop. This entails the creation and replication of clusters with the aim of optimizing cluster configurations to balance workloads. Data must be processed once it has been stored and defined, and traditional languages like Structured Query Language (SQL) may not be sufficient for this task. Developing new languages and tools for data processing is a critical challenge. After data processing, the next step involves understanding and effectively applying the results. A complete list of abbreviations is shown in *Appendix I.*

Acknowledgment

The authors would like to express their gratitude to Ms. Celina T. Teri for her contributions in both psychological support and encouragement, which greatly facilitated the completion of this work.

Conflicts of interest

The authors have no conflicts of interest to declare.

References

[1] Rodionov AA, Fayziev RA, Gulyamov SS. Experience in using big data technology for digitalization of information. In proceedings of the 6th international conference on future networks & distributed systems 2022 (pp. 412-5).

[2] Garg A, Popli R, Sarao BS. Growth of digitization and its impact on big data analytics. In IOP conference series: materials science and engineering 2021 (pp. 1-9). IOP Publishing.

[3] Lampropoulos G. Artificial intelligence, big data, and machine learning in industry 4.0. In encyclopedia of data science and machine learning 2023 (pp. 2101-9). IGI Global.

[4] Molčan M, Čajková A. Knowledge management and local government. In developments in information and knowledge management systems for business applications 2023 (pp. 293-312). Cham: Springer Nature Switzerland.

[5] Momen MN, Ferdous J. Governance in Bangladesh: innovations in delivery of public service. Springer Nature; 2023.

[6] Sahid NZ, Sani MK, Noordin SA, Zaini MK, Baba J. Determinants factors of intention to adopt big data analytics in Malaysian public agencies. Journal of Industrial Engineering and Management. 2021; 14(2):269-93.

[7] Cronemberger F, Gil-garcia JR. Big data and analytics as strategies to generate public value in smart cities: proposing an integrative framework. Setting Foundations for the Creation of Public Value in Smart Cities. 2019:247-67.

[8] Pramanik S, Bandyopadhyay SK. Analysis of big data. In encyclopedia of data science and machine learning 2023 (pp. 97-115). IGI Global.

[9] Supriyadi RS, Abraham I, Giyanto B, Asropi A. Big data in the public sector: systematic literature review and bibliometric analysis. Jurnal Ilmiah Mandala Education. 2023; 9(1):1-5.

[10] Chen J, Tao Y, Wang H, Chen T. Big data based fraud risk management at Alibaba. The Journal of Finance and Data Science. 2015; 1(1):1-10.

[11] Fan S, Lau RY, Zhao JL. Demystifying big data analytics for business intelligence through the lens of marketing mix. Big Data Research. 2015; 2(1):28-32.

[12] Xiang Z, Schwartz Z, Gerdes JH, Uysal M. What can big data and text analytics tell us about hotel guest experience and satisfaction? International Journal of Hospitality Management. 2015; 44:120-30.

[13] Vassakis K, Petrakis E, Kopanakis I. Big data analytics: applications, prospects and challenges. Mobile Big Data: A Roadmap from Models to Technologies. 2018:3-20.

[14] Gamage P. New development: leveraging 'big data'analytics in the public sector. Public Money & Management. 2016; 36(5):385-90.

[15] Moorthy J, Lahiri R, Biswas N, Sanyal D, Ranjan J, Nanath K, et al. Big data: prospects and challenges. Vikalpa. 2015; 40(1):74-96.

