

Big Data Governance in Healthcare Institutions: A Systematic Review and Implications for Tanzania

Matendo Didas*, James Samuel Chambua, Godfrey Njulumi Justo

University of Dar es Salaam, Dar es Salaam, Tanzania

Background and Purpose: The rapid adoption of digital health technologies has led to unprecedented growth in healthcare data, commonly referred to as big data. To guarantee data quality, privacy protection, and trustworthy decision-making within healthcare organizations, this data must be governed effectively in nexus with context-aware aspects.

Methods: This paper adopts a systematic literature review (SLR) guided by the PRISMA 2020 framework to synthesize literature on big data governance in healthcare, with emphasis on the Tanzanian healthcare context. Studies published between 2010 and 2025 were identified from Scopus, PubMed, IEEE Xplore, ScienceDirect, and Web of Science. Of the 210 studies initially identified, 28 met the eligibility criteria and were included in the analysis.

Results: The review identified four main findings: (1) key big data governance components, including data quality, privacy and security, interoperability, and regulatory compliance; (2) governance objectives focused on improving data quality, security, accessibility, accountability, and effective data use; (3) recurring challenges, particularly inadequate contextual infrastructure, data silos, skills shortages, regulatory gaps, and implementation costs; and (4) different types of governance frameworks, ranging from conceptual and organizational models to healthcare-specific frameworks. Overall, the findings indicate that existing frameworks are predominantly conceptual and insufficiently adapted to the contextual realities of resource-constrained healthcare settings.

Conclusion: This study proposes a context-aware big data governance context-aware conceptual model, identifies critical gaps in the current literature, and highlights the need for context-aware governance approaches. It further outlines implications for Tanzania's healthcare system adoption, providing a foundation for the development of tailored big data governance frameworks suitable for resource-constrained healthcare systems

Keywords: Big data governance, Data governance elements, Healthcare data, Systematic review, Big data governance frameworks

1 Introduction

Healthcare systems around the world are undergoing fast digital transformation, with a greater reliance on data-driven decision-making to improve service delivery, efficiency, and patient outcomes [1]. The widespread use of electronic health records (EHRs), health information systems such as District Health Information Software 2 (DHIS2), mobile health applications, and advanced analytics tools has resulted in the creation of massive amounts of heterogeneous healthcare data termed as big data [2]. These innovations offer enormous prospects to improve clinical decision-making, illness surveillance, resource allocation, and health-care system planning [2, 3].

However, governing this big data has become a crucial area in healthcare systems. The policies, standards, procedures, and institutional structures that guarantee the accessibility, usability, quality, integrity, and security of data throughout its lifecycle are collectively referred to as big data governance [4, 5]. In addition to promoting innovation and research, effective big data governance frameworks help healthcare organizations maintain high data quality, guarantee regulatory compliance, safeguard patient privacy, and build stakeholder confidence [5, 6]. Data governance is especially important in healthcare

*Corresponding author address: University of Dar es Salaam, P.O. Box 33335, Dar es Salaam, Tanzania, email: mdidas@sjut.ac.tz
© 2026 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2026;13(2):65-84. DOI: 10.12856/JHIA-2026-v13-i2-696

since health data is generally sensitive and includes individually identifiable and confidential patient information. Poorly executed data governance measures can result in data breaches, poor quality data, ethical violations, and loss of public trust among others, compromising the benefits of digital health efforts [7, 8]. It is also imperative to note that the growing usage of big data technologies such as artificial intelligence and machine learning creates new governance issues in data ownership, interoperability, algorithmic transparency, and accountability [9].

Despite worldwide progress, there are considerable differences in the application of big data governance among regions as asserted by [10]. Like any other developing countries, Tanzania's digital health systems are rapidly emerging which result in more big data. In particular, Tanzania's measures such as the national eHealth plan and digital health investments have even increased data collection at all levels of the healthcare system, particularly in district hospitals [11]. Nevertheless, data governance frameworks frequently fall behind technical adoption in this region. Limited infrastructure, poor regulatory frameworks, data silos, staff capability shortages, and fragmented health information systems and other contextual-aware issues contribute to ineffective big data governance [12, 13]. Nonetheless, systematic and context-aware big data governance approaches in healthcare systems are lacking [14]. The existing research on big data governance in healthcare has been concentrated on high-income settings, with little attention paid to low- and middle-income countries (LMICs), and even less emphasis on subnational healthcare institutions such as district hospitals [15]. This creates a major knowledge vacuum in the existing studies as big data governance facets that is to say requirements, problems, and implementation techniques are heavily context-dependent. Understanding these contextual variables is critical for developing effective and sustainable governance frameworks in resource-constrained settings like Tanzania healthcare system. Therefore, this study conducts a systematic review of big data governance in healthcare with the aim of synthesizing key governance components, objectives, challenges, and framework types reported in the literature. Furthermore, the study contextualizes these findings to derive practical implications for Tanzania healthcare system. This systematic review is guided by the following research questions (RQ):

RQ1: What are the key components of big data governance in healthcare systems?

RQ2: What objectives do healthcare data governance frameworks aim to achieve?

RQ3: What challenges affect the implementation of big data governance in healthcare?

RQ4: What types of governance frameworks exist in the literature?

RQ5: What gaps exist in current governance models, particularly for low- and middle-income countries?

RQ6: How can big data governance be adapted to the context of Tanzanian district hospitals?

1.1 Study Contribution

By bridging global evidence with local context, this review contributes to the development of robust, context-aware big data governance strategies that can enhance healthcare big data use, improve healthcare outcomes, and strengthen trust in Tanzania's healthcare system. In particular, this study makes three primary contributions to the field of healthcare big data governance. First, it systematically synthesizes existing evidence on big data governance in healthcare by identifying and consolidating key governance components, objectives, challenges, and framework types reported in the relevant literature. This provides a structured understanding of how governance is conceptualized and implemented across different healthcare contexts.

Secondly, the study identifies critical gaps in existing big data governance models, particularly their limited applicability to low- and middle-income countries (LMICs). Current data governance frameworks are designed for well-resourced environments, non-operationalized and fail to adequately address contextual challenges [16, 17]. The third contribution is that the study provides practical implications and proposes a context-aware big data governance model tailored for Tanzania's healthcare system, particularly district hospitals. Unlike generic governance data governance models, the proposed model integrates policy, data quality, security, and contextual adaptation to ensure feasibility and relevance in resource-constrained healthcare settings. This contribution bridges the gap between theoretical governance models and real-world implementation.

The rest of the paper is organized as follows: the next section (section 2) presents the methodology followed for the execution of the study. Section three and four give the results and discussion of the study respectively. The conclusion, limitations, and future research are provided in section 5.

2 Materials and methods

The review adapted systematic literature review (SLR) methodology complemented by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses 2020 (PRISMA 2020) framework [18]. Systematic reviews are well known for their straightforward, organized, and repeatable techniques of synthesizing evidence. Previous research has effectively used systematic methods to identify governance frameworks and components in the field of healthcare governance [19]. Several scholarly reasons support the selection of an SLR. First, systematic reviews make it possible to synthesize empirical data about the advantages and disadvantages of particular techniques or strategies. Second, they make it easier to find study gaps, which helps direct future studies. Thirdly, they offer a well-organized context that facilitates the placement of recent research within the body of existing knowledge [20]. Additionally, an SLR offers a rigorous and standalone methodology for conducting literature reviews [21], while its systematic nature helps minimize bias in the review process [22]. These justifications are directly relevant to this study, which seeks to systematically examine big data governance in healthcare and derive context-specific implications for Tanzania.

It is inevitable not to highlight that the use of an SLR improves methodological rigor and transparency in the review process. Compared to typical narrative evaluations, SLR offer significant benefits. They use a well-established and systematic process to reduce bias in research selection and synthesis, albeit they do not completely eradicate publication bias in the literature [23]. Furthermore, SLR allow for the consolidation of evidence from many empirical methodologies and situations, which improves the generalizability and transferability of results. When supported by reliable investigations, systematic reviews give strong evidence for observed phenomena across settings. In quantitative studies, SLRs also allow for the use of meta-analytic approaches, which can reveal effects that would be difficult to detect in smaller individual studies [23]. These advantages make the SLR approach particularly suitable for analyzing complex and multidisciplinary topics such as big data governance in healthcare. The study selection process followed the PRISMA guidelines, including identification, screening, eligibility, and inclusion stages as summarized in Figure 3.

2.1 Review protocol

This study followed a structured systematic review (SR) protocol based on PRISMA 2020 guidelines [18]. The review protocol was internally developed and approved to ensure methodological consistency and transparency. While it was not registered in external databases such as PROSPERO which is an International Prospective Register of Systematic Reviews, all steps were predefined, including search strategy in a, inclusion/exclusion criteria in b, screening and quality assessment procedures in c and d respectively.

a. Search strategy.

To guarantee thorough coverage of pertinent studies, the literature search was conducted across Scopus, PubMed, IEEE Xplore, ScienceDirect, and Web of Science. These databases together offer access to scholarly publications, peer-reviewed journal articles, and conference proceedings. These databases were selected to capture interdisciplinary viewpoints from computer science, data governance, health informatics, and information systems. Searches were applied to title, abstract, and keywords. The search string in Figure 1 was used with Boolean operators. A combination of keywords and search terms was used as a search strategy to retrieve relevant studies. The keywords were grouped into thematic categories as presented in Figure 2. This aspect of search strategy ensured that relevant literature addressing both global perspectives and context-specific challenges in healthcare data governance was captured.

