

UNIVERSITY MIDOCEAN
College of Management
Master of Health Administration



"A Research Project Titled:"
Analyzing Data Silos Challenges in Public Hospitals in Saudi Arabia and
Proposing AI-Based Solutions Using Open Health Data"

Prepared by the Student:

Osama Jehad Alostaz

Supervisor

Dr. AL Faisal Abdelhamid Mohamed

Dr. Fayez Ali Jrad

Academic Year: 2025/2026

All printing rights are reserved by the university.

Acknowledgment

This research project, titled “Analyzing Data Silos Challenges in Public Hospitals in Saudi Arabia and Proposing AI-Based Solutions Using Open Health Data,” was prepared by Osama Jehad Alostaz under the supervision of Dr. AL Faisal Abdelhamid Mohamed and Dr. Fayez Ali Jrad.

The researcher would like to express his deepest appreciation to his supervisors for their continuous guidance, valuable insights, and constructive feedback throughout all stages of this study. Their expertise and encouragement have been instrumental in shaping the research direction and ensuring its academic rigor.

Sincere gratitude is also extended to all participants and professionals who contributed their time and knowledge to the data collection process, and to the academic institution for providing the resources and support necessary to complete this project successfully.

Table of Contents	
Abstract	4
Chapter One: Introduction	
1.1Introduction	8
1.2 Problem Statement	9
1.3Research Theoretical	11
1.4 Study Objectives:	13
1.5Research Hypotheses	13
1.6Research Model	14
1.7The Scientific Theoretical Framework Adopted in the Study	14
Chapter Two: Theoretical Framework and Literature Review	
<u>2.1 Introduction</u>	19
<u>2.2Theoretical Framework</u>	20
<u>2.3 Previous Studies</u>	32
2.3. 1 Studies on Data Silos and AI Integration in Healthcare	33
2.3.2 Studies on Open Health Data and Interoperability	39

2.3.3 Integrated Frameworks and Comparative Insights for Healthcare Transformation	42
2.4 summary of the Theoretical and Empirical Frameworks	45
2.5 Summary of Chapter Two	47
Chapter Three: Methodology	
3.1 Introduction:	50
3.2 Research Design	50
3.3 Research Population and Sample	51
3.4 Research Variables	53
3.5 Research Instrumentation	54
3.6 Data Collection Procedures	56
3.7 Data Analysis Techniques	57
3.8 Study Boundaries	58
Chapter Four: Results and Recommendations	
4.1 Introduction	61
4.2 Descriptive Statistics for Questionnaire Dimensions	61
4.3 Hypotheses Testing :	70
4.4 Discussion of Findings	75
4.5 Recommendations	78
References	80
Appendices	84

List of Tables

Table Title	Page
Table (3.1): Demographic Characteristics of the Study Participants	52
Table 3.2Cronbach's Alpha Coefficients and Reliability Levels for the Study Instrument Dimensions	56
Table (4. 1): Descriptive Statistics for the Dimension "Technological Fragmentation"	62
Table (4.2): Descriptive Statistics for the Dimension "Organizational and Administrative Fragmentation"	64
Table (4.3): Descriptive Statistics for the Dimension "Open Health Data Quality and Availability"	65
Table (4. 4): Descriptive Statistics for the Dimension "Operational Efficiency"	67
Table (4.5): Descriptive Statistics for the Dimension "Data Usage and Integration Practices"	68
Table (4.6): Regression Analysis for the Effect of Data Silos on Operational Inefficiency	70
table (4.7) Regression Analysis for the Effect of Lack of Standardized Interoperability Protocols on Fragmented Health Information Systems	71
table (4.8) Regression Analysis for the Effect of Availability and Quality of Open Health Data on AI-Based Integration Feasibility	72
Table 4.9: Regression Analysis of the Impact of AI-Powered Integration Solutions on Operational Efficiency and Data Usage & Integration Practices in Saudi Public Hospitals	74

Abstract

This study aims to analyze the impact of data integration, open data, and artificial intelligence on improving the efficiency and quality of healthcare services in public hospitals in the Kingdom of Saudi Arabia. The study adopted a descriptive-analytical methodology, with data collected from a sample of 310 respondents from public hospital staff, in addition to utilizing open data and official sources. The results indicated that data integration combined with artificial intelligence applications enhances decision-making speed and improves healthcare service quality, while highlighting the need to strengthen technological infrastructure and train human resources to ensure effective use of these systems. The findings also emphasized that clear policies and procedures for data governance are essential for achieving integration across different hospital departments. Based on these results, the study provided a set of strategic, technical, and administrative recommendations to support digital transformation and maximize the benefits of open data and artificial intelligence in enhancing healthcare performance.

Keywords

Data Integration – Open Data – Artificial Intelligence – Healthcare Efficiency – Public Hospitals

Study Terminology

Term	Definition (English)	Reference
Data Silos	Situations in which healthcare data is stored in isolated systems or departments, preventing effective sharing, integration, and unified decision-making in hospitals.	Alharthi et al. (2019)
Public Hospitals	Government-owned healthcare institutions that provide medical services to the population, often forming the backbone of national health systems.	WHO (2020)
Health Information Systems (HIS)	Digital systems designed to collect, store, manage, and transmit patient health information, supporting clinical and administrative operations.	Al-Kahtani et al. (2022)
Open Health Data	Health datasets made publicly accessible to enhance transparency, innovation, and evidence-based decision-making within the healthcare sector.	OECD (2021)
Artificial Intelligence (AI)	Computational techniques that enable machines to perform tasks such as learning, reasoning, and pattern recognition based on large-scale health data.	Rajkomar et al. (2019)
Machine Learning (ML)	A subset of AI that uses algorithms to automatically learn patterns from health data and generate predictions or classifications.	Topalović et al. (2020)
Interoperability	The ability of different health information systems to exchange, interpret, and use shared data seamlessly across departments or institutions.	Sheikh et al. (2021)
Electronic Health Records (EHRs)	Comprehensive digital records that store a patient's medical history, clinical data, and treatment information to improve continuity of care.	Alaa et al. (2021)

Data Integration	The process of merging data from multiple clinical, administrative, and operational sources into a unified system that supports analytics and decision-making.	Mzoughi et al. (2022)
Data Governance	A set of policies and procedures that ensure the accuracy, security, privacy, and quality of healthcare data within hospitals.	OECD (2021)
Predictive Analytics	The use of AI and ML techniques to analyze historical health data and predict future events such as disease risk or patient outcomes.	Rajkomar et al. (2019)
Digital Transformation in Healthcare	The adoption of digital technologies including AI, EHRs, and automation to improve quality, safety, and efficiency in health service delivery.	Al-Kahtani et al. (2022)
Healthcare Data Quality	The degree to which health data is accurate, complete, reliable, and suitable for clinical and managerial decision-making.	WHO (2020)
Big Data in Healthcare	Large and complex datasets generated from diverse sources such as EHRs and medical devices, requiring advanced analytics to extract insights.	Alaa et al. (2021)

Chapter One:

Introduction

1.1 Introduction

The rapid digital transformation of the global healthcare sector has intensified the reliance on data-driven systems, advanced informatics, and integrated health technologies. Modern healthcare delivery increasingly depends on the ability of institutions to access, share, and analyze vast amounts of clinical, administrative, and operational data to improve decision-making, enhance patient outcomes, and ensure service efficiency (Kumar et al., 2025a). However, despite continued investments in electronic health records (EHRs), digital health tools, and hospital information systems, many public hospitals still struggle with fragmented data environments. This fragmentation is most evident in the form of data silos, where information is stored in isolated departmental systems that lack interoperability, thereby hindering coordinated care and limiting the potential of advanced analytics (Bayer & Eisawi, 2025).