[16] Surya L. An exploratory study of AI and big data, and it's future in the United States. International Journal of Creative Research Thoughts. 2015; 3(2):991-5.
[17] Adusumalli HP, Pasupuleti MB. Applications and practices of big data for development. Asian Business Review. 2017; 7(3):111-6.
[18] Mittal P. Big data and analytics: a data management perspective in public administration. International Journal of Big Data Management. 2020; 1(2):152-65.
[19] Naeem M, Jamal T, Diaz-martinez J, Butt SA, Montesano N, Tariq MI, et al. Trends and future perspective challenges in big data. In advances in intelligent data analysis and applications: proceeding of the sixth euro-China conference on intelligent data analysis and applications 2022 (pp. 309-25). Springer Singapore.
[20] Väyrynen H, Helander N, Jalonen H. Public Innovation and Digital Transformation. Taylor & Francis; 2023.
[21] Hossain E, Khan I, Un-noor F, Sikander SS, Sunny MS. Application of big data and machine learning in smart grid, and associated security concerns: a review. IEEE Access. 2019; 7:13960-88.
[22] Shi Y. Advances in big data analytics. Adv Big Data Anal. Springer Nature; 2022.
[23] Emmanuel I, Stanier C. Defining big data. In proceedings of the international conference on big data and advanced wireless technologies 2016 (pp. 1-6). ACM.
[24] Grover V, Chiang RH, Liang TP, Zhang D. Creating strategic business value from big data analytics: a research framework. Journal of Management Information Systems. 2018; 35(2):388-423.
[25] Kalmbach P. o 'zapft is: Tap Your Networking Algorithm's Big Data!. 2017.
a. https://pdfs.semanticscholar.org/a7c2/236528b5ea2a7d36ccb3ea4cb7ff0d15a301.pdf.
[26] Van DHG, Klievink AJ, Arnaboldi M, Meijer AJ. Rationality and politics of algorithms. will the promise of big data survive the dynamics of public decision making? Government Information Quarterly. 2019; 36(1):27-38.
[27] Van DH, Arnaboldi M, Dew BH, Klievink B. Will the promise of big data survive the dynamics of public decision making? the need for checks and balances. In EGPA annual conference, European group for public administration 2016.
[28] Marr B. Data strategy: how to profit from a world of big data, analytics and artificial intelligence. Kogan Page Publishers; 2021.
[29] Talaoui Y, Kohtamäki M, Ranta M, Paroutis S. Recovering the divide: a review of the big data analytics-strategy relationship. Long Range Planning. 2023.
[30] Visco G, Plattner SH, Fortini P, Sammartino M. A multivariate approach for a comparison of big data matrices. Case study: thermo-hygrometric monitoring inside the Carcer Tullianum (Rome) in the absence and in the presence of visitors. Environmental Science and Pollution Research. 2017:13990-4004.