("big data governance" OR "data governance" OR "health data governance")
AND ("healthcare" OR "health systems" OR "health information systems")

OR was used
to capture
synonyms

Fig. 1. Search string used for literature search. **Source:** Synthesized by the Author(s) (2026)

Big Data	Big data governance, health data
Healthcare	Healthcare data governance
Systems	Health information governance
Context	Developing countries/ LMC

Fig. 2. Search Keywords. **Source:** Synthesized by the Author(s) (2026)

b. Inclusion and exclusion criteria.

To ensure the relevance and quality of the selected studies, predefined inclusion and exclusion criteria were applied as shown in Table 1.

Table 1. Inclusion and exclusion criteria

Criteria type	Inclusion	Exclusion
Publication Year	2010–2025	Before 2010
Language	English	Non-English
Article Type	Peer-reviewed journal & conference papers	Editorials, opinions
Domain	Healthcare	Non-health domains
Focus	Data governance / big data governance	General IT without governance
Availability	Full-text available	Abstract-only

Source: Synthesized by the Author(s) (2026)

c. Screening procedure.

The screening process was conducted in three stages: firstly, title screening which was concerned with Removal of clearly irrelevant studies. Title screening was followed by abstract screening in the second stage in which evaluation of relevance to healthcare governance was conducted. The third stage involved full-text review where final eligibility based on criteria was executed.

To improve reliability, screening criteria were applied consistently across all studies, and ambiguous cases were reviewed multiple times before inclusion.

d. Quality assessment.

To ensure the reliability and validity of the included studies, a quality appraisal checklist was applied based on established SLR guidelines proposed by [21, 24]. This step was essential for evaluating the methodological soundness and relevance of each study before inclusion in the final analysis.

Each selected study was assessed against four key criteria: clarity of research objectives, methodological rigor, relevance to data governance, and strength of evidence presented. A scoring scale ranging from 0 to 2 was used for each criterion, where 0 indicates poor or absent, 1 indicates moderate, and 2 indicates high quality. This resulted in a maximum possible score of 8 for each study. The criteria and scoring scheme are summarized in Table 2. Studies that achieved a minimum score of 5 out of 8 were considered to have sufficient quality and were therefore included in the final review. This threshold ensured that only studies with acceptable methodological rigor and relevance contributed to the synthesis, thereby enhancing the credibility and robustness of the review findings.

Table 2. Quality assessment criteria and scoring scheme used to evaluate the methodological rigor and relevance of included studies

Criterion	Score
Clear objectives	0–2
Methodological rigor	0–2
Relevance to governance	0–2
Evidence support	0–2

Source: Synthesized by the Author(s) (2026)

2.2 Data extraction

Following the selection of eligible studies, a structured data extraction process was conducted to systematically capture key information relevant to the research objectives. A standardized data extraction form was developed to ensure consistency and comparability across studies. The extracted data included bibliographic details (author and year), study context (country and setting), study characteristics (study type), as well as substantive elements such as governance components, framework type, and key findings. This process enabled the identification of common themes, patterns, and gaps in big data governance within healthcare systems, particularly in low- and middle-income countries. The extracted information from the included studies is summarized in Table 3.

Table 3. Summary of Extracted Data from Selected Studies on Big Data Governance in Healthcare Systems (n = 28)

Author	Year	Country	Study Type	Governance Components	Framework Type	Setting	Key Findings
[2]	2014	USA	Conceptual	Data analytics, governance	Conceptual	Hospitals	Big data enhances healthcare decision-making and efficiency
[16]	2018	Global	Review	Interoperability, scalability	Policy framework	LMIC	Highlights infrastructure and scaling gaps in digital health
[17]	2010	LMIC	Empirical	HIS governance, integration	Organizational	Hospitals	Fragmented systems hinder effective implementation
[25]	2017	USA	Empirical	EHR adoption, policy	Policy-driven	Hospitals	HITECH Act significantly improved EHR adoption
[26]	2009	Global	Review	Data quality assessment	Methodological framework	General	Presents structured approaches for improving data quality
[27]	2026	Global	Conceptual	Security, privacy, ethics	Integrated governance	Healthcare systems	Emphasizes emerging risks in big data governance
[28]	2021	Global	Book/Policy	Ethics, equity	Ethical framework	Global health	Focus on ethical challenges in global healthcare governance
[3]	2021	OECD countries	Policy analysis	Data protection, governance	OECD framework	International	Highlights OECD standards in data governance evolution
[14]	2025	LMIC	Systematic review	Digital health systems	Umbrella review	LMIC healthcare	Technology improves care delivery but infrastructure limits persist
[29]	2021	Global	Literature review	Security management (ISO 27001)	Standards-based	Organizations	ISO 27001 improves information security governance maturity
[9]	2019	Global	Review	Big data analytics, governance	Analytical framework	Healthcare systems	Big data supports predictive healthcare and planning
[30]	2024	Global	Empirical	Data governance, data quality	Organizational	Healthcare institutions	Strong governance improves data quality outcomes
[31]	2021	Global	Conceptual	Risk management, security	Security framework	Healthcare systems	Emphasizes structured risk-based governance models
[7]	2024	Global	Conceptual	Privacy, governance, security	Integrated framework	Healthcare systems	Proposes unified governance framework for healthcare data
[32]	2018	Global	Ethical framework	Ethics, accountability, transparency	AI ethics framework	Global	AI governance requires ethical and transparent principles
[5]	2024	Global	Scoping review	Governance structure, policy	Health governance framework	Health systems	Identifies fragmented governance structures globally
[33]	2025	EU and China	Doctoral research	Legal governance, privacy	Legal framework	Digital health	Contrasts regulatory approaches in EU and China
[12]	2024	Tanzania and Germany	Policy analysis	Digital health governance	Comparative framework	National health systems	Tanzania aligns partially with WHO digital health strategy
[34]	2014	OECD	Policy review	Privacy protection	OECD guidelines	International	OECD privacy frameworks influence global standards

[35]	2025	Global	Book chapter	Governance structure, data control	Organizational framework	Healthcare organizations	Governance is critical for data-driven healthcare systems
[36].	2006	USA	Technical report	Access control systems	Security framework	Healthcare IT systems	Strong access control improves data protection
[37]	2026	Global	Systematic review	AI governance maturity	Maturity model	Healthcare AI systems	Proposes maturity model for AI governance in healthcare
[38]	2017	Global	Framework	Data management, governance	Industry standard	Organizations	Provides comprehensive data governance knowledge areas
[39]	2020	Global	Systematic mapping	Security, privacy	Security framework	mHealth systems	Privacy risks remain significant in mobile health systems
[1]	2025	Global	Analytical	Decision-making, governance	Data-driven framework	Healthcare business	Data governance improves strategic healthcare decisions
[40]	2025	Global	Review	Privacy, scalability, accuracy	Big data governance framework	Healthcare systems	Balancing scalability and privacy remain a key challenge
[41]	2018	Global	Empirical	Data quality dimensions	Big data governance model	Healthcare industry	Identifies key dimensions of healthcare data quality
[4]	2010	USA	Conceptual	Governance design, accountability	Governance framework	Enterprises	Foundational model for data governance design

Source: Synthesized by the Author(s) (2026)

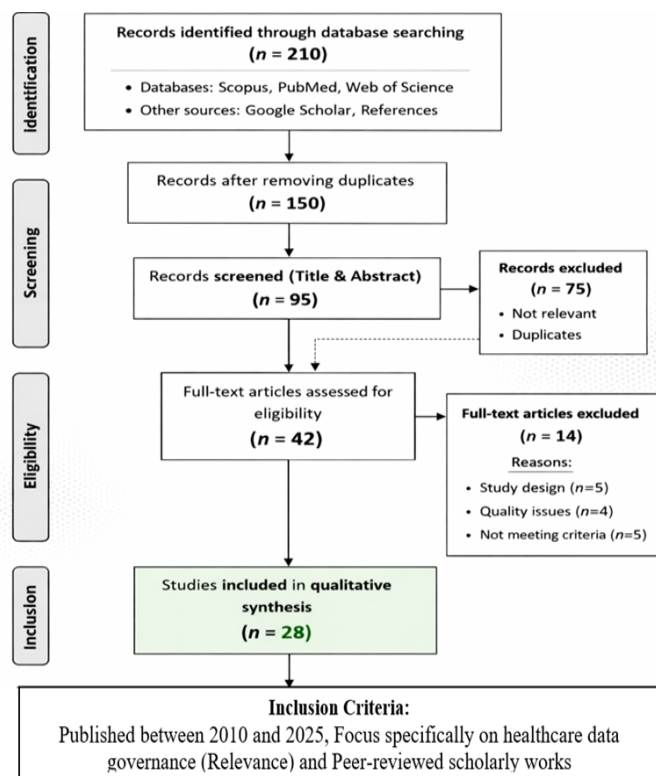


Fig. 3. PRISMA 2020 flow diagram illustrating the study selection. Source: Synthesized by the Author(s) (2026)

The PRISMA 2020 framework was used in the study selection process, as shown in Figure 1, to guarantee reproducibility and transparency in finding pertinent studies. The initial pool of potentially relevant studies consisted of 210 papers that were found through database searches and other sources. There were then 95 papers left for the screening step after duplicate records were eliminated, and obviously unrelated research was eliminated based on their abstracts and titles. 42 studies were found to be eligible after these studies

underwent full-text evaluation to determine their applicability in relation to predetermined inclusion and exclusion criteria.