In Saudi Arabia, the challenge of data silos has become increasingly critical as the healthcare sector experiences rapid expansion, high patient demand, and growing expectations aligned with Vision 2030. Public hospitals often operate multiple disconnected systems across clinical departments, such as radiology, laboratories, emergency services, nursing units, and outpatient care. These systems frequently fail to exchange information efficiently, resulting in delays in diagnosis, duplicated laboratory and imaging tests, incomplete patient profiles, and reduced capability for predictive or preventive analysis (Alquayt et al., 2025; Moafa et al., 2024). The lack of unified data integration not only weakens the quality of care but also restricts administrators and policymakers from utilizing data to optimize resource allocation, design targeted interventions, or implement data governance strategies (Alsaleem et al., 2025).

The emergence of artificial intelligence (AI) presents a transformative opportunity to address these structural limitations. AI-driven systems can process large-scale datasets, merge information from heterogeneous sources, and generate actionable insights that support clinicians and decision-makers. Through machine learning, natural language processing, and intelligent decision-support tools, AI can mitigate the impact of data silos by identifying hidden patterns, predicting risk trajectories, and improving the accuracy of clinical assessments (Sulaiman et al., 2025; Solaiman et al., 2025). Importantly, AI offers the capability to automate data harmonization processes and enhance interoperability, ensuring that patient information becomes accessible in real-time across departments.

One promising avenue supporting this transformation is the use of open health data, which promotes transparency, collaboration, and innovation in healthcare. Open

data initiatives allow standardized datasets to be shared across institutions, researchers, and technology developers, enabling AI models to be trained on diverse and comprehensive information. This contributes to better disease monitoring, improved patient stratification, and more accurate prediction systems (Abdalzaher et al., 2025). By integrating open health data into AI-based decision frameworks, public hospitals can overcome traditional barriers related to data fragmentation, reduce operational inefficiencies, and support more personalized and cost-effective care delivery (Kumar et al., 2025b).

Given these developments, this project aims to analyze the nature, causes, and implications of data silos in public hospitals in Saudi Arabia and to propose AI-based solutions that leverage open health data to improve data integration, clinical workflow efficiency, and overall healthcare quality. The study seeks to highlight the root causes of data fragmentation—whether technological, organizational, or governance-related—and evaluate how AI tools can effectively bridge these gaps. The expected contribution of this research extends beyond identifying the challenges; it offers practical, evidence-based recommendations that support hospital administrators, IT departments, and policymakers in advancing digital transformation.

Ultimately, this project contributes to the broader agenda of developing a sustainable, data-driven healthcare system in Saudi Arabia—one that aligns with global best practices in digital health and supports Vision 2030 objectives for integrated, innovative, and high-quality healthcare services. By addressing data silos through AI-powered, open-data-driven strategies, the proposed solutions aim to enhance patient safety, improve service efficiency, and enable more informed decision-making at all levels of the healthcare system.

1.2 Research Problem:

Public hospitals in Saudi Arabia are undergoing rapid digital transformation as part of national efforts to enhance clinical efficiency, improve patient outcomes, and enable data-driven decision-making. Despite substantial investments in Health Information Systems (HIS), Electronic Health Records (EHRs), and AI-ready digital infrastructures, many hospitals continue to suffer from persistent data silos fragmented repositories of clinical, administrative, and operational data that remain isolated across departments. These silos limit the flow of information, hinder real-time analytics, and reduce the ability of healthcare institutions to leverage advanced technologies such as AI and predictive analytics for improved service delivery, as emphasized by recent studies on digital health integration in the region (Kumar et al., 2025; Alquayt et al., 2025).

The challenge of data silos is particularly critical in public hospitals, where the volume of patients is high and where unified, high-quality data is essential for standardized care and national-level health planning. Fragmented datasets lead to duplicated records, incomplete patient histories, inconsistent data formats, and delays in care coordination issues highlighted across multiple healthcare assessments in Saudi Arabia (Bayer & Eisawi, 2025; Alshararl et al., 2025). Moreover, the lack of interoperability between systems prevents hospitals from taking full advantage of emerging digital innovations such as AI-enabled clinical support, distributed decision-support systems, and smart hospital infrastructure (Almadani et al., 2025; Moafa et al., 2024).

Recent literature also underscores that the absence of accessible and well-structured open health data further restricts the development of national AI applications and limits the ability of researchers, policymakers, and practitioners to generate insights that could enhance quality, efficiency, and equity in healthcare delivery (Lichtenauer et al., 2023; Chen et al., 2024). In other words, data silos do not only affect hospital-level operations; they also impede the country's broader transition toward smart, interoperable healthcare ecosystems aligned with Vision 2030.

Even though digital health initiatives have advanced considerably in Saudi Arabia, many studies point to ongoing barriers in data sharing, standardization, governance, and cross-platform integration (Solaiman et al., 2025; Abdalzaher et al., 2025). These obstacles reduce the readiness of public hospitals to deploy AI-based tools that rely on comprehensive, high-quality, and integrated datasets. Addressing these challenges is vital for building sustainable, future-ready healthcare systems.

Given these gaps, there is a pressing need to analyze the core drivers, manifestations, and consequences of data silos within Saudi public hospitals, and to explore how AI-based solutions supported by open health data frameworks can overcome fragmentation and unlock the value of existing digital infrastructures. Understanding this issue is essential for advancing evidence-based policies, improving operational efficiency, and supporting national AI transformation strategies.

And in light of the above, and given the growing need for interoperable healthcare systems capable of utilizing data efficiently, the scientific and practical gap becomes evident regarding understanding the nature of data silos in Saudi public hospitals, their implications, and how they can be addressed through AI-based

solutions supported by open health data. Accordingly, the research problem can be summarized in the following research question:

What are the main challenges associated with data silos in public hospitals in Saudi Arabia, and how can AI-based solutions using open health data be proposed to effectively address these challenges?

From this, the following sub-questions emerge:

1. What are the key causes and implications of data silos within public hospitals in Saudi Arabia?
2. To what extent do technological, administrative, and regulatory factors contribute to maintaining fragmented healthcare information systems?
3. What types of open health data exist in Saudi Arabia, and how accessible and usable are they for research and data-integration purposes?
4. How can artificial intelligence (AI) be utilized to analyze and integrate fragmented health data across public healthcare institutions?
5. What international best practices exist in leveraging AI to address data silos in healthcare systems, and how can these practices be adapted to the Saudi context?
6. How aligned are AI-driven integration solutions with Saudi Arabia's Vision 2030 objectives, particularly regarding healthcare efficiency, innovation, and digital transformation?

1.3 Research Theoretical

1. Theoretical Significance

The theoretical significance of this study lies in its contribution to the growing body of knowledge on healthcare data management, particularly within the context of public hospitals in Saudi Arabia. Despite rapid digital transformation efforts, the persistence of data silos remains a major barrier to developing integrated and intelligent health systems. Existing literature highlights the complexity of fragmented health information infrastructures and the challenges they pose in terms of interoperability, data governance, and knowledge sharing (Bayer & Eisawi, 2025; Alshararl et al., 2025).