[31] Gul R, Leong K, Mubashar A, Al-faryan MA, Sung A. The empirical nexus between data-driven decision-making and productivity: evidence from Pakistan's banking sector. Cogent Business & Management. 2023; 10(1):1-17.
[32] Gärtner B, Hiebl MR. Issues with big data. In the Routledge companion to accounting information systems 2017 (pp. 161-72). Routledge.
[33] Lohr SL, Raghunathan TE. Combining survey data with other data sources. 2017:293-312.
[34] Hartmann T, Moawad A, Fouquet F, Nain G, Klein J, Traon YL, et al. Model-driven analytics: connecting data, domain knowledge, and learning. ArXiv preprint arXiv:1704.01320. 2017.
[35] Blenk A, Kalmbach P, Kellerer W, Schmid S. O'zapft is: tap your network algorithm's big data!. In proceedings of the workshop on big data analytics and machine learning for data communication networks 2017 (pp. 19-24).
[36] Merhi MI, Bregu K. Effective and efficient usage of big data analytics in public sector. Transforming Government: People, Process and Policy. 2020; 14(4):605-22.
[37] Rogge N, Agasisti T, De WK. Big data and the measurement of public organizations' performance and efficiency: the state-of-the-art. Public Policy and Administration. 2017; 32(4):263-81.
[38] Desouza KC, Jacob B. Big data in the public sector: lessons for practitioners and scholars. Administration & Society. 2017; 49(7):1043-64.
[39] Visvizi A, Troisi O, Grimaldi M. Mapping and conceptualizing big data and its value across issues and domains. In big data and decision-making: applications and uses in the public and private sector 2023 (pp. 15-25). Emerald Publishing Limited.
[40] Sarkis-onofre R, Catalá-lópez F, Aromataris E, Lockwood C. How to properly use the PRISMA statement. Systematic Reviews. 2021; 10(1):1-3.
[41] Nazari E, Afkanpour M, Tabesh H. Big data from A to Z. Frontiers in Health Informatics. 2019; 8(1):1-5.
[42] Cartledge C. How Many VS are there in big data. 2006:1-4.
[43] Khanna D, Jindal N, Singh H, Rana PS. Applications and challenges in healthcare big data: a strategic review. Current Medical Imaging. 2023; 19(1):27-36.
[44] Mcdermott S, Turk MA. What are the implications of the big data paradigm shift for disability and health? Disability and Health Journal. 2015; 3(8):303-4.
[45] Reinsel D, Gantz J, Rydning J. Data age 2025: the evolution of data to life-critical. From Seagate. Framingham, MA, US: International Data Corporation. 2017.
[46] Athey S. Beyond prediction: using big data for policy problems. Science. 2017; 355(6324):483-5.
[47] Gill SS, Chana I, Buyya R. IoT based agriculture as a cloud and big data service: the beginning of digital India. Journal of Organizational and End User Computing (JOEUC). 2017; 29(4):1-23.