Subsequently, a detailed quality assessment and methodological evaluation were conducted to ensure the robustness and relevance of the selected studies. As a result, 28 publications were included in the final systematic review. Generally, the study selection process involved multiple stages of filtering and rigorous evaluation to ensure that only high-quality and relevant studies were incorporated, thereby strengthening the validity and reliability of the review findings.

3 Results

This section presents the findings derived from the analysis of the 28 studies included in the SLR.

3.1 Descriptive analysis of selected publications

Key characteristics of the chosen studies, such as publishing trends, database distribution, and theme patterns, as well as governance components, objectives, and problems, among others, were examined to offer a deeper understanding of the reviewed literature. The 28 articles that made up the final list of studies that were part of the systematic review represented a fair distribution of research methodologies and contributions to the field. Based on their principal area of interest, these studies were divided into four major groups. Nine of the studies were conceptual in character, offering theoretical explanations and insights into important ideas and concepts related to the subject of the study. In order to solve the highlighted research difficulties, seven articles suggested frameworks that provided methodological techniques or structured models. Eight empirical studies, including case studies and experimental investigations, presented conclusions derived from actual data collection and analysis. Furthermore, four of the studies were review papers that summarized the body of literature and identified key trends and unmet research needs.

Geographically, the studies came from many parts of the world, including Europe, Asia, Africa, and North America, demonstrating a global study effort. However, the majority of the studies were conducted in industrialized countries, which have a bigger research presence as well as more established technological and institutional infrastructures. Fewer studies focused on the context of developing countries, emphasizing a significant research gap and the need for further context-specific investigations, particularly in low- and middle-income settings.

Figure 4 depicts the classification and geographical distribution of the included studies. It depicts how the 28 selected studies were classified into four categories, as well as the locations from which the studies originated, and the prevalence of research undertaken in developed countries.

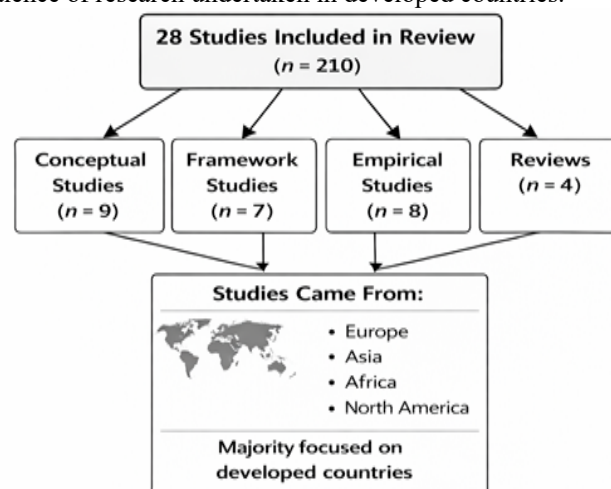


Fig. 4. Classification and geographical distribution of studies included in the systematic review. **Source:** Synthesized by the Author(s) (2026)

Distribution of publications by year: This analysis demonstrates how research interest in big data governance in healthcare has increased over time. Figure 5 shows a strong upward trend in publications, with a peak between 2020 and 2025, indicating increased interest in digital health expansion and AI integration.

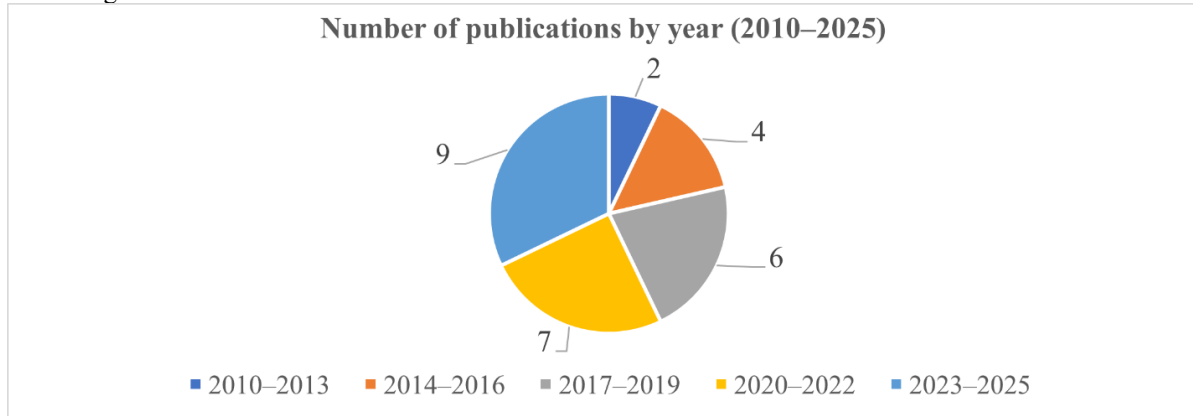


Fig. 5. Number of publications by year (2010–2025). **Source:** Synthesized by the Author(s) (2026)

Distribution by database source: This demonstrates the contribution of each database to the final selected studies. Figure 6 shows that Scopus and PubMed supplied the bulk of studies, demonstrating a high presence of both technical and medical research disciplines.

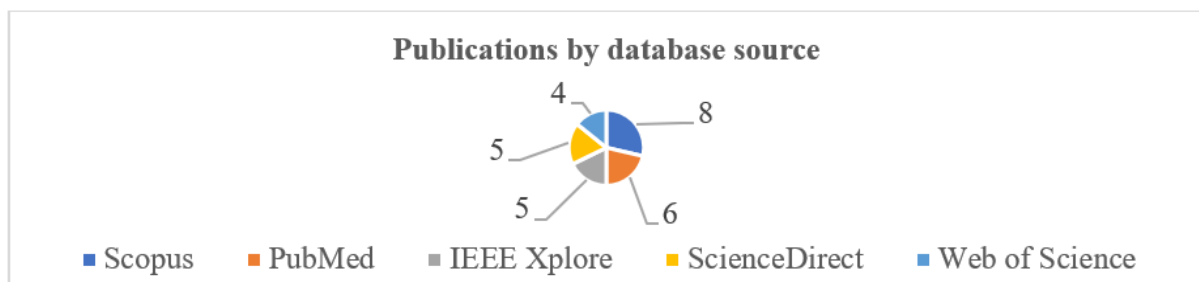


Fig. 6. Publications by database source. **Source:** Synthesized by the Author(s) (2026)

Common big data governance components: Figure 7 shows the frequently referenced big data governance components identified across the 28 studies examined. The components were derived by systematically coding each study for the governance components discussed and tallying their frequency of occurrence across the studies. The numbers shown in the chart represent the number of studies (out of 28) that referenced each component, rather than percentages. Data quality was the most frequently referenced component (26 studies), followed by data privacy and security (25), interoperability (21), regulatory compliance (20), data stewardship (18), data architecture (15), and metadata management (12).

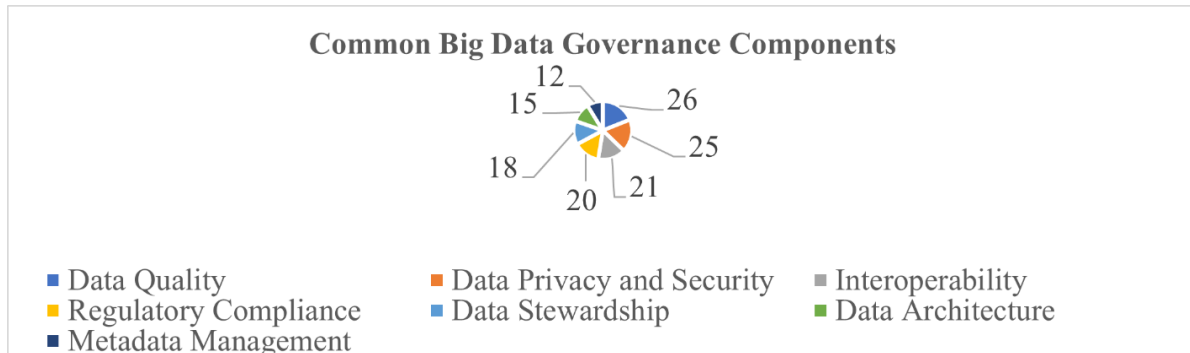


Fig. 7. Common big data governance components. **Source:** Synthesized by the Author(s) (2026)

Data governance objectives: Figure 8 presents the data governance objectives identified across the 28 studies examined. The objectives were derived by systematically coding the stated governance purposes in each study and tallying the frequency with which each objective was reported. The numbers in the chart represent the number of studies (out of 28) that referenced each objective, rather than percentages. Protecting patient privacy and ensuring data quality were the most frequently reported objectives (25 studies each), followed by ensuring regulatory compliance (21), supporting decision-making (20), enabling data sharing (19), and enhancing accountability (17).