This study enriches theoretical understanding by examining data silos within the specific policy and organizational context of Saudi Arabia, offering insights into how administrative, technological, and regulatory factors interact to exacerbate data fragmentation. It also contributes to theoretical models of AI-driven data integration, building on the work of researchers who emphasize the importance of

AI for multi-source data fusion, predictive analytics, and system-wide optimization (Chen et al., 2024; Abedalrhman & Alzaydi, 2025).

Additionally, the study advances theoretical discourse on open health data, an emerging area in the Middle East. Prior work identifies openness as a cornerstone for transparency, research advancement, and digital innovation, yet significant barriers remain related to accessibility, standardization, and data protection (Lichtenauer et al., 2023). By investigating the availability and usability of open health data in the Kingdom, this study addresses a critical theoretical gap in regional scholarship.

Finally, the study offers a conceptual contribution to understanding how AI-based solutions such as federated learning, distributed decision support systems, and intelligent data pipelines can mitigate the limitations of data silos (Albshaier et al., 2025; Almadani et al., 2025). This integration strengthens the theoretical link between AI capabilities and national digital transformation frameworks, particularly within the context of Vision 2030.

2. Practical Significance

The practical significance of this research is substantial, as the challenges associated with fragmented data directly affect public healthcare delivery in Saudi Arabia. Data silos limit clinicians' ability to access complete patient histories, reduce coordination between departments, and ultimately impede decision-making processes especially in high-risk environments such as emergency care or mass gatherings like Hajj (Alquayt et al., 2025).

By proposing AI-based solutions that leverage open health data, this study has the potential to support healthcare administrators, policymakers, and IT leaders in improving efficiency, accuracy, and patient outcomes. Prior research emphasizes that AI can enhance diagnostic accuracy, automate clinical workflows, and support interdepartmental coordination when supported by integrated datasets (Moafa et al., 2024; Almotairi et al., 2025).

Additionally, the study aligns with Saudi Arabia's Vision 2030, which prioritizes digital health innovation, interoperability, and advanced analytics as pillars of healthcare reform. AI-driven integration solutions can contribute directly to improving patient care, reducing operational costs, and enhancing system reliability, as supported by evidence from Saudi and international healthcare systems (Kumar et al., 2025; Solaiman et al., 2025).

Practically, addressing data silos is also crucial for effective research, national health surveillance, and public health preparedness. Enhancing access to open health data supports innovation in hospitals, academic institutions, and health

startups accelerating the development of predictive models, personalized medicine, and smart hospital ecosystems (Nguyen, 2025; Awad & Mahmoud, 2025).

the practical value of this study lies in its ability to offer actionable AI-based strategies that can transform fragmented health information environments into integrated, intelligent, and future-ready systems.

1.4 Study Objectives:

The main objective of this study is to analyze the challenges of data silos in public hospitals in Saudi Arabia and to develop an AI-based framework, supported by open health data, to enhance data integration, accessibility, and healthcare decision-making.

From this main objective, the following specific sub-objectives emerge:

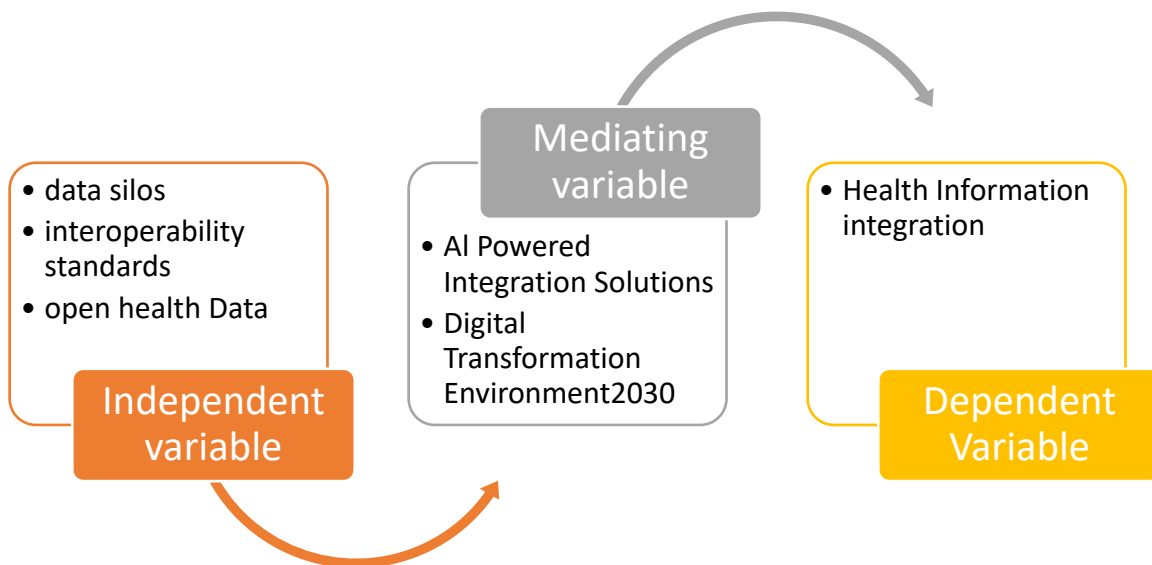
1. To investigate the underlying technological, administrative, and regulatory causes that contribute to the persistence of data silos in public hospitals.
2. To assess the availability, accessibility, and usability of open health data in Saudi Arabia and determine its suitability for supporting integrated healthcare systems.
3. To explore the potential of artificial intelligence (AI) tools and techniques in analyzing and integrating fragmented healthcare datasets, based on regional and global best practices.
4. To propose a customized AI-driven solution aligned with Saudi Arabia's Vision 2030 to reduce data fragmentation and improve operational efficiency within public healthcare institutions.

1.5 Research Hypotheses :

1. **H1:** There is a statistically significant relationship between data silos and the inefficiency of clinical and administrative operations in public hospitals in Saudi Arabia.
2. **H2:** The lack of standardized interoperability protocols significantly contributes to the persistence of fragmented health information systems.
3. **H3:** The availability and quality of open health data in Saudi Arabia positively influence the feasibility of developing and implementing AI-based integration models.
4. **H4:** AI-powered integration solutions significantly reduce the impact of data silos by enhancing data accessibility, accuracy, and system-wide interoperability across public healthcare institutions.

1.6 Research Model

Relationships Between Variables



1.7 The Scientific Theoretical Framework Adopted in the Study

This research adopts a multidimensional theoretical framework that integrates three major theories deemed most appropriate for analyzing the challenges of data silos in public hospitals in Saudi Arabia: the Systems Approach, Institutional Theory, and Data Governance Theory. These theories were selected for their ability to explain the complex interactions between data, technological infrastructures, and the regulatory environment—key elements underlying the fragmentation of health information systems.

1. Systems Approach

The Systems Approach is widely used in analyzing healthcare institutions, as it considers the hospital a unified system consisting of inputs, processes, outputs, and feedback loops. According to *Bertalanffy (1968)*, every component of a system interacts with and influences other components; therefore, the existence of isolated databases or non-communicating systems disrupts the natural flow of information. *Sweeney & McFarlin (2020)* emphasize that data-driven healthcare systems require a high level of internal integration to achieve efficiency and quality, making this approach particularly relevant to explaining how data silos impede clinical and administrative operations.