[48] De MA, Greco M, Grimaldi M. A formal definition of big data based on its essential features. Library Review. 2016; 65(3):122-35.
[49] Galetsi P, Katsaliaki K, Kumar S. Big data analytics in health sector: theoretical framework, techniques and prospects. International Journal of Information Management. 2020; 50:206-16.
[50] Mergel I, Rethemeyer K, Isett KR. What does big data mean to public affairs research? understanding the methodological and analytical challenges. Impact of Social Sciences Blog. 2016.
[51] Khanra S, Dhir A, Islam AN, Mäntymäki M. Big data analytics in healthcare: a systematic literature review. Enterprise Information Systems. 2020; 14(7):878-912.
[52] Bahri S, Zoghlami N, Abed M, Tavares JM. Big data for healthcare: a survey. IEEE Access. 2018; 7:7397-408.
[53] Khan A, Tao M. Knowledge absorption capacity's efficacy to enhance innovation performance through big data analytics and digital platform capability. Journal of Innovation & Knowledge. 2022; 7(3):1-13.
[54] Wang Y, Kung L, Byrd TA. Big data analytics: understanding its capabilities and potential benefits for healthcare organizations. Technological Forecasting and Social Change. 2018; 126:3-13.
[55] Watson HJ. Update tutorial: big data analytics: concepts, technology, and applications. Communications of the Association for Information Systems. 2019; 44(1):1-14.
[56] Statistics UB. Statistical abstract. Kampala: Uganda Bureau of Statistics. 2013.
[57] Mishra D, Luo Z, Hazen B, Hassini E, Foropon C. Organizational capabilities that enable big data and predictive analytics diffusion and organizational performance: a resource-based perspective. Management Decision. 2019; 57(8):1734-55.
[58] Galetsi P, Katsaliaki K. A review of the literature on big data analytics in healthcare. Journal of the Operational Research Society. 2020; 71(10):1511-29.
[59] Kamble SS, Gunasekaran A, Goswami M, Manda J. A systematic perspective on the applications of big data analytics in healthcare management. International Journal of Healthcare Management. 2018; 12(3):226-40.
[60] Boutabba MA, Lardic S. EU emissions trading scheme, competitiveness and carbon leakage: new evidence from cement and steel industries. Annals of Operations Research. 2017; 255:47-61.
[61] Ozminkowski RJ, Wells TS, Hawkins K, Bhattarai GR, Martel CW, Yeh CS. Big data, little data, and care coordination for medicare beneficiaries with medigap coverage. Big Data. 2015; 3(2):114-25.
[62] Amirian P, Van LF, Lang T, Thomas A, Peeling R, Basiri A, et al. Using big data analytics to extract disease surveillance information from point of care diagnostic machines. Pervasive and Mobile Computing. 2017; 42:470-86.
[63] Singh RK, Agrawal S, Sahu A, Kazancoglu Y. Strategic issues of big data analytics applications for managing health-care sector: a systematic literature review and future research agenda. The TQM Journal. 2023; 35(1):262-91.
[64] Kalid N, Zaidan AA, Zaidan BB, Salman OH, Hashim M, Muzammil HJ. Based real time remote health monitoring systems: A review on patients prioritization and related" big data" using body sensors information and communication technology. Journal of Medical Systems. 2018; 42:1-30.
[65] Prasser F, Spengler H, Bild R, Eicher J, Kuhn KA. Privacy-enhancing ETL-processes for biomedical data. International Journal of Medical Informatics. 2019; 126:72-81.
[66] Harerimana G, Jang B, Kim JW, Park HK. Health big data analytics: a technology survey. IEEE Access. 2018; 6:65661-78.
[67] Zhang R, Simon G, Yu F. Advancing Alzheimer's research: a review of big data promises. International Journal of Medical Informatics. 2017; 106:48-56.
[68] Alharthi A, Krotov V, Bowman M. Addressing barriers to big data. Business Horizons. 2017; 60(3):285-92.
[69] Amalina F, Hashem IA, Azizul ZH, Fong AT, Firdaus A, Imran M, et al. Blending big data analytics: review on challenges and a recent study. IEEE Access. 2019; 8:3629-45.
[70] Zhuang YT, Wu F, Chen C, Pan YH. Challenges and opportunities: from big data to knowledge in AI 2.0. Frontiers of Information Technology & Electronic Engineering. 2017; 18:3-14.
[71] Sáez C, García-gómez JM. Kinematics of big biomedical data to characterize temporal variability and seasonality of data repositories: functional data analysis of data temporal evolution over non-parametric statistical manifolds. International Journal of Medical Informatics. 2018; 119:109-24.
[72] Coyne EM, Coyne JG, Walker KB. Big Data information governance by accountants. International Journal of Accounting & Information Management. 2018; 26(1):153-70.
[73] Vitale G, Cupertino S, Riccaboni A. Big data and management control systems change: the case of an agricultural SME. Journal of Management Control. 2020; 31:123-52.
[74] Boell SK, Cecez-kecmanovic D. On being 'systematic'in literature reviews in IS. Journal of Information Technology. 2015; 30(2):161-73.
[75] Paré G, Kitsiou S. Methods for literature reviews. In Handbook of eHealth Evaluation: an evidence-based approach [Internet] 2017. University of Victoria.
[76] Paul J, Barari M. Meta-analysis and traditional systematic literature reviews-What, why, when, where, and how? Psychology & Marketing. 2022; 39(6):1099-115.
[77] Selçuk AA. A guide for systematic reviews: PRISMA. Turkish Archives of Otorhinolaryngology. 2019; 57(1):57-8.
[78] Zainal NZ, Hussin H, Nazri MN. Big data initiatives by governments--issues and challenges: a review. In 6th international conference on information and