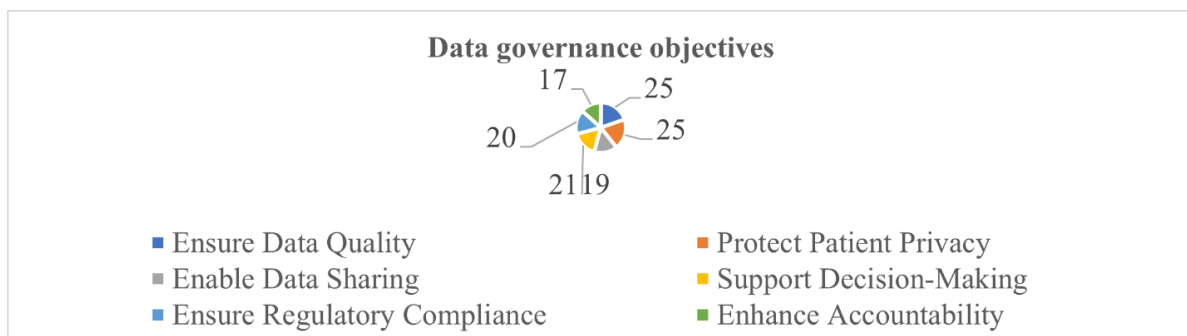
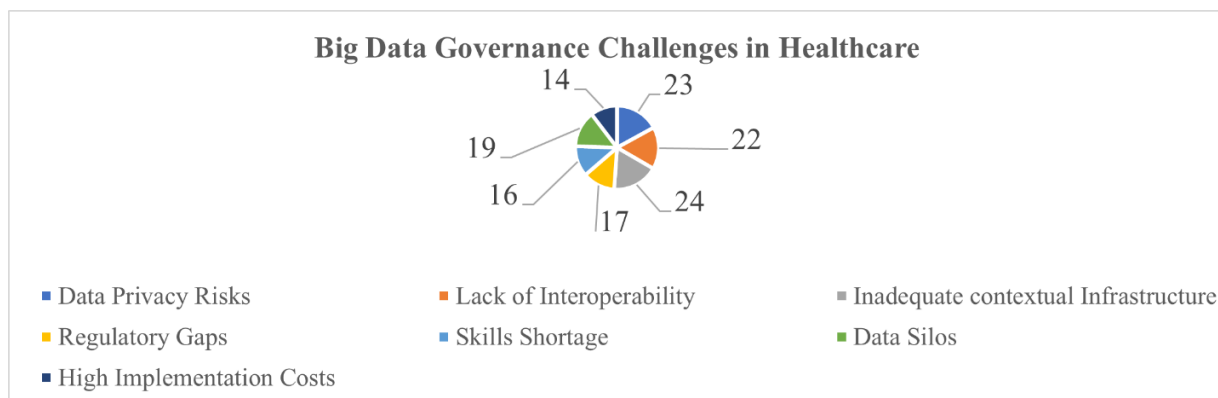


Fig. 8. Data governance objectives. **Source:** Synthesized by the Author(s) (2026)

Big data governance challenges in healthcare: Figure 9 depicts recurring issues reported across the 28 examined studies. The numbers shown in the figure represent the number of studies that reported each respective governance challenge. For example, inadequate contextual infrastructure was reported in 24 of



the 28 studies, making it the most frequently identified challenge and particularly pertinent to low-resource settings.

Fig. 9. Common governance challenges. **Source:** Synthesized by the Author(s) (2026)

Data governance framework types: This taxonomy illustrates the various types of governance frameworks recognized. Figure 10 shows that the majority of research proposes general conceptual frameworks, with few offering practical or non-implementable structures.

3.2 Key insights from selected literature attribute analysis

The preceding analysis yields numerous crucial insights. Research on big data governance in healthcare has increased dramatically in recent years, indicating a growing interest and recognition of its significance. Data governance models frequently prioritize data quality, privacy, and interoperability as essential foundations. However, most extant frameworks are essentially conceptual and developed for highly resourced contexts, limiting their application in low-resource settings. Critical difficulties, such as contextual infrastructure limits, skill gaps, and the persistence of data silos, are especially pertinent to the Tanzanian setting and must be addressed accordingly.

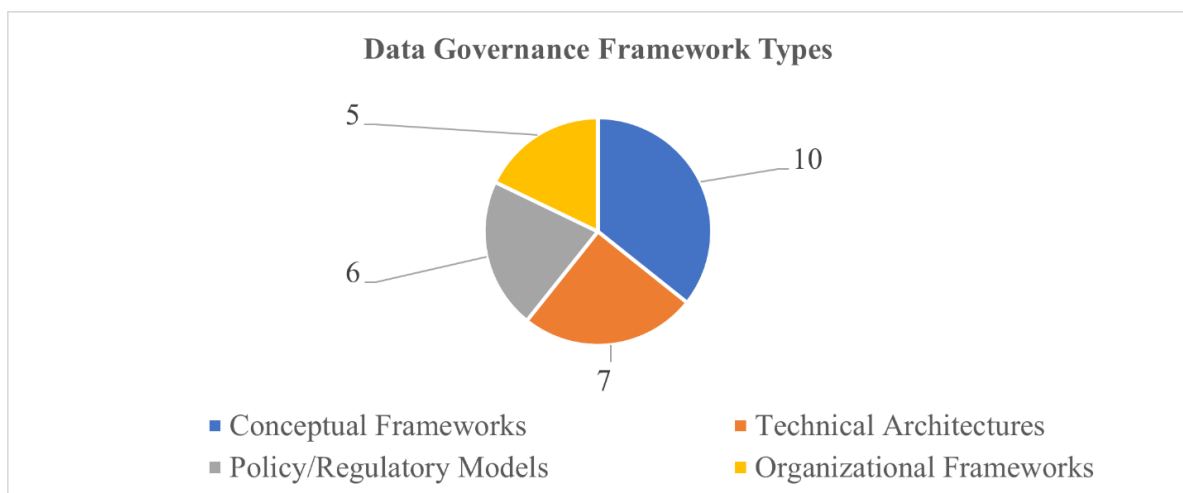


Fig. 10. Types of governance frameworks. **Source:** Synthesized by the Author(s) (2026)

3.2.1 Content findings of the study

The thematic analysis presented in this part is directly informed by the structured data extraction process summarized in Table 4. Each extracted study was systematically coded according to governance components, framework type, study context, and key findings. These coded elements were then synthesized to identify recurring patterns, governance priorities, implementation challenges, and contextual differences across healthcare settings. Table 4 served as the analytical foundation for deriving themes presented in the findings of this study.

A. Big data governance components

The identification of key big data governance components is based on cross-case synthesis of studies presented in Table 4, particularly columns on “governance components” and “key findings” along with author’s interpretations. Big data governance structures in healthcare are made up of fundamental components that ensure data reliability, security, and responsible use throughout its lifecycle. The literature regularly emphasizes data quality, privacy, security, access control, interoperability, and regulatory compliance as vital to effective governance. These components work together to support structured data management techniques, allowing healthcare companies to safely and efficiently use massive datasets.

Governance frameworks often provide policies, standards, and procedures to assure data accuracy, integrity, protection, and appropriate access to sensitive information [5].

Since accurate, comprehensive, and consistent data is crucial for healthcare decisions, data quality is one of these elements. To improve system stability, governance techniques, including data validation, auditing, and standardization, are frequently used. While poor data quality can lead to inaccurate diagnoses and ineffective operations, high-quality data has been demonstrated to enhance clinical decision-making and organizational performance [41]. Because healthcare data is sensitive, security and privacy are equally crucial. To safeguard patient privacy and stop illegal access or disclosure, governance frameworks include safeguards, such as encryption, authentication, and access limitations. According to [39], these protections are crucial for preserving confidence and guaranteeing the ethical use of data in healthcare systems.

Additionally, ensuring that data is used in line with organizational policies and legal requirements depends heavily on access control and regulatory compliance. While compliance frameworks guarantee adherence to laws, rules, and ethical norms, access control techniques, such as role-based access control, govern who can view or alter data. When combined, these elements improve risk management and accountability in intricate big data systems [42].

Data stewardship, metadata management, data architecture, and interoperability are further supporting elements found in the literature that help execute data governance by defining roles, enhancing data comprehension, and permitting system interaction. According to [5], these governance elements are generally constant across many models and offer a solid basis for creating successful big data governance frameworks in healthcare contexts. The fundamental governance elements, their implementation aspects, and related processes in healthcare organizations are compiled in Table 4. Sensitive health information must be protected, since the investigation shows that data quality and interoperability are crucial. The most important governance elements, which are crucial for guaranteeing trustworthy decision-making and efficient data integration across healthcare systems, are data privacy and security.