2. Institutional Theory

Institutional Theory explains how organizational practices are shaped by regulatory, normative, and mimetic pressures. *DiMaggio & Powell (1983)* argue that organizations comply with governmental regulations, professional norms, and competitive expectations to maintain legitimacy. In the Saudi healthcare context, bodies such as the Ministry of Health, the Saudi Data and AI Authority (SDAIA), and the National Health Information Center establish mandatory data policies and interoperability standards. Recent studies, including *Al-Kahtani (2022)*, indicate that fragmented health data often results from inconsistent adherence to national digital policies. Thus, Institutional Theory provides a strong foundation for understanding how regulatory gaps contribute to persistent data silos.

3. Data Governance Theory

Data Governance Theory views data as a strategic asset that must be managed through clear rules, defined responsibilities, and quality standards. According to *Otto (2011)*, data governance involves four pillars: data quality, lifecycle management, accountability, and compliance. In healthcare, *Lee et al. (2018)* stress that the success of electronic health record integration depends on robust governance mechanisms that ensure privacy protection, standardization, and clarity in data ownership. This theory is central to the present study because it explains how the quality and structure of open health data influence the feasibility of implementing AI-based integration solutions.

4. Integration of the Three Theories

The integration of the Systems Approach, Institutional Theory, and Data Governance Theory provides a comprehensive and multidimensional framework for analyzing the persistence of data silos in public hospitals and exploring the potential of artificial intelligence (AI) and open health data to overcome these challenges. Each theory contributes a unique analytical lens, and together they enable a deeper understanding of the structural, organizational, and data-centric dynamics that shape healthcare information fragmentation within the Saudi Arabian context.

From the perspective of the Systems Approach, public hospitals function as complex and interdependent organizations composed of clinical departments, administrative units, health information systems, and decision-support structures. When data become fragmented across isolated systems, the hospital loses its systemic coherence; feedback loops weaken, coordination between units becomes inefficient, and delays in clinical and administrative decision-making increase. Thus, data silos are not merely technical glitches but structural disruptions that

undermine the harmony and efficiency of the entire healthcare system. This theoretical lens highlights that improving interoperability is essential for restoring systemic balance and enabling hospitals to operate as integrated, adaptive, and responsive entities.

Institutional Theory adds a deeper layer by explaining why data fragmentation persists despite ongoing digital transformation efforts. Public hospitals in Saudi Arabia operate within a dense institutional environment shaped by ministries, regulatory bodies, accreditation agencies, and national digital health strategies. Institutional pressures whether coercive, normative, or mimetic—may promote digitalization, but inconsistencies in policy enforcement, unequal digital maturity across hospitals, and organizational resistance to change can hinder interoperability. Moreover, institutional norms and routines often compel hospitals to adopt information systems that reflect existing practices rather than transformative requirements. This theory therefore elucidates how regulatory landscapes, cultural expectations, and organizational legitimacy influence decisions surrounding data management and technology adoption.

Data Governance Theory complements the previous two lenses by outlining the conditions under which data can be effectively integrated, shared, and used for AI-enabled insights. Even when systems are interconnected and regulations are supportive, interoperability cannot be achieved without high-quality data, standardized formats, clear rules of ownership and accountability, and robust mechanisms for privacy and cybersecurity. This theory emphasizes that open health data must be accurate, well-structured, ethically managed, and accessible in order to serve as a foundation for AI-based integration. It establishes the operational and technical prerequisites for creating an ecosystem where data flow seamlessly across clinical and administrative boundaries.

When synthesized, these theories offer a strong analytical foundation for the present study. The Systems Approach reveals the functional consequences of data silos, highlighting the need for systemic integration. Institutional Theory clarifies the regulatory, cultural, and organizational forces that shape data fragmentation and influence hospitals' readiness to adopt AI-based solutions. Data Governance Theory defines the data-related standards and safeguards required to ensure that integration efforts are feasible, sustainable, and aligned with national priorities. Together, they create a holistic model explaining not only how data silos emerge and persist, but also how AI and open data can be strategically leveraged to address them.

Ultimately, this integrated theoretical perspective supports the study's aim of proposing AI-driven interoperability solutions that are technically robust,

institutionally feasible, and systemically transformative. It also aligns closely with the strategic vision of Saudi Arabia's healthcare transformation under Vision 2030, where digital innovation, data-driven decision-making, and improved service efficiency are central priorities. By framing the issue through these three complementary theories, the research gains both conceptual depth and practical relevance, enabling a nuanced analysis of the pathways through which healthcare data fragmentation can be mitigated in public hospitals.

Chapter Two:
Theoretical Framework Previous
Studies and

2.1 Introduction

The second chapter provides the theoretical and empirical foundation for understanding the challenges of data silos in public hospitals in Saudi Arabia and exploring the potential of artificial intelligence (AI) and open health data in overcoming these challenges. This chapter is divided into two main parts: the theoretical framework, which outlines the conceptual underpinnings and models relevant to data fragmentation, information integration, and digital transformation in healthcare; and the review of previous studies, which synthesizes existing research to identify patterns, gaps, and lessons learned from both global and local contexts.

Theoretically, the study draws upon multiple perspectives, including the Systems Approach, which views healthcare as an interconnected system where inefficiencies in data flow disrupt coordination and decision-making (Almadani et al., 2025); Institutional Theory, which explains how regulatory frameworks, norms, and organizational structures influence data management practices and technology adoption (Solaiman et al., 2025); and Data Governance Theory, which emphasizes the policies, standards, and accountability mechanisms that enable data interoperability and quality (Sulaiman et al., 2025). Together, these theories form an integrated framework that supports understanding how AI and open data can enable more cohesive, transparent, and efficient healthcare systems.

Empirically, this chapter reviews a broad range of studies that have examined data silos, AI applications, and open health data initiatives in healthcare. Several recent works have highlighted the persistence of fragmented health information systems and their negative impact on efficiency, patient care, and policy formulation (Bayer & Eisawi, 2025; Alshararl et al., 2025). Other studies have demonstrated how AI can facilitate real-time integration, improve data quality, and enhance clinical decision support systems (Kumar et al., 2025; Moafa et al., 2024). Moreover, research on open health data has shed light on the importance of accessibility, interoperability, and ethical governance in data-driven healthcare transformation (Lichtenauer et al., 2023; Chen et al., 2024).

By synthesizing these theoretical insights and empirical findings, this chapter establishes a strong conceptual base for the current research. It aims to identify the critical factors contributing to data silos in Saudi public hospitals, evaluate global best practices in AI-driven integration, and position the study within the broader discourse on healthcare digitalization and Vision 2030 objectives.

2.2 Theoretical Framework

The theoretical framework serves as the foundation upon which this study is built. It provides the conceptual basis for analyzing the problem of data silos in public healthcare institutions and exploring how artificial intelligence (AI) and open health data can enhance integration and efficiency. By drawing on relevant theories, this framework helps explain the relationships among the study variables and guides the interpretation of findings. Specifically, the framework relies on three key theories: the Systems Approach, which emphasizes the interdependence of healthcare components; Institutional Theory, which examines regulatory and organizational influences; and Data Governance Theory, which focuses on the mechanisms ensuring data quality, accessibility, and interoperability. Together, these perspectives offer a comprehensive understanding of the structural and technological dimensions of healthcare digital transformation in line with Saudi Vision 2030.

2.2.1 Conceptualizing Data Silos in the Healthcare Sector

1. Definition and characteristics of data silos

Data silos in healthcare refer to isolated repositories of information that are not easily accessible or interoperable across different departments, systems, or institutions. According to Alshararl et al. (2025), data silos emerge when clinical, administrative, and operational systems operate independently, preventing seamless communication and data exchange. Similarly, Kumar et al. (2025) define data silos as “fragmented data environments where patient records, operational metrics, and administrative data remain confined within specific software systems or organizational boundaries.”