communication technology for the Muslim world 2016 (pp. 304-9). IEEE.
[79] Al-sai ZA, Abdullah R. Big data impacts and challenges: a review. In Jordan international joint conference on electrical engineering and information technology 2019 (pp. 150-5). IEEE.
[80] Kamilaris A, Kartakoullis A, Prenafeta-boldú FX. A review on the practice of big data analysis in agriculture. Computers and Electronics in Agriculture. 2017; 143:23-37.
[81] Desai PV. A survey on big data applications and challenges. In second international conference on inventive communication and computational technologies 2018 (pp. 737-40). IEEE.
[82] Mureddu F, Schmeling J, Kanellou E. Research challenges for the use of big data in policy-making. Transforming Government: People, Process and Policy. 2020; 14(4):593-604.
[83] Praharaj S. Development challenges for big data command and control centres for smart cities in India. Data-driven Multivalence in the Built Environment. 2020:75-90.
[84] Cannataci J, Falce V, Pollicino O. Legal challenges of big data. Edward Elgar Publishing; 2020.
[85] Tse D, Chow CK, Ly TP, Tong CY, Tam KW. The challenges of big data governance in healthcare. In 17th IEEE international conference on trust, security and privacy in computing and communications/12th IEEE international conference on big data science and engineering (TrustCom/BigDataSE) 2018 (pp. 1632-6). IEEE.
[86] Pradhan P, Shakya S. Big data challenges for e-Government services in Nepal. Journal of the Institute of Engineering. 2018; 14(1):216-22.
[87] Albqowr A, Alsharairi M, Alsoussi A. Big data analytics in supply chain management: a systematic literature review. VINE Journal of Information and Knowledge Management Systems. 2022.
[88] Shastri A, Deshpande M. A review of big data and its applications in healthcare and public sector. Big Data Analytics in Healthcare. 2020:55-66.
[89] Younas M. Research challenges of big data. Service Oriented Computing and Applications. 2019; 13:105-7.
[90] Kondraganti A, Narayanamurthy G, Sharifi H. A systematic literature review on the use of big data analytics in humanitarian and disaster operations. Annals of Operations Research. 2022; 21:1-38.
[91] Eachempati P, Srivastava PR. Systematic literature review of big data analytics. In proceedings of the SIGMIS conference on computers and people research 2017 (pp. 177-8). ACM.
[92] Candra S, Joartha JJ, Moniaga JV, Hidayat MF, Jabar BA. A systematic literature review of big data analytics challenges in cloud computing. In 2nd international conference on electronic and electrical engineering and intelligent system 2022 (pp. 246-51). IEEE.
[93] Pappas IO, Mikalef P, Giannakos MN, Krogstie J, Lekakos G. Big data and business analytics ecosystems: paving the way towards digital transformation and sustainable societies. Information Systems and e-business Management. 2018:479-91.
[94] Kolajo T, Daramola O, Adebiyi A. Big data stream analysis: a systematic literature review. Journal of Big Data. 2019; 6(1):1-30.
[95] Mikalef P, Pappas IO, Krogstie J, Giannakos M. Big data analytics capabilities: a systematic literature review and research agenda. Information Systems and e-Business Management. 2018; 16:547-78.
[96] Vanani IR, Majidian S. Literature review on big data analytics methods. Social Media and Machine Learning; 2019.
[97] Aggarwal AK. Opportunities and challenges of big data in public sector. Web Services: Concepts, Methodologies, Tools, and Applications. 2019:1749-61.
[98] Morabito V, Morabito V. Big data and analytics for government innovation. Big Data and Analytics: Strategic and Organizational Impacts. 2015:23-45.
[99] Batko K, Ślęzak A. The use of big data analytics in healthcare. Journal of Big Data. 2022; 9(1):1-24.
[100] Verma S. Sentiment analysis of public services for smart society: literature review and future research directions. Government Information Quarterly. 2022; 39(3):101708.
[101] Rafiq F, Awan MJ, Yasin A, Nobanee H, Zain AM, Bahaj SA. Privacy prevention of big data applications: a systematic literature review. SAGE Open. 2022; 12(2).1582440221096445.
[102] Lazarevska ZB, Tocev T, Dionisijev I. How to improve performance in public sector auditing through the power of big data and data analytics? - the case of the republic of north macedonia. Journal of Accounting, Finance and Auditing Studies. 2022; 8(3):187-209.
[103] Jin W. Challenges and innovative countermeasures faced by public administration in the context of big data and internet of things. Mathematical Problems in Engineering. 2022; 2022:1-10.
[104] Barham H, Daim T. The use of readiness assessment for big data projects. Sustainable Cities and Society. 2020; 60:102233.
[105] Gandomi AH, Chen F, Abualigah L. Machine learning technologies for big data analytics. Electronics. 2022; 11(3):421.
[106] Kaufmann M. Big data management canvas: a reference model for value creation from data. Big Data and Cognitive Computing. 2019; 3(1):1-18.
[107] Lee I. Big data: dimensions, evolution, impacts, and challenges. Business Horizons. 2017; 60(3):293-303.
[108] Williamson B. Governing through infrastructural control: artificial intelligence and cloud computing in the data-intensive state. The SAGE Handbook of Digital Society. Thousand Oaks, CA: SAGE. 2023.
[109] Hagsten E. ICT infrastructure in firms and online sales. Electronic Commerce Research. 2022:1-20.