Table 4. Core big data governance components and implementation dimensions for healthcare organizations

Component	Description	Key Objectives	Governance Mechanisms	Example Technologies/Tools	Risks if Absent	Performance Indicators	Author(s)
Data Quality	Ensures data is accurate, complete, consistent, and reliable	Improve decision-making and analytics accuracy	Data validation, cleansing, standardization, and audits	Data quality tools, ETL systems	Incorrect diagnoses, poor planning	Error rate, completeness %, consistency score	[4], [38]
Privacy	Protects patient confidentiality and personal information	Maintain trust and ethical data use	Anonymization, consent management, privacy policies	Encryption tools, anonymization software	Privacy breaches, loss of trust	Number of breaches, privacy audit scores	[34], [43], [44]
Security	Protects data from unauthorized access or attacks	Safeguard sensitive health information	Encryption, firewalls, intrusion detection	Security platforms, identity management systems	Data theft, cyberattacks	Incident rate, vulnerability score	[29], [31]
Access Control	Regulates who can access data and actions permitted	Ensure authorized data usage	Role-based access control (RBAC), authentication	Access management systems	Unauthorized access	Access violation rate	[36], [38]
Compliance	Ensures adherence to legal and regulatory requirements	Meet regulatory standards	Policies, audits, and reporting procedures	Compliance monitoring tools	Legal penalties	Audit success rate	[4], [34]
Data Stewardship	Defines responsibility for data management	Ensure accountability	Data ownership roles, stewardship policies	Data catalogs	Unclear responsibilities	Stewardship coverage	[4]
Metadata Management	Provides information about data sources and structure	Improve data understanding	Metadata standards, data catalogs	Metadata repositories	Data misinterpretation	Metadata completeness	[38]

Data Lifecycle Management	Controls data from creation to deletion	Ensure efficient data handling	Retention policies, archiving	Database management systems	Data loss	Retention compliance	[38]
Data Integrity	Ensures data remains accurate and unchanged	Maintain data reliability	Validation rules, checksums	Integrity monitoring tools	Corrupted data	Integrity error rate	[4], [29]
Audit and Monitoring	Tracks data usage and changes	Improve accountability	Logging, monitoring systems	Audit software	Undetected misuse	Audit coverage	[29]
Interoperability	Enables data sharing across systems	Improve integration and coordination	Standards and protocols (e.g., HL7, FHIR)	Interoperability platforms	Data silos	Integration success rate	[34], [36]
Risk Management	Identifies and mitigates data-related risks	Reduce threats and vulnerabilities	Risk assessments, mitigation strategies	Risk management tools	Uncontrolled risks	Risk reduction index	[31]

Source: Synthesized by the Author(s) (2026)

B. Big data governance objectives

Big data governance in healthcare is intended to achieve numerous major goals, including trust-building, efficiency improvement, patient safety, and innovation support [33]. A primary goal is to build trust, which is achieved when health data are maintained in a reliable, accurate, and transparent manner. This allows healthcare providers and decision-makers to confidently use data for clinical and administrative objectives [30], [45]. In this context, well-managed data considerably improves decision-making processes and overall healthcare quality [46].

The goal of guaranteeing data quality is closely tied to trust and was identified as one of the most frequently underlined goals throughout the examined studies (25 studies). High-quality data is essential for making precise healthcare decisions, planning effectively, and generating reliable analytics. Similarly, protecting patient privacy (as recognized in 25 studies) is a top governance priority, highlighting the sensitive nature of healthcare data and the importance of maintaining patient confidentiality through adequate rules and controls.

Another significant goal of data governance is to increase efficiency in healthcare workflows. Governance frameworks encourage uniform data collection and management procedures, which reduce duplication of effort and streamline operations in healthcare facilities [47]. This helps to optimize resource utilization and improve service delivery. Furthermore, facilitating decision-making (mentioned in 21 research) is a critical goal, as strong governance ensures that high-quality and easily accessible data are available to inform both clinical and administrative decisions.

Ensuring regulatory compliance (20 studies) and improving accountability (17 studies) are other important goals. Compliance mechanisms ensure that legal, ethical, and institutional standards are met, whereas accountability structures specify roles for data management and utilization, which strengthen governance and monitoring. Furthermore, data sharing (19 studies) is increasingly seen as a critical goal, particularly for facilitating integrated care, research collaboration, and health system coordination.

Patient safety remains a top governance priority, as accurate, timely, and easily available data prevent medical errors, enhance continuity of treatment, and improve clinical outcomes [35]. At the same time, governance mechanisms ensure that data is safely preserved while still being accessible for valid clinical purposes.

Finally, data governance supports innovation by enabling advanced analytics, research, and data-driven healthcare improvements. Well-managed health data facilitates the development of new healthcare services, enhances disease surveillance, and contributes to overall health system strengthening [48, 49].

Ultimately, the findings reveal that the primary goals of healthcare data governance are data quality and privacy protection, emphasizing the sensitive nature of health information. These goals work together to enable the safe, efficient, and ethical use of data, eventually promoting evidence-based decision-making and better healthcare outcomes.

C. Big data governance challenges in healthcare

The big data governance challenges presented in this section were derived from thematic coding of the columns in Table 3, with additional interpretation based on contextual classification of studies for example

LMIC in nexus with high-income settings. The classification of data governance challenges in low-resource settings is strongly supported by studies conducted in LMIC contexts such as [16, 17], as shown in Table 3.

Big data governance in healthcare raises several interconnected issues that affect the successful management and use of health information systems. These problems can be roughly classified into technical, organizational, and ethical dimensions, with each playing an important role in determining governance success. Addressing these issues is critical to guaranteeing data quality, improving decision-making, and promoting dependable healthcare services. Data governance offers the structures and rules required for effective data management [4], whereas healthcare big data environments require robust governance mechanisms to work effectively, especially in resource-constrained settings [2].

Various common issues were noted across the study under consideration. The most common are inadequate contextual infrastructure (24 studies), data privacy issues (23 studies), and a lack of interoperability (22 studies), followed by data silos (19 studies), regulatory gaps (17 studies), skill shortages (16 studies), and high implementation costs (14 studies). These findings show that both technical restrictions and institutional constraints greatly impede effective governance, especially in low-resource healthcare settings.

Technical Challenges: Infrastructure, system integration, and interoperability constraints are the main causes of technical challenges. Healthcare companies frequently run several digital systems that aren't meant to talk to one another, which leads to disjointed data environments and restricted data exchange opportunities. The efficacy of data-driven healthcare services is diminished, and access to complete patient data is limited due to the lack of interoperability seen in 22 research studies [2].

Furthermore, the most commonly mentioned issue (24 research) is insufficient contextual infrastructure, which poses a significant obstacle to efficient governance. Many healthcare facilities lack the computational power, data storage capacity, and dependable internet access needed to handle massive amounts of medical data. Sustainable digital health adoption requires robust digital infrastructure and interoperable systems [50, 51]. Healthcare organizations find it difficult to develop integrated technology infrastructures that facilitate safe data storage and smooth data interchange without these capabilities [52].

Organizational Challenges: Another significant challenge to efficient big data governance is organizational difficulties. According to 16 studies, a significant problem is the lack of qualified staff, which hinders healthcare organizations' capacity to efficiently handle, evaluate, and use massive amounts of data [52]. The availability of qualified experts who can oversee health information systems is crucial for the effective implementation of digital health systems [51].

Furthermore, ambiguous roles, responsibilities, and accountability mechanisms are frequently the consequence of poor governance structures and regulatory loopholes (17 research studies). For data management procedures to be coordinated and consistent, well-defined governance structures are necessary [4]. Healthcare businesses may suffer from disjointed data governance procedures, lower data quality, and insufficient control in their absence. Furthermore, data silos continue to be a major organizational problem (19 studies), making data integration and sharing between departments and organizations even more difficult. Adoption of comprehensive governance systems is also hampered by high implementation costs (14 studies), especially in environments with limited resources. Therefore, enhancing healthcare data governance requires strengthening institutional capability and governance frameworks.

Ethical Challenges: Because health information is sensitive, ethical issues are crucial to the regulation of healthcare data. According to 23 studies, data privacy threats are among the most important issues, highlighting the necessity of robust protections to preserve patient privacy [28], [53]. Data collection, storage, and use must respect individual rights and adhere to legal and regulatory requirements in order to be considered ethical.

Fairness, privacy protection, and data use openness are important ethical factors. According to [32], ethical data governance models must guarantee responsible data practices while prohibiting abuse or prejudice. Fairness guarantees that data-driven decisions do not add prejudice or unfairness, while privacy protection is a basic concept of digital health governance [50]. Conversely, transparency fosters trust by making apparent how data are gathered, handled, and used [37].

In general, the results show that the biggest obstacles to big data governance in healthcare are infrastructure constraints, interoperability problems, and data protection concerns. These difficulties are most noticeable in low-resource environments where organizational, ethical, and technical limitations

collide. Developing efficient, context-aware governance systems that can enable trustworthy, safe, and fair use of healthcare data requires addressing these problems.

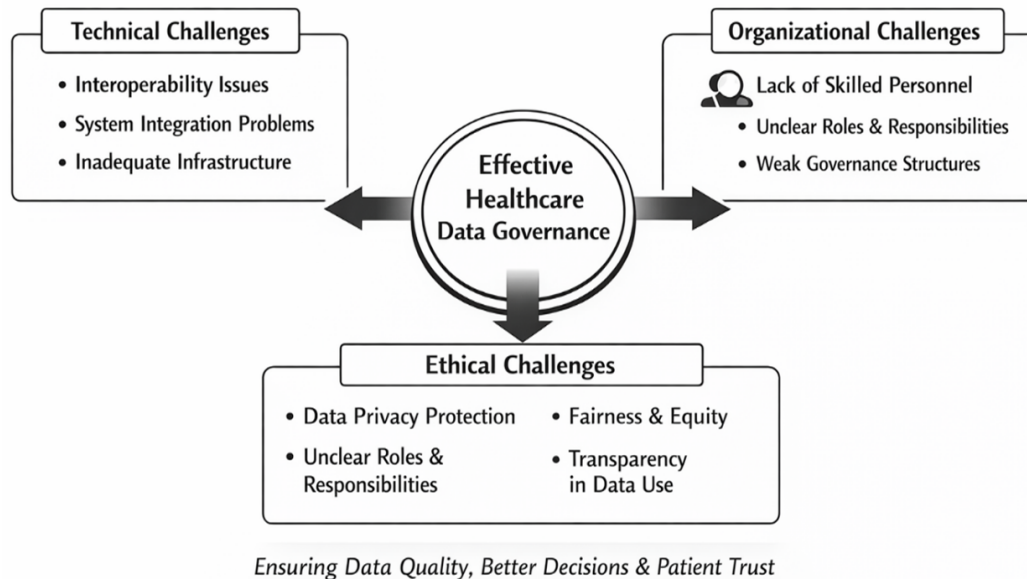


Fig. 11. The Summary of big data governance challenges in healthcare. **Source:** Synthesized by the Author(s) (2026)

D. Data governance framework types

The categorization of data governance frameworks into conceptual, policy, technical, and organizational types was derived from the “Framework Type” column in Table 4 and author’s interpretations. Depending on their main purpose and method of implementation, data governance frameworks in the healthcare industry can be roughly divided into three types: policy-based, technical, and organizational. These frameworks offer organized methods for handling health data and guaranteeing its efficient, safe, and responsible use. The intricacy of healthcare information systems frequently necessitates the integration of several governance approaches. Governance frameworks, on the other hand, provide rules, responsibilities, and procedures that guide data management within businesses [4]. Across the reviewed studies, different types of governance frameworks were identified, including conceptual frameworks (10 studies), technical architectures (7 studies), policy or regulatory models (6 studies), and organizational frameworks (5 studies). This distribution indicates that most studies focus on high-level conceptual approaches, with relatively fewer providing practical or fully implementable architectures.