Moafa et al. (2024) emphasize that such fragmentation leads to inefficiencies, duplication of effort, and data inconsistencies, which ultimately affect decision-making and patient care quality. In the same vein, Almadani et al. (2025) describe healthcare data silos as “non-integrated databases that hinder comprehensive analysis and evidence-based healthcare planning.” Furthermore, Albshaier et al. (2025) highlight that the absence of common interoperability standards reinforces data isolation, making it difficult to create unified electronic health records or utilize data-driven innovations.

Conceptually, data silos in the healthcare context represent both a technical and organizational challenge. Technically, they arise from incompatible systems, differing data formats, and security restrictions. Organizationally, they reflect

fragmented governance structures, lack of collaboration, and departmental resistance to data sharing (Bayer & Eisawi, 2025; Sulaiman et al., 2025).

In this study, the term “data silos” is operationally defined as:

“The existence of separate, non-communicating data systems within public hospitals in Saudi Arabia that limit the integration, accessibility, and utilization of healthcare information across administrative and clinical functions.”

2. Causes and Implications of Fragmented Health Information Systems

Fragmentation of health information systems represents one of the major challenges hindering digital transformation in the healthcare sector. This issue arises from a combination of technical, organizational, and regulatory factors, leading to multiple adverse effects on operational efficiency, data reliability, and evidence-based decision-making.

First: Causes of Fragmentation

- 1) **Reliance on legacy systems and lack of interoperability standards** One of the primary causes of health information fragmentation is the continued use of legacy systems that do not support unified interoperability standards such as HL7 and FHIR. These outdated systems prevent data exchange across healthcare organizations, creating isolated “data islands” within the health sector. *Bayer and Eisawi (2025)* and *Moafa et al. (2024)* emphasized that the absence of technical standardization remains a major obstacle to effective data integration, particularly in public hospitals with traditional infrastructures, thereby hindering digital healthcare transformation.
- 2) **Vertically implemented health programs without institutional coordination** In many healthcare systems, information systems are developed vertically to serve specific programs such as immunization, maternal health, or chronic disease management, without an overarching national framework ensuring integration among them. This vertical design leads to data duplication, inconsistent reporting, and multiple disconnected databases. According to *Alshararl et al. (2025)*, the lack of coordination among health programs has weakened data flow between departments and reduced the ability to make comprehensive, data-driven decisions.
- 3) **Weak data governance and lack of institutional alignment** Ineffective data governance is another critical factor contributing to fragmentation. Many healthcare systems lack a national governance framework that enforces standardized data management and sharing practices. As a result, each organization develops its own database structures and reporting mechanisms, causing inconsistencies and redundancy. *Solaiman et al. (2025)* noted that the

absence of clear governance policies undermines interoperability efforts and limits the potential use of artificial intelligence in nationwide health data analysis.

- 4) **Regulatory constraints and inconsistent application of privacy laws** Although data privacy laws are essential for protecting patient information, their inconsistent enforcement across institutions often restricts data sharing. In Saudi Arabia, differences between public and private sectors in how data exchange is regulated have created communication gaps between systems. *Solaiman et al. (2025)* highlighted that the lack of unified legal and regulatory standards causes institutions to hesitate in sharing data, further exacerbating information fragmentation.
- 5) **Limited technical and human capacities in health data management** A shortage of professionals specialized in data science and health informatics also aggravates the fragmentation problem. Insufficient training and low digital literacy among healthcare workers lead to poor data handling and ineffective system integration. *Kumar et al. (2025)* emphasized that limited digital competencies among healthcare staff represent a major obstacle to implementing AI-driven integration solutions.

Second: Implications of Fragmentation

- 1) **Reduced operational efficiency and duplication of medical procedures** Fragmented systems often lead to repeated laboratory tests and redundant diagnostic procedures due to the lack of unified patient records. This not only increases operational costs but also wastes human and technological resources. *Alquayt et al. (2025)* reported that data fragmentation directly affects service efficiency and delays clinical response within healthcare institutions.
- 2) **Poor data quality and inconsistent performance indicators** Fragmentation results in the generation of incomplete and conflicting datasets, undermining the reliability of analytical reports and policy decisions. *Almadani et al. (2025)* pointed out that weak data integration negatively affects the accuracy of health performance indicators and limits the development of effective evidence-based policies.
- 3) **Barriers to artificial intelligence implementation** AI applications rely on large, diverse, and standardized datasets. Fragmented systems, however, restrict access to integrated data, leading to weak model performance and inaccurate predictions. *Moafa et al. (2024)* asserted that fragmented data environments are among the most critical barriers to adopting AI technologies in healthcare systems.
- 4) **Disrupted continuity of care and decreased service quality** Data fragmentation disrupts the continuity of patient information across healthcare facilities. When medical records are not seamlessly transferred, clinicians face

incomplete patient histories, increasing the likelihood of diagnostic errors. *Alsaleem et al. (2025)* confirmed that fragmented information systems negatively affect the quality of care and patient safety.

5) **Hindrane to achieving Saudi Vision 2030 healthcare transformation goals**

At a strategic level, persistent data fragmentation poses a major barrier to the successful implementation of Saudi Arabia's healthcare digital transformation initiatives. Overcoming this issue is essential to building an integrated, data-driven health system. *Bayer and Eisawi (2025)* emphasized that addressing data silos is a crucial step toward achieving Vision 2030 goals in improving healthcare efficiency, innovation, and overall service quality.

3. Impact on Data Accessibility, Quality, and Patient Outcomes

The impact of data silos on data accessibility and quality represents one of the major challenges facing modern healthcare systems, as they hinder the exchange of information between different departments or medical institutions, thereby reducing the ability of physicians and researchers to obtain integrated patient data (Bates et al., 2018). When health data are confined within disconnected systems, accessibility becomes limited, which reduces both the speed and accuracy of medical decision-making (Vest & Kash, 2016).

In terms of quality, fragmented data are often characterized by redundancy or inconsistency due to the repeated and uncoordinated entry of the same information across multiple systems, leading to deteriorating data integrity and reliability (Menachemi & Collum, 2011). Moreover, the lack of integration among health systems reduces the effectiveness of statistical and clinical analysis and limits the application of artificial intelligence tools for disease diagnosis and risk prediction (Ben-Assuli, 2015).

At the patient level, several studies have shown that the absence of comprehensive health data increases the likelihood of medical errors and negatively affects the continuity and quality of care (Bates et al., 2018; Kruse et al., 2016). Physicians without access to complete patient histories or prior test results may make inaccurate clinical decisions, which in turn undermines treatment effectiveness and patient safety. Therefore, addressing data silos is a critical step toward improving healthcare quality and enhancing the efficiency of health systems.

2.2.2 Artificial Intelligence in Health Data Integration

1. The role of AI in healthcare transformation

Artificial Intelligence (AI) plays a transformative role in modern healthcare systems by enabling data-driven decision-making, improving diagnostic accuracy, and enhancing operational efficiency (Topol, 2019). Through the use of machine learning, natural language processing, and predictive analytics, AI allows for the integration and analysis of massive volumes of health data that were previously fragmented across various systems (Rajkomar et al., 2019). This integration supports more comprehensive insights into patient care, disease trends, and healthcare resource allocation.