[110]Eldow A, Alsharida RA, Hammood M, Shakir M, Malik SI, Muttar AK, et al. Information communication technology infrastructure in Sudanese Governmental Universities. Recent Advances in Intelligent Systems and Smart Applications. 2021:363-75.
[111]Shin DH, Choi MJ. Ecological views of big data: perspectives and issues. Telematics and Informatics. 2015; 32(2):311-20.
[112]Anna K, Nikolay K. Survey on big data analytics in public sector of Russian federation. Procedia Computer Science. 2015; 55:905-11.
[113]Özköse H, Arı ES, Gencer C. Yesterday, today and tomorrow of big data. Procedia-Social and Behavioral Sciences. 2015; 195:1042-50.
[114]Ullrich A. Opportunities and challenges of big data and predictive analytics for achieving the UN's SDGs. Pacific Asia Conference on Information Systems. 2022:1-17.
[115]Soomro K, Bhutta MN, Khan Z, Tahir MA. Smart city big data analytics: an advanced review. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. 2019; 9(5):e1319.
[116]Raad E, Al BB, Chbeir R. Preventing sensitive relationships disclosure for better social media preservation. International Journal of Information Security. 2016; 15:173-94.
[117]Malomane R, Musonda I, Okoro CS. The opportunities and challenges associated with the implementation of fourth industrial revolution technologies to manage health and safety. International Journal of Environmental Research and Public Health. 2022; 19(2):846.
[118]Pandey R, Dhoundiyal M. Quantitative evaluation of big data categorical variables through R. Procedia Computer Science. 2015; 46:582-8.
[119]Sureephong P, Komanee P, Trongpanyachot C. Data sharing and electronic medical record privacy protection of out-patient-department using blockchain. In 25th international symposium on wireless personal multimedia communications 2022 (pp. 303-7). IEEE.
[120]Rossi R, Hirama K. Characterizing big data management. arXiv preprint arXiv:2201.05929. 2022.
[121]Moreno A, Molano-pulido J, Gomez-morantes JE, Gonzalez RA. ADACOP: a big data platform for open government data. In proceedings of the 15th international conference on theory and practice of electronic governance 2022 (pp. 369-75).
[122]Da STL, Magalhães RP, Brilhante IR, De MJA, Araújo D, Rego PA, et al. Big data analytics technologies and platforms: a brief review. LADaS@ VLDB. 2018:25-32.
[123]Liu SM, Yuan Q. The evolution of information and communication technology in public administration. Public Administration and Development. 2015; 35(2):140-51.
[124]Biswas R. Outlining big data analytics in health sector with special reference to Covid-19. Wireless Personal Communications. 2022; 124(3):2097-108.
[125]Coulthart S, Riccucci R. Putting big data to work in government: the case of the United States border patrol. Public Administration Review. 2022; 82(2):280-9.
[126]Hassenstein MJ, Vanella P. Data quality—concepts and problems. Encyclopedia. 2022; 2(1):498-510.
[127]Spangenberg N, Wilke M, Franczyk B. A big data architecture for intra-surgical remaining time predictions. Procedia Computer Science. 2017; 113:310-7.
[128]Manson SM. Big Data and Human-Environment Systems. Cambridge University Press; 2023.
[129]Weibl J, Hess T. Success or failure of big data: Insights of managerial challenges from a technology assimilation perspective. Proceedings of the Multikonferenz Wirtschaftsinformatik (MKWI). 2018:12-59.
[130]Munné R. Big data in the public sector. New Horizons for a Data-Driven Economy: A Roadmap for Usage and Exploitation of Big Data in Europe; 2016:195-208.
[131]Hardy K, Maurushat A. Opening up government data for Big Data analysis and public benefit. Computer Law & Security Review. 2017; 33(1):30-7.
[132]Mavriki P, Karyda M. Big data analytics in e-government and e-democracy applications: privacy threats, implications and mitigation. International Journal of Electronic Governance. 2022; 14(1-2):58-82.
[133]Abouelmehdi K, Beni-hessane A, Khaloufi H. Big healthcare data: preserving security and privacy. Journal of Big Data. 2018; 5(1):1-8.
[134]Zulkarnain N, Anshari M, Hamdan M, Fithriyah M. Big data in business and ethical challenges. In international conference on information management and technology 2021 (pp. 298-303). IEEE.
[135]Gumbus A, Grodzinsky F. Era of big data: danger of descrimination. ACM SIGCAS Computers and Society. 2016; 45(3):118-25.
[136]Pavone P, Ricci P, Calogero M. New challenges for public value and accountability in the age of big data: a bibliometric analysis. Meditari Accountancy Research. 2023.
[137]Kacem A, Belkaroui R, Jemal D, Ghorbel H, Faiz R, Abid IH. Towards improving e-government services using social media-based citizen's profile investigation. In proceedings of the 9th international conference on theory and practice of electronic governance 2016 (pp. 187-90).
[138]Bello-orgaz G, Jung JJ, Camacho D. Social big data: recent achievements and new challenges. Information Fusion. 2016; 28:45-59.
[139]Mohan DA. Big data analytics: recent achievements and new challenges. International Journal of Computer Applications Technology and Research. 2016; 5(7):460-4.
[140]Lev-on A, Steinfeld N. Local engagement online: municipal Facebook pages as hubs of interaction. Government Information Quarterly. 2015; 32(3):299-307.