Policy-based frameworks are primarily concerned with rules, standards, and formal guidelines that control the collection, use, and exchange of healthcare data. These frameworks specify the legal and regulatory standards that healthcare organizations must follow to maintain data security, privacy, and appropriate information exchange [54]. National digital health policies generally serve as the foundation for such frameworks, defining criteria for system installation and data administration. Tanzania's national digital health initiatives provide policies for the use of information technology in healthcare services, with the Ministry of Health highlighting its role in leading the creation and control of health information systems [55, 56]. As a result, policy-based frameworks play an important role in guaranteeing compliance and supporting standardized data management methods.

The technology infrastructure needed to support healthcare data governance is the main emphasis of technical frameworks. The systems, instruments, and standards required for data gathering, storing, processing, and sharing are specified by these frameworks.

Data architectures, interoperability standards, and system integration models that enable effective data sharing across platforms are essential components. Because they offer the infrastructure required to handle vast and complicated datasets, technical frameworks are crucial for supporting big data analytics in the healthcare industry [2]. Additionally, the successful adoption of digital health depends on standardized data structures and interoperable health information systems [50]. Therefore, the dependability and effectiveness of healthcare information systems are supported by technical frameworks.

The institutional structures, roles, and operational procedures that support data governance in healthcare organizations are covered by organizational frameworks [57]. These frameworks lay forth protocols for handling healthcare data, provide accountability systems, and specify roles for data stewardship. To ensure accountability and coordination in data management, clearly defined responsibilities and decision-making procedures are crucial [4]. Effective digital health governance necessitates well-organized organizational procedures and clearly defined roles, according to the [50]. Therefore, organizational frameworks guarantee that data management procedures are coordinated, consistent, and in line with institutional goals.

In summary, the results indicate that although there are many different kinds of governance frameworks, conceptual models predominate, with fewer studies providing specific organizational or technical implementations. This demonstrates a gap between theoretical advancement and real-world implementation, especially in healthcare settings with limited resources where context-aware and implementable governance frameworks are crucial.

4 Discussion and Interpretation of the Key Findings

This study provides a comprehensive synthesis of big data governance in healthcare by examining governance components, objectives, challenges, and framework types, while highlighting their implications for Tanzania's healthcare system. The findings demonstrate that healthcare data governance is inherently multi-dimensional and requires alignment between technical infrastructure, organizational processes, and regulatory frameworks. Effective governance cannot be achieved through isolated interventions, but rather through integrated systems that collectively ensure data quality, privacy, security, interoperability, and regulatory compliance [4] [6]. These governance components directly support broader healthcare goals such as patient safety, trust, efficiency, innovation, and improved decision-making. For instance, strong data quality mechanisms enhance clinical decisions and healthcare outcomes, while privacy and security measures strengthen patient trust and confidentiality [4], [8]. Similarly, interoperability and data-sharing capabilities facilitate integrated care delivery and improve system-wide efficiency [2]. This suggests that governance mechanisms are not merely technical requirements, but foundational enablers of healthcare performance and service delivery.

Despite this conceptual alignment, the findings reveal a persistent gap between theoretical governance models and practical implementation, particularly in LMICs. The literature is dominated by conceptual frameworks because data governance remains an evolving field in which researchers often prioritize theoretical generalization over practical deployment. Furthermore, implementation and empirical validation require substantial financial resources, access to operational healthcare systems, and the ability to manage highly sensitive health data, all of which are constrained by ethical, legal, and institutional barriers [8]. Consequently, organizational and technical implementations remain limited despite the abundance of conceptual and policy-oriented frameworks. Existing policy-based frameworks provide important regulatory direction, but their effectiveness is often undermined by inadequate institutional capacity and weak technical infrastructure. Likewise, technical frameworks cannot function effectively without organizational structures that clearly define governance roles, accountability mechanisms, and operational responsibilities [2], [4]. These findings highlight the importance of integrated governance approaches that combine policy, organizational, and technical dimensions.

The review further demonstrates that LMICs and district-level healthcare settings remain significantly underrepresented in healthcare data governance research. This underrepresentation is largely associated with limited research funding, weak digital infrastructure, restricted access to healthcare data, and insufficient institutional research capacity. In many cases, healthcare governance studies focus predominantly on national systems or large tertiary hospitals, while district hospitals receive comparatively little attention [16]. This creates a major gap in localized governance understanding, particularly in contexts where healthcare systems operate under severe resource constraints. In LMICs, governance implementation is additionally hindered by technical limitations such as fragmented systems, poor interoperability, inadequate infrastructure, and limited human resource capacity [50], [17]. Ethical concerns related to data privacy, fairness, and balancing accessibility with patient confidentiality further complicate governance efforts [32], [58].

The findings also emphasize the critical role of contextual factors in determining governance effectiveness. Most existing frameworks were developed in high-income environments and therefore

assume the availability of advanced ICT infrastructure, skilled personnel, and mature regulatory systems. However, Tanzania's district hospitals operate under substantially different conditions characterized by fragmented information systems, including parallel use of DHIS2 and paper-based systems, limited ICT infrastructure, shortages of skilled personnel, and high patient loads. These realities make many sophisticated governance frameworks impractical within local healthcare settings. Studies indicate that digital health initiatives in low-resource environments frequently fail because they do not adequately reflect local institutional capacities and operational realities [16, 17]. This underscores the need for governance frameworks that are simplified, adaptive, and context-aware.

In response to these identified gaps, the proposed governance model contributes both theoretically and practically to healthcare data governance research. The model in Figure 12 integrates policy, technical, and organizational dimensions while introducing context-awareness as a central governance component. Unlike many existing frameworks, it emphasizes feasibility, scalability, and alignment with local workflows and resource limitations. Theoretically, the study extends data governance theory by incorporating contextual realities associated with LMIC healthcare environments and bridging the gap between governance theory and implementation practice. It also provides a structured synthesis of existing governance literature within healthcare contexts. Practically, the proposed model offers a ready-to-adapt governance structure that can support policy formulation, strengthen healthcare data management practices, and enhance data-driven decision-making at district hospital level in Tanzania. Overall, the findings suggest that effective big data governance in healthcare requires holistic, integrated, and context-sensitive approaches capable of bridging the persistent divide between theoretical frameworks and practical implementation realities

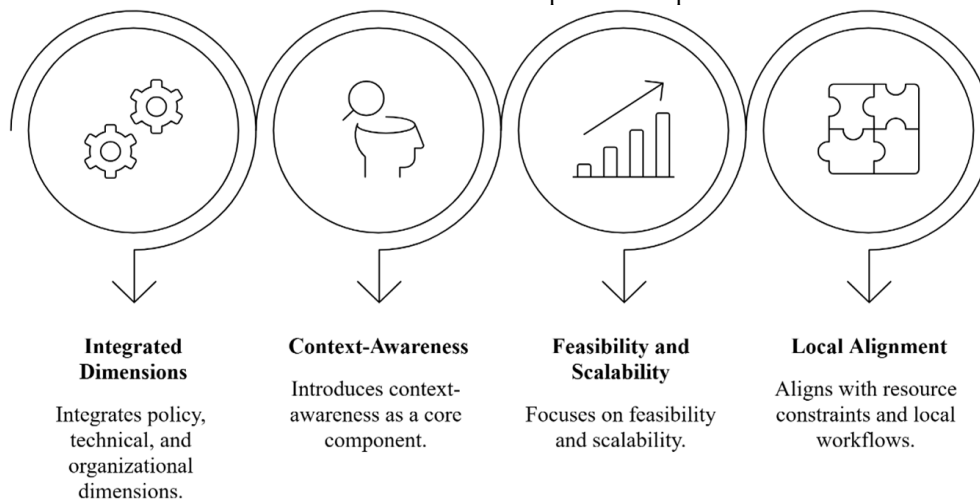


Fig. 12. Proposed context-aware big data governance model for healthcare systems in resource-constrained settings. **Source:** Synthesized by the Author(s) (2026)

4.1 Implications for district hospitals in Tanzania and research gaps

The findings of this study demonstrate that effective data governance in Tanzanian district hospitals depends on several interrelated requirements. First, clear policies and regulatory frameworks are essential for defining governance standards, institutional responsibilities, and accountability mechanisms, thereby promoting consistent data management practices across healthcare facilities [4]. Second, training and capacity building are critical for enabling healthcare workers to effectively manage and utilize health data. Existing evidence emphasizes that the availability of skilled personnel is fundamental to the successful implementation of digital health systems [50]. Without adequate technical and managerial competencies, even well-structured governance frameworks may fail to achieve their intended outcomes.