AI contributes to healthcare transformation on multiple levels. Clinically, it facilitates early disease detection, personalized treatment plans, and improved patient monitoring through predictive models that identify potential risks before they become critical (Obermeyer & Emanuel, 2016). Administratively, AI optimizes workflows by automating routine tasks such as scheduling, documentation, and claims processing, which enhances efficiency and reduces human error (Davenport & Kalakota, 2019).

Moreover, AI serves as a bridge for interoperability between disparate systems, translating and harmonizing data from different formats and standards. This capability not only supports clinical integration but also enhances the effectiveness of public health surveillance and research (Shickel et al., 2018). Therefore, AI represents a cornerstone of healthcare digital transformation, aligning with national strategies—such as Saudi Vision 2030—that aim to build smart, efficient, and patient-centered health systems.

2. Machine Learning and Predictive Analytics in Clinical Decision-Making

1) Scope of Clinical Applications of Machine Learning

Machine learning (ML) encompasses a wide spectrum of applications in healthcare, including automated image diagnosis, disease classification, prediction of disease progression (e.g., sepsis risk or hospital readmission), personalized treatment planning, and chronic disease management. These applications extend beyond outcome modeling to include remote patient monitoring and precision medicine. Recent studies have demonstrated that ML systems can achieve diagnostic accuracy comparable to or exceeding that of human experts (Anjum et al., 2024; Chen et al., 2024).

2) Sources and Types of Data Used

ML models rely on integrating diverse data sources—structured electronic health records (EHRs), medical imaging data (DICOM), unstructured clinical texts

(physician notes, reports), and behavioral or sensor data from wearable devices. This **multimodal data fusion** enhances predictive accuracy and enables a holistic understanding of patient health (Chen et al., 2024).

3) Common Models and Methodological Approaches

- **Supervised Learning:** Logistic regression, decision trees, random forests, and deep learning (CNNs for imaging, RNNs/Transformers for textual data).
- **Unsupervised Learning:** Clustering and pattern discovery in large-scale datasets.
- **Federated Learning:** Allows model training across multiple hospitals without sharing raw data, preserving privacy while overcoming data silos — an approach increasingly adopted in Saudi and regional studies (Albshaier et al., 2025; Alasmari et al., 2025).

4) Importance of Feature Engineering

Feature quality is a key determinant of model performance. Data preprocessing includes cleaning, normalization, ICD coding alignment, and extraction of temporal and textual features using NLP. Proper feature engineering reduces bias and enhances model interpretability (Kumar et al., 2025).

5) Natural Language Processing (NLP) and Knowledge Extraction from Clinical Texts

NLP enables the extraction of valuable information from clinical notes and radiology reports—often containing details not encoded in structured fields. Advanced transformer-based language models can convert unstructured text into analytical features while ensuring patient de-identification and data privacy (Chen et al., 2024; Moafa et al., 2024).

6) Model Validation and Generalizability

A model's success in one hospital is insufficient; external validation across multiple datasets is necessary to ensure generalizability. Evaluation should employ performance metrics (AUC, sensitivity, specificity, calibration) and fairness testing to detect bias caused by fragmented data or underrepresentation (Juneja, 2025; Moafa et al., 2024).

7) Explainability and Transparency (XAI)

For clinical and managerial acceptance, ML outputs must be interpretable. Explainable AI (XAI) techniques help illustrate how each feature contributes to a prediction, enhancing accountability and regulatory compliance (Solaiman et al., 2025).

8) Operational Integration with Hospital Information Systems

Deploying ML models in clinical workflows requires integration with hospital EHRs, user-friendly interfaces, and real-time responsiveness. Success depends on adherence to interoperability standards (HL7/FHIR) and a robust deployment infrastructure (Alsharar et al., 2025; Almadani et al., 2025).

9) Data Privacy and Federated Training Approaches

Given the sensitivity of health data, federated learning, encrypted model updates, and differential privacy mechanisms are essential for collaborative learning without raw data exchange. These approaches balance analytical utility with data protection (Albshaier et al., 2025; Sulaiman et al., 2025).

10) Governance, Ethics, and Regulatory Considerations

Implementing ML in healthcare demands a governance framework encompassing data quality standards, accountability mechanisms, model certification protocols, and risk management strategies. Ethical considerations—especially regarding fairness and transparency—must be embedded throughout the lifecycle (Juneja, 2025; Solaiman et al., 2025).

11) Human and Institutional Capacity Building

Successful adoption requires building workforce capacity in data science, information systems, and clinical informatics. Training healthcare professionals to interpret and apply ML outputs is crucial, as is preparing IT teams to manage, monitor, and maintain deployed models (Kumar et al., 2025; Awad & Mahmoud, 2025).

12) Impact Evaluation and Post-Deployment Monitoring

Continuous monitoring is essential to assess clinical and economic impacts—such as reduced mortality, shortened hospital stays, or cost savings. Models must be routinely recalibrated to prevent performance drift and ensure sustained reliability (Nguyen, 2025; Regin & Rajest, 2024).

13) Contextual Challenges in Saudi Arabia

In Saudi Arabia, implementation challenges include varying digital maturity among hospitals, data ownership fragmentation, and alignment with national e-health initiatives like *Seha Virtual Hospital* and *Nafis*. Therefore, scalable and adaptable ML models are essential for achieving integration and supporting Vision 2030 objectives (Kumar et al., 2025; Bayer & Eisawi, 2025).

3. AI Applications for Data Unification and Interoperability

The integration of artificial intelligence (AI) into healthcare data systems has become a transformative solution to overcoming the long-standing challenges of

fragmented and non-interoperable health information systems. Data unification refers to the process of consolidating heterogeneous datasets—originating from multiple hospitals, departments, and technologies—into a standardized, interoperable framework that facilitates seamless data exchange, accessibility, and analytical consistency (Alshararl et al., 2025; Moafa et al., 2024).

1) Automated Data Mapping and Standardization

AI-driven tools use machine learning algorithms and natural language processing (NLP) techniques to automatically map, clean, and standardize medical records from different systems. These applications convert inconsistent terminologies, formats, and coding systems (e.g., ICD-10, SNOMED CT, LOINC) into unified data structures, thus enabling efficient cross-platform interoperability (Kumar et al., 2025; Almadani et al., 2025). For instance, automated ontology alignment powered by deep learning has proven effective in harmonizing data across electronic health record (EHR) systems within Saudi public hospitals (Bayer & Eisawi, 2025).

2) Federated and Distributed Learning Models

AI-based federated learning models allow institutions to collaboratively train algorithms without transferring sensitive patient data. This decentralized approach enhances interoperability while maintaining privacy, as each hospital's local data contributes to a shared model update rather than a central database (Albshaier et al., 2025; Sulaiman et al., 2025). In Saudi Arabia, these models are gaining traction for integrating patient data from geographically dispersed healthcare facilities under unified analytical platforms aligned with Vision 2030's digital transformation goals.

3) Semantic Interoperability through Ontology-Based Systems

Semantic interoperability—the ability of systems to exchange data with unambiguous, shared meaning—is a core AI application area. Ontology-based frameworks use machine reasoning and knowledge graphs to interpret medical terminology and relationships automatically. These frameworks ensure that different healthcare providers can exchange patient information without semantic distortion, improving data consistency and clinical decision-making (Chen et al., 2024; Awad & Mahmoud, 2025).