[141]Lande D, Subach I, Puchkov A. A system for analysis of big data from social media. Information & Security. 2020; 47(1):44-61.
[142]Saggi MK, Jain S. A survey towards an integration of big data analytics to big insights for value-creation. Information Processing & Management. 2018; 54(5):758-90.
[143]Jordan L. The problem with big data in translational medicine. a review of where we've been and the possibilities ahead. Applied & Translational Genomics. 2015; 6:3-6.
[144]Corallo A, Fortunato L, Matera M, Alessi M, Camillò A, Chetta V, et al. Sentiment analysis for government: an optimized approach. In machine learning and data mining in pattern recognition: 11th international conference, MLDM 2015, Hamburg, Germany, 2015 (pp. 98-112). Springer International Publishing.
[145]Li L, Ma S, Wang R, Wang Y, Zheng Y. Citizen participation in the co-production of urban natural resource assets: analysis based on social media big data. Journal of Global Information Management. 2021; 30(6):1-21.
[146]Dash S, Shakyawar SK, Sharma M, Kaushik S. Big data in healthcare: management, analysis and future prospects. Journal of Big Data. 2019; 6(1):1-25.
[147]Meng L, Li W. Application of big data technology in the field of public service: precise demand management. In proceedings of the 5th international conference on computer science and software engineering 2022 (pp. 670-4).
[148]Xu J. An accurate management method of public services based on big data and cloud computing. Journal of Cloud Computing. 2023; 12(1):80.
[149]De OCDK, Da SGG, Porto RB. Big data in the brazilian public health sector: concept, characteristics, benefits, and challenges. Revista do Serviço Público. 2022; 73(4):594-621.
[150]Keikhosrokiani P. Big data analytics for healthcare: datasets, techniques, life cycles, management, and applications. Academic Press; 2022.
[151]Goundar S, Masilamani K, Bhardwaj A, Dhasarathan C. Big data analytics in healthcare: a developing country survey. In applications of big data in large-and small-scale systems 2021 (pp. 85-98). IGI Global.
[152]Mishra S, Tripathy HK, Mishra BK, Sahoo S. Usage and analysis of big data in E-health domain. In research anthology on big data analytics, architectures, and applications 2022 (pp. 417-30). IGI Global.
[153]Schulte T, Bohnet-joschko S. How can big data analytics support people-centred and integrated health services: a scoping review. International Journal of Integrated Care. 2022; 22(2):23.
[154]Sazu MH, Akter JS. Impact of big data analytics on government organizations. Management & Datascience. 2022; 6(2).
[155]Khan S. Barriers of big data analytics for smart cities development: a context of emerging economies. International Journal of Management Science and Engineering Management. 2022; 17(2):123-31.
[156]Dey N, Ella HA, Bhatt C, S AA, Chandra SS. Internet of things and big data analytics toward next-generation intelligence; 2018.
[157]Losurdo F, Dileo I, Siergiejczyk M, Krzykowska K, Krzykowski M. Innovation in the ICT infrastructure as a key factor in enhancing road safety: a multi-sectoral approach. In 25th international conference on systems engineering 2017 (pp. 157-62). IEEE.
[158]Rathore MM, Ahmad A, Paul A, Rho S. Urban planning and building smart cities based on the internet of things using big data analytics. Computer Networks. 2016; 101:63-80.
[159]Di VA, Hassan R, Alavoine C. Data intelligence and analytics: a bibliometric analysis of human–artificial intelligence in public sector decision-making effectiveness. Technological Forecasting and Social Change. 2022; 174:121201.
[160]Wibowo S, Sandikapura T. Improving data security, interoperability, and veracity using blockchain for one data governance, case study of local tax big data. In international conference on ICT for smart society 2019 (pp. 1-6). IEEE.
[161]Leonelli S. Data governance is key to interpretation: reconceptualizing data in data science. Harvard Data Science Review. 2019;1(1):1-9.
[162]Moutselos K, Kyriazis D, Maglogiannis I. A web based modular environment for assisting health policy making utilizing big data analytics. In 9th international conference on information, intelligence, systems and applications 2018 (pp. 1-5). IEEE.
[163]Saunders GH, Christensen JH, Gutenberg J, Pontoppidan NH, Smith A, Spanoudakis G, et al. Application of big data to support evidence-based public health policy decision-making for hearing. Ear and Hearing. 2020; 41(5):1057-63.
[164]Rahmanto F, Pribadi U, Priyanto A. Big data: what are the implications for public sector policy in society 5.0 era. In IOP conference series: earth and environmental science 2021 (pp. 1-7). IOP Publishing.
[165]Amankwah-amoah J. Safety or no safety in numbers? Governments, big data and public policy formulation. Industrial Management & Data Systems. 2015; 115(9):1596-603.
[166]El MIG, Mardiyanta A, Suryono A, Rizkika HL. Who owns big data? examine the policies of rural government in Indonesia. Promoting Adaptive System to The Current Turbulence Within Crisis Environments. 2023:298.
[167]Panori A, Kakderi C, Komninos N. Transformation of smart city public services through AI and big data analytics: towards universal cross-sector solutions. Chapters. 2023:292-307.
[168]Barkham R, Bokhari S, Saiz A. Urban big data: city management and real estate markets. Artificial Intelligence, Machine Learning, and Optimization Tools for Smart Cities: Designing for Sustainability. 2022:177-209.
[169]Liu X, Chang G. Construction of a public service cloud platform for disabled people based on the big data management model of the internet of things.