In addition, effective governance requires adequate ICT infrastructure to support data collection, storage, integration, and exchange. However, many LMICs, including Tanzania, continue to experience infrastructural limitations such as unstable connectivity, inadequate technological resources, and fragmented information systems, which constrain governance implementation [16, 17]. The adoption of

standardized data formats and procedures is equally important for ensuring interoperability, improving data quality, and supporting integrated healthcare delivery across multiple health information systems [1].

Based on these findings, this study proposes a context-aware governance model specifically designed for Tanzanian district hospitals. The model is built around four key dimensions: policies, data quality, security, and context awareness. Policies provide the institutional and regulatory foundation for governance, while data quality ensures that healthcare information remains accurate, complete, and timely for effective decision-making. Security mechanisms safeguard sensitive patient data against unauthorized access and misuse, thereby supporting trust and ethical compliance [8]. Context awareness represents the distinguishing feature of the proposed model, ensuring that governance mechanisms are aligned with local realities such as limited resources, infrastructural constraints, and workforce capabilities. This adaptive approach has been recognized as essential for successful digital health implementation in resource-constrained environments [16].

To operationalize the proposed model, the study introduces a three-layered governance architecture. The first layer consists of core data sources, including Health Management Information Systems (HMIS), Electronic Health Records (EHR), and laboratory information systems, which generate healthcare data necessary for service delivery and decision-making [61]. The second layer comprises governance mechanisms such as policies, standards, and clearly defined roles and responsibilities that regulate data management and ensure compliance with governance requirements [4]. The final layer focuses on application and decision-support functions, where governed data are utilized to strengthen clinical decision-making, planning, and performance monitoring [45] [64, 65]. This integrated architecture reflects established best practices in health information systems by linking data generation, governance, and utilization within a unified framework [50].

The review further identifies several critical research gaps in healthcare big data governance. First, there is limited research focusing on African healthcare contexts, particularly within sub-Saharan Africa, despite the unique institutional and infrastructural challenges affecting governance implementation in these settings [4], [51]. Second, existing studies rarely develop context-aware governance models capable of addressing the operational realities of low-resource healthcare environments. Many current frameworks remain highly generalized and difficult to implement in practice, thereby limiting their effectiveness and practical relevance [16] [59]. Third, district-level healthcare facilities remain underrepresented in governance research, as most studies concentrate on national health systems or large tertiary hospitals, despite district hospitals serving as primary healthcare access points for a significant proportion of the population in countries such as Tanzania.

Finally, there is a shortage of practical and operational data governance models that translate theoretical principles into implementable strategies [60, 61]. Although many studies acknowledge the importance of healthcare data governance, few provide validated implementation frameworks capable of guiding real-world practice [62, 63]. Addressing these gaps is essential for the development of scalable, context-sensitive, and operational data governance frameworks that can effectively support healthcare systems in resource-constrained settings.

5 Conclusion, Limitations, and Future Research

To synthesize knowledge on data governance components, objectives, challenges, and framework types, and implications for the Tanzania healthcare system, this study conducted a comprehensive assessment of the literature on big data governance in healthcare. The results verify that data governance is essential to guaranteeing data security, privacy, quality, and interoperability, all of which promote efficiency, trust, and better healthcare outcomes. But even with this solid conceptual basis, the majority of current governance strategies are still mostly theoretical and not properly modified for real-world application, especially in settings with limited resources. The ongoing disconnect between governance theory and practical implementation is a crucial finding from this review. Particularly in low- and middle-income nations, existing frameworks sometimes overlook contextual issues including inadequate infrastructure, a lack of workers, and institutional limitations. To close this gap, the study suggested a layered architecture specifically designed for Tanzanian district hospitals and a context-aware governance model that integrated policy, data quality, security, and contextual adaptation as essential elements. Overall, by encouraging a more integrated and practical approach to healthcare data governance and stressing congruence across

technical, organizational, and policy elements while being attentive to local settings, the study advances both theory and practice. For district hospitals to support resilient health systems, improve clinical decision-making, and increase data reliability, data governance must be strengthened. Notwithstanding these contributions, some limitations need to be taken into account. First, the study's capacity to accurately represent actual procedures in district hospitals is constrained by the fact that it is mostly literature-based and dependent on secondary data sources. Second, because there was no fieldwork, important contextual elements, including organizational culture, staff capabilities, and infrastructure limitations, were not properly evaluated, which limited the findings' application to particular Tanzanian situations. Third, the practical viability of the suggested governance model was not established because it was not empirically validated through implementation or stakeholder involvement. Furthermore, the investigation was limited to publications published between 2010 and 2025, which may have eliminated pertinent older works. It's also possible that important studies found in physical libraries or institutional archives were overlooked due to the dependence on digital databases and online search engines. Additionally, the depth of analysis may have been limited by the exclusive use of an SLR technique, which restricted data collection to titles, abstracts, and keywords. Future research should prioritize developing realistic and context-sensitive big data governance solutions for district hospitals. This includes working with healthcare stakeholders to determine particular local requirements, restrictions, and goals, which will guide the building of bespoke governance structures. Subsequent research should focus on real-world application and testing of these frameworks in healthcare settings to assess their usability and viability. Rigorous evaluation is required to assess effectiveness, uncover issues, and facilitate iterative development. Furthermore, future research should prioritize the validation and operationalization of context-aware governance models, particularly in African and other low-resource contexts, to close the gap between conceptual frameworks and practical application.

Acknowledgements

None

Statement on conflicts of interest

None

References

- [1] Iyer, S. S. (2025). Data-driven decision making: The key to future health care business success. *RA Journal Of Applied Research (RAJAR)*, 11(3), 115-136.
- [2] Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: promise and potential. *Health information science and systems*, 2(1), 3.
- [3] Canton, H. (2021). Organisation for economic co-operation and development—OECD. In *The Europa directory of international organizations 2021* (pp. 677-687). Routledge.
- [4] Khatri, V., & Brown, C. V. (2010). Designing data governance. *Communications of the ACM*, 53(1), 148-152.
- [5] [Ghaffari Heshajin, S., Sedghi, S., Panahi, S., & Takian, A. (2024). A framework for health information governance: a scoping review. *Health Research Policy and Systems*, 22(1), 109.
- [6] Otto, B. (2011). Organizing data governance: Findings from the telecommunications industry and consequences for large service providers. *Communications of the Association for Information Systems*, 29(1), 3.
- [7] Faridoon, A., & Kechadi, M. T. (2024, February). Healthcare data governance, privacy, and security: a conceptual framework. In *EAI International Conference on Body Area Networks* (pp. 261-271). Cham: Springer Nature Switzerland.
- [8] Keshta, I., & Odeh, A. (2021). Security and privacy of electronic health records: Concerns and challenges. *Egyptian Informatics Journal*, 22(2), 177-183.
- [9] Dash, S., Shakyawar, S. K., Sharma, M., & Kaushik, S. (2019). Big data in healthcare: management, analysis, and prospects. *Journal of Big Data*, 6(1), 54.
- [10] Odionu, C. S., & Ibeh, C. V. (2023). Big data analytics in healthcare: A comparative review of USA and global use cases. *Journal Name*, 4(6), 1109-1117.