4) AI-Enhanced Data Governance and Quality Assurance

AI supports data governance by continuously monitoring data quality, detecting anomalies, and enforcing standard compliance rules. Advanced algorithms identify missing, inconsistent, or duplicate data in real time, thereby improving data

accuracy and reliability (Moafa et al., 2024; Abdalzaher et al., 2025). Such mechanisms are critical for achieving a unified national health information infrastructure.

5) Real-Time Interoperability Platforms

AI enables the development of intelligent interoperability middleware that connects disparate hospital information systems in real time. These platforms employ predictive synchronization and adaptive learning to ensure that patient data are updated dynamically across departments and regions. This has been particularly effective during large-scale events such as Hajj, where real-time integration of clinical data is crucial for patient safety and operational efficiency (Alquayt et al., 2025).

6) AI in Open Health Data Utilization

AI facilitates the integration and analysis of open health datasets—both national and international—by structuring unorganized public health data into standardized, machine-readable formats. This accelerates epidemiological research, supports predictive modeling, and enhances public health surveillance (Lichtenauer et al., 2023; Kumar et al., 2025). Saudi initiatives increasingly rely on open data portals and AI-driven analytics to strengthen transparency and evidence-based health policy design.

7) Cybersecurity and Ethical Integration

Ensuring interoperability must go hand in hand with robust cybersecurity. AI assists in encrypting data, detecting intrusion, and safeguarding patient confidentiality across interconnected systems. This ethical dimension is emphasized in the convergence of AI and cybersecurity frameworks, which protect sensitive data flows within unified digital health ecosystems (Sulaiman et al., 2025; Solaiman et al., 2025).

8) Strategic Implications for Saudi Vision 2030

In line with Saudi Vision 2030, the adoption of AI for health data unification promotes efficient resource utilization, reduces duplication of services, and strengthens the country's healthcare digital infrastructure. Through standardized interoperability frameworks, Saudi Arabia moves closer to achieving a data-driven, patient-centered healthcare ecosystem (Bayer & Eisawi, 2025; Kumar et al., 2025).

2.2.3 Open Health Data and Digital Transformation

1. The concept of open data in healthcare

Open health data refers to health-related data that is freely available for use, reuse, and redistribution by anyone without restrictions, while ensuring privacy and confidentiality of individuals. According to *OECD (2021)*, open data in healthcare encompasses datasets such as hospital performance indicators, disease surveillance reports, and health expenditure statistics that are made publicly accessible to promote transparency and innovation. *Janssen et al. (2017)* emphasize that open health data enables collaboration between governments, researchers, and technology developers, fostering evidence-based policymaking and public engagement. In this research, open health data is conceptually understood as a strategic enabler for integrating fragmented health systems and driving AI-based healthcare transformation.

2. Benefits and challenges of open health data initiatives

The implementation of open health data initiatives offers multiple benefits across healthcare systems. *Safarov (2020)* notes that open data enhances transparency and accountability, allowing citizens to evaluate healthcare performance and policy effectiveness. Moreover, *Kitchin (2018)* highlights those open datasets stimulate innovation, particularly in digital health startups and AI-driven analytics, by providing the raw material for algorithmic model training. However, significant challenges remain, particularly concerning data privacy, standardization, and interoperability. *Wang et al. (2019)* argue that while open data improves accessibility, inconsistent data formats and inadequate governance frameworks often hinder its full utilization. Furthermore, ethical concerns about patient consent and data misuse persist, underscoring the need for robust governance structures.

3. Global trends in data openness and interoperability standards

Globally, countries have been advancing toward data openness and interoperability as part of digital health transformation. The European Union's *Health Data Space (EU, 2022)* initiative, for instance, aims to ensure secure and standardized access to health data for research and innovation across member states. Similarly, the *United States Office of the National Coordinator for Health Information Technology (ONC)* developed interoperability frameworks such as the *FHIR (Fast Healthcare Interoperability Resources)* standard, which facilitates real-time data sharing between systems (ONC, 2020). According to *OECD (2021)*, nations that have invested in harmonizing data standards and promoting open data ecosystems have achieved greater efficiency, research productivity, and public trust in digital health systems. Therefore, studying these global practices provides valuable

insights for countries like Saudi Arabia seeking to enhance data integration through AI and open health initiatives aligned with Vision 2030.

2.2.4 Relationship Between AI Integration and Open Health Data

1) Theoretical linkage between AI and open data

Artificial intelligence (AI) and open health data are theoretically complementary: open data supplies the diverse, large-scale, and (potentially) representative datasets that enable AI algorithms to learn, generalize, and produce robust models; in turn, AI provides the analytical mechanisms that convert raw open datasets into actionable knowledge, predictive models, and decision-support tools. From an information-systems perspective, this linkage rests on two core assumptions. First, the **data-availability assumption** posits that AI performance improves with data volume, variety, and quality—conditions typically enhanced by open data initiatives (Kumar et al., 2025; Lichtenauer et al., 2023). Second, the **value-extraction assumption** holds that algorithmic methods (machine learning, NLP, knowledge graphs) can synthesize heterogeneous open datasets into semantically aligned inputs suitable for clinical and policy use (Chen et al., 2024; Abdalzaher et al., 2025). The synergy is thus both empirical (more/better data → better AI) and structural (AI enables harmonization, imputation, and semantic mapping that make open data practically usable). However, theoretical accounts also caution about feedback effects: biased open datasets can perpetuate algorithmic bias, and AI-driven inferences may change data-production incentives—hence the need for governance-aware theory (Juneja, 2025; Solaiman et al., 2025).

2) Conceptual models explaining their combined impact on healthcare innovation

Several conceptual models can be used to explain how AI and open health data jointly drive healthcare innovation. Three parsimonious but complementary models are particularly useful:

1. **The Data-to-Decision Pipeline model:** This linear-cybernetic model maps stages from data publication (open datasets) → data ingestion and harmonization (NLP, ontology alignment) → model development (ML, federated learning) → deployment (CDSS, population analytics) → impact (clinical outcomes, policy). The pipeline emphasizes transformation stages where AI converts raw open data into decision-ready artifacts (Kumar et al., 2025; Chen et al., 2024).
2. **The Socio-Technical Feedback model:** This model situates AI and open data within interacting technical, organizational, and regulatory layers. It highlights feedback loops: (a) AI outputs inform policy and clinical

workflows, which change data generation patterns; (b) governance rules reshape what open data are published and how. This model underscores co-evolution: technological advances and institutional incentives must be aligned for sustained innovation (Solaiman et al., 2025; Sulaiman et al., 2025).

3. **The Layered Interoperability-Value model:** This multi-layer model posits concentric layers—syntactic (formats/standards), semantic (ontologies), organizational (workflows/policies), and analytics (AI models/XAI). Innovation arises when value (improved outcomes, efficiency, research acceleration) accrues across layers—i.e., when syntactic interoperability enables semantic mapping, which enables reliable AI analytics and trustworthy decision support (Alshararl et al., 2025; Awad & Mahmoud, 2025).

Together, these models explain not only the mechanics of how AI and open data interact but also where failure modes occur (e.g., poor semantics, weak governance, biased datasets) and where interventions (technical, organizational, ethical) are most effective.

3) Frameworks for achieving holistic health data ecosystems

Operationalizing the theoretical linkage and conceptual models requires an integrated framework that addresses technical, organizational, legal/ethical, and capacity-building dimensions. A practicable framework—drawn from the literature and adapted to the Saudi context—includes four interdependent pillars:

1. Technical Foundation (Standards & Architecture)

- Adopt and enforce interoperability standards (FHIR, HL7, SNOMED CT, LOINC) and promote machine-readable open-data formats (CSV/JSON with clear metadata).
- Implement ontology-driven middleware and data harmonization pipelines (NLP for unstructured notes, automated code mapping) to prepare open datasets for AI consumption (Kumar et al., 2025; Chen et al., 2024).

2. Privacy-Preserving Analytics (Governance + Methods)

- Use federated learning, secure multiparty computation, and differential privacy to enable cross-institutional model training without centralizing raw data, balancing utility and confidentiality (Albshaier et al., 2025; Sulaiman et al., 2025).

- Establish data-access tiers and standardized data-use agreements that specify permissible uses of open datasets for model development and research.

3. Ethical & Regulatory Governance

- Create clear policies for data publication, consent, de-identification, and algorithmic accountability. Embed requirements for model explainability (XAI), fairness assessments, and continuous auditing (Juneja, 2025; Solaiman et al., 2025).
- Align national strategies (e.g., Vision 2030 initiatives such as Nafis and Seha Virtual Hospital) with legal frameworks to reduce fragmentation in ownership and stewardship roles (Bayer & Eisawi, 2025).

4. Capacity & Institutional Alignment

- Invest in workforce development (data scientists, clinical informaticians, ethics officers) and change-management programs for clinicians and administrators to adopt AI-augmented workflows.
- Foster cross-sector collaborations (academia, public health agencies, industry) and a national coordinating body to monitor standards adoption and open-data curation (Kumar et al., 2025; Lichtenauer et al., 2023).

Finally, the framework must incorporate **monitoring and learning**: defined KPIs (data quality scores, model performance metrics, clinical outcome indicators), continual re-validation of models across diverse datasets, and mechanisms for stakeholder feedback. When implemented coherently, such a framework transforms open health data from static repositories into dynamic substrates for AI-driven innovation—enabling predictive public health, personalized care, and evidence-based policy—while managing risks of bias, privacy breaches, and institutional resistance.

2.3 Previous Studies

The review of previous studies provides a comprehensive understanding of the theoretical and empirical foundations related to the topic of health data integration and digital transformation. This section highlights the most relevant research addressing issues of data silos in healthcare, the application of artificial intelligence (AI) in data integration, and the role of open health data in promoting interoperability and innovation. Reviewing these studies helps identify the

strengths and limitations of existing literature, clarify the relationship between AI and open data in healthcare systems, and reveal the research gaps that the current study aims to fill.

2.3 .1 Studies on Data Silos and AI Integration in Healthcare

- 1. Kumar, Rakesh, et al. (2025). “Leveraging Artificial Intelligence to Achieve Sustainable Public Healthcare Services in Saudi Arabia: A Systematic Literature Review of Critical Success Factors.” *Computer Modeling in Engineering & Sciences*, 142(2), 1289.**

This study conducted by Kumar and colleagues (2025) represents a comprehensive systematic literature review that explores the critical success factors influencing the deployment of artificial intelligence (AI) in Saudi Arabia’s public healthcare sector. The primary objective of the research was to identify technological, organizational, and infrastructural determinants that contribute to the successful adoption and sustainability of AI-driven healthcare systems.

Using the Systematic Literature Review (SLR) methodology, the researchers examined over fifty peer-reviewed studies published between 2018 and 2024, encompassing empirical, conceptual, and case-based investigations. The review process adhered to the PRISMA framework to ensure methodological rigor, transparency, and replicability. Findings revealed that data integration and unified digital infrastructure are essential preconditions for effective AI implementation in healthcare. The fragmentation of data often stored in isolated systems or incompatible formats—was identified as one of the most critical obstacles hindering AI accuracy and scalability. The authors emphasized that data silos weaken predictive analytics and compromise the interoperability required for efficient decision-making and clinical support systems. Moreover, the study highlighted that achieving sustainable AI-driven healthcare in alignment with Saudi Vision 2030 requires a strategic focus on data governance, cross-institutional interoperability, and workforce digital literacy. The authors concluded that the elimination of data silos through integrated electronic health record (EHR) systems and centralized data platforms could significantly enhance the effectiveness of predictive models, optimize resource allocation, and improve patient outcomes across the public health ecosystem.

- 2. Kumar, Rakesh, et al. (2025). “Adoption Challenges to Artificial Intelligence Literacy in Public Healthcare: An Evidence-Based Study in Saudi Arabia.” *Frontiers in Public Health*, 13, 1558772.**

In this empirical study, Kumar and his co-authors (2025) investigated the barriers to developing artificial intelligence literacy (AI literacy) among healthcare

professionals working in Saudi Arabia's public health sector. The research aimed to examine the readiness, awareness, and competence of healthcare workers in adopting AI-based technologies and to explore the institutional and structural challenges that impede this process. The study employed a descriptive-analytical methodology, utilizing a structured questionnaire administered to a sample of 312 physicians, nurses, and administrative staff from various public hospitals across the Kingdom. The survey instrument was designed to capture both individual and organizational dimensions of AI literacy, including knowledge of AI applications, attitudes toward data sharing, and perceived institutional support for digital innovation.

The results revealed that limited data exchange between departments and the absence of unified data standards constitute significant barriers to effective AI implementation. Participants reported that fragmented data environments reduce the capacity of AI systems to learn from comprehensive datasets, ultimately diminishing the accuracy and reliability of algorithmic decision support tools. Furthermore, the study identified a deficiency in AI-related training and professional development programs, which limits the workforce's ability to interact with and leverage intelligent systems effectively. The authors emphasized the importance of cultivating a data-driven culture within healthcare institutions and proposed a national strategy to embed AI literacy within continuous professional education. They concluded that successful integration of AI requires not only technological investment but also organizational commitment to interoperability, open data sharing, and human capital development. Strengthening AI literacy across healthcare roles, the study argues, is a critical enabler for bridging existing data silos and ensuring sustainable digital transformation within the Saudi healthcare ecosystem.

3. Moafa, Kholood Mohammed Yahya, et al. (2024). "Artificial Intelligence for Improved Health Management: Applications, Uses, Opportunities, and Challenges – A Systematic Review." *Egyptian*

Moafa and her colleagues (2024) conducted a systematic review aimed at examining the current landscape of artificial intelligence (AI) applications in health management systems and exploring how AI technologies are leveraged to enhance both clinical decision-making and administrative efficiency. The study analyzed a wide range of peer-reviewed publications to identify patterns, opportunities, and constraints shaping AI implementation across different healthcare contexts. Employing a comparative analytical approach, the researchers contrasted AI utilization in developed and developing nations, highlighting disparities in digital readiness, data infrastructure, and institutional governance. The review revealed

that countries with well-integrated health information systems benefit more substantially from predictive analytics, early disease detection, and efficient resource allocation than those facing fragmented data environments.

One of the most prominent findings underscored that data fragmentation remains a central obstacle to realizing the full potential of AI in healthcare management. In developing countries, including many within the Middle East and Africa, data silos across hospitals, laboratories, and insurance systems prevent the creation of comprehensive patient records, thereby limiting the precision of AI-driven insights. Conversely, the integration of heterogeneous data sources—such as electronic medical records, imaging databases, and laboratory information systems—was found to significantly improve chronic disease prediction accuracy and optimize health resource distribution. The study also discussed technical and ethical challenges, such as data privacy, algorithmic