Computational Intelligence and Neuroscience. 2023; 2023:1-11.

[170]Ohemeng FL, Ofosu-adarkwa K. One way traffic: the open data initiative project and the need for an effective demand side initiative in Ghana. Government Information Quarterly. 2015; 32(4):419-28.

[171]Huang T, Lan L, Fang X, An P, Min J, Wang F. Promises and challenges of big data computing in health sciences. Big Data Research. 2015; 2(1):2-11.

Matendo Didas completed his Certificate and Diploma in Computer Science at Makerere University Business School (MUBS), Uganda, in 2012. He earned his BSc. in Software Engineering from Makerere University (MAK), Uganda, in 2017. Currently, he is pursuing a Master's degree in Information Technology, specializing in big data analytics, big data architectures, and Management at the College of Information and Virtual Education (CIVE), The University of Dodoma (UDOM). Until 2021, he worked as a remote instructor in business intelligence and database systems at Makerere University. His research interests encompass artificial intelligence, big data analytics in public institutions, mobile data in lifelong learning, and Data Mining.

Email: matendodidas@gmail.com

Appendix I

S. No.	Abbreviation	Description
1	4thIR	Fourth Industrial Revolution
2	AD	Autonomous Driving
3	AI	Artificial Intelligence
4	BDA	Big Data Analytics
5	HDFS	Hadoop Distributed File System
6	IoT	Internet of Things
7	MR	MapReduce
8	NoSQL	Not Only Structured Query Language
9	PII	Personal Identification Information
10	PRISMA	Preferred Report Items for Systematic Review and Meta-Analysis
11	SALSA	Search, Appraisal, Synthesis, and Analysis
12	SLRA	Systematic Literature Review Analysis
13	SQL	Structured Query Language