- [11] Mwogosi, A. (2026). Revolutionizing primary health care in Tanzania: unravelling the contextual factors on electronic health record systems implementation for effective decision support. *Journal of Science and Technology Policy Management*, 17(3), 642-663.
- [12] Holl, F., Kircher, J., Hertelendy, A. J., Sukums, F., & Swoboda, W. (2024). Tanzania's and Germany's digital health strategies and their consistency with the World Health Organization's global strategy on digital health 2020-2025: comparative policy analysis. *Journal of Medical Internet Research*, 26, e52150.
- [13] Tun, S. Y. Y., & Madanian, S. (2023). Clinical information system (CIS) implementation in developing countries: requirements, success factors, and recommendations. *Journal of the American Medical Informatics Association*, 30(4), 761-774.
- [14] Craig, A., Lawford, H., Miller, M., Chen-Cao, L., Woods, L., Liaw, S. T., & Godinho, M. A. (2025). Use of technology to support health care providers delivering care in low-and lower-middle-income countries: systematic umbrella review. *Journal of Medical Internet Research*, 27, e66288.
- [15] [Knight, S. R., Ots, R., Maimbo, M., Drake, T. M., Fairfield, C. J., & Harrison, E. M. (2019). Systematic review of the use of big data to improve surgery in low-and middle-income countries. *Journal of British Surgery*, 106(2), e62-e72.
- [16] Labrique, A. B., Wadhwani, C., Williams, K. A., Lampthey, P., Hesp, C., Luk, R., & Aerts, A. (2018). Best practices in scaling digital health in low and middle income countries. *Globalization and health*, 14(1), 103.
- [17] Fraser, H. S., & Blaya, J. (2010, November). Implementing medical information systems in developing countries: what works and what doesn't. AMIA Annu Symp Proc. 2010 Nov 13;2010:232-6. PMID: 21346975; PMCID: PMC3041413.
- [18] Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372.
- [19] Ngesimani, N. L., Ruhode, E., & Harpur, P. A. (2022). Data governance in healthcare information systems: A systematic literature review. *South African Journal of Information Management*, 24(1), 1475.
- [20] Keele, S. (2007). *Guidelines for performing systematic literature reviews in software engineering* (Vol. 5). Technical report, ver. 2.3 ebse technical report. ebse.
- [21] Okoli, C. (2015). A guide to conducting a standalone systematic literature review. *Communications of the association for information systems*, 37.
- [22] Siddaway, A. (2014). What is a systematic literature review, and how do I do one? *University of Stirling*, 1(1), 1-13.
- [23] Ryan, G. (2010). Guidance notes on planning a systematic review. *James Hardiman Library*, 6.
- [24] Brereton, P., Kitchenham, B. A., Budgen, D., Turner, M., & Khalil, M. (2007). Lessons from applying the systematic literature review process within the software engineering domain. *Journal of systems and software*, 80(4), 571-583.
- [25] Adler-Milstein, J., & Jha, A. K. (2017). The HITECH Act drove large gains in hospital electronic health record adoption. *Health affairs*, 36(8), 1416-1422.
- [26] Batini, C., Cappiello, C., Francalanci, C., & Maurino, A. (2009). Methodologies for data quality assessment and improvement. *ACM computing surveys (CSUR)*, 41(3), 1-52.
- [27] Bena, Y. A., Muchtar, F., Ibrahim, R., Mahmood, J., Chan, W. H., Shah, M. Z. M. Z., ... & Fattah, S. (2026). Enhancing Big Data Governance Practices: Addressing Security, Privacy, and Ethical Challenges. *Journal of Advanced Research Design*, 142(1), 159-176.
- [28] Benatar, S., & Brock, G. (Eds.). (2021). *Global health: ethical challenges*. Cambridge University Press.
- [29] Culot, G., Nassimbeni, G., Podrecca, M., & Sartor, M. (2021). The ISO/IEC 27001 information security management standard: literature review and theory-based research agenda. *The TQM Journal*, 33(7), 76-105.
- [30] Deghati, S. (2024). Impact of Data Governance on Data Quality in Healthcare Institutions. *Am. J. Data Inf. Knowl. Manag.*, 5, 39-48.
- [31] Dias, F. M., Martens, M. L., de Paula Monken, S. F., da Silva, L. F., & Santibanez-Gonzalez, E. D. R. (2021). Risk management focusing on the best practices of data security systems for healthcare. *International Journal of Innovation*, 9(1), 45-78.
- [32] Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and machines*, 28(4), 689-707.
- [33] He, Z. (2025). *Health Data Governance in the Era of Digital Health: Legal Approaches of the EU and China* (Doctoral dissertation, Department of Law, Stockholm University).
- [34] Horodyski, D. (2014). 2013 OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data as an Example of Recent Trends in Personal Data Protection. *Collective human rights*, 255.
- [35] Housawi, A., & Lytras, M. D. (2025). Data governance in healthcare organizations. In *Next Generation eHealth* (pp. 13-32). Academic Press.
- [36] Hu, V. C., Ferraiolo, D., & Kuhn, D. R. (2006). *Assessment of access control systems* (Vol. 76). Gaithersburg, MD: US Department of Commerce, National Institute of Standards and Technology.
- [37] Hussein, R., Zink, A., Ramadan, B., Howard, F. M., Hightower, M., Shah, S., & Beaulieu-Jones, B. K. (2026). Advancing healthcare AI governance through a comprehensive maturity model based on a systematic review. *npj Digital Medicine*.
- [38] International, D. (2017). *DAMA-DMBOK: Data management body of knowledge*. Technics Publications, LLC.

- [39] Iwaya, L. H., Ahmad, A., & Babar, M. A. (2020). Security and privacy for mHealth and uHealth systems: a systematic mapping study. *IEEE Access*, 8, 150081-150112.
- [40] Javeedullah, M. (2025). Big Data and Health Informatics: Managing Privacy, Accuracy, and Scalability. *Global Trends in Science and Technology*, 1(3), 29-47.
- [41] Juddoo, S., George, C., Duquenoy, P., & Windridge, D. (2018). Data governance in the health industry: investigating data quality dimensions within a big data context. *Applied System Innovation*, 1(4), 43.
- [42] Sargiotis, D. (2024). Overview and importance of data governance. In *Data Governance: A Guide* (pp. 1-85). Cham: Springer Nature Switzerland.
- [43] Rindfleisch, T. C. (1997). Privacy, information technology, and health care. *Communications of the ACM*, 40(8), 92-100.
- [44] Ronchi, E., & Reimsbach-Kounatze, C. (2022). A decade and a half of OECD action on data governance policy-making. In *Annales des Mines-Réalités industrielles* (Vol. 2022, No. 3, pp. 71-74). Institut Mines-Télécom.
- [45] World Health Organization. (2025a). *Global strategy on digital health 2020-2027*. World Health Organization.
- [46] Orlu, G. U., Abdullah, R. B., Zaremohzzabieh, Z., Jusoh, Y. Y., Asadi, S., Qasem, Y. A., ... & Mohd Nasir, W. M. H. B. (2023). A Systematic Review of Literature on Sustaining decision-making in healthcare organizations amid imperfect information in the big data era. *Sustainability*, 15(21), 15476.
- [47] D. Tse, C. -k. Chow, T. -p. Ly, C. -y. Tong and K. -w. Tam. (2018). "The Challenges of Big Data Governance in Healthcare," *2018 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)*, New York, NY, USA, 2018, pp. 1632-1636, doi: 10.1109/TrustCom/BigDataSE.2018.00240
- [48] Oktaviana, S., Handayani, P. W., Hidayanto, A. N., & Siswanto, B. B. (2025). Healthcare data governance assessment based on hospital management perspectives. *International Journal of Information Management Data Insights*, 5(1), 100342.
- [49] Nomura, S., Sakamoto, H., Ishizuka, A., Katsuma, Y., Akashi, H., & Miyata, H. (2020). Ongoing debate on data governance principles for achieving Universal Health Coverage: a proposal to post-G20 Osaka Summit meetings. *Global Health Action*, 13(1), 1859822.
- [50] World Health Organization. (2021). *Guidance for health information system governance* (No. WHO/EURO: 2021-1999-41754-57182). Guidance for health information system governance.
- [51] World Health Organization. (2025b). *Health data governance in the age of artificial intelligence: policy imperatives for the WHO European Region* (No. WHO/EURO: 2025-11462-51234-78079 (PDF)). World Health Organization. Regional Office for Europe.
- [52] Mwogosi, A., Mallya, M., Asenga, S. J., Mchome, A. J., Audi, M. J., & Gunda, J. J. (2025). The impact of digital health records systems on workflow and service efficiency in Tanzania: A systematic review and framework development. *Health Information Management Journal*, 18333583251363641.
- [53] Maseme, M. (2022). Ethical considerations for health research data governance. In *Data Integrity and Data Governance*. IntechOpen.
- [54] Mwogosi, A., & Simba, R. (2026). Digital policy and governance frameworks for EHR systems in Tanzania: a scoping review. *Digital Policy, Regulation and Governance*, 28(1), 52-74.
- [55] Li, T., Wandella, A., Gomer, R., & Al-Mafazy, M. H. (2024). Operationalizing health data governance for AI innovation in low-resource government health systems: a practical implementation perspective from Zanzibar. *Data & Policy*, 6, e63.
- [56] Watts, G. (2020). The Tanzanian digital health agenda. *The Lancet Digital Health*, 2(2), e62-e63
- [57] Korhonen, J. J., Melleri, I., Hiekkanen, K., & Helenius, M. (2013). Designing data governance structure: an organizational perspective. *GSTF Journal on Computing*, 2(4), 11-17.
- [58] Shaw, J., & Sekalala, S. (2023). Health data justice: building new norms for health data governance. *NPJ Digital Medicine*, 6(1), 30.
- [59] Osamika, D., Adelusi, B. S., Kelvin-Agwu, M. T. C., Mustapha, A. Y., Forkuo, A. Y., & Ikhalea, N. (2025). A critical review of health data interoperability standards: FHIR, HL7, and beyond. *World Scientific News*, 203, 194-233.
- [60] Noor, M., Baharom, F., & Mohd, H. (2025). Big Data Governance Framework: Current and Future Trends. *Journal of Research and Digital Innovation*, 1(1), 26-36.
- [61] Khan, J. (2025). Ensuring Data Accuracy and Uniformity in Real-Time ETL for Streaming Systems: A Comparative Study of Contemporary ETL Frameworks.
- [62] Mamuye, A. L., Yilma, T. M., Abdulwahab, A., Broomhead, S., Zondo, P., Kyeng, M., ... & Tilahun, B. C. (2022). Health information exchange policy and standards for digital health systems in Africa: a systematic review. *PLOS Digital Health*, 1(10), e0000118.
- [63] Nauman, M., Almadhor, A. S., Albekairi, M., Ansari, A. R., Fayyaz, M. A., & Nawaz, R. (2025). The role of big data analytics in revolutionizing diabetes management and healthcare decision-making. *IEEE Access*, 13, 10767-10785.
- [64] O 'Neil, S., Taylor, S., & Sivasankaran, A. (2021). Data equity to advance health and health equity in low-and middle-income countries: a scoping review. *Digital health*, 7, 20552076211061922.
- [65] Zhang, Y., Callaghan-Koru, J. A., & Koru, G. (2024). The challenges and opportunities of continuous data quality improvement for healthcare administration data. *JAMIA open*, 7(3), ooae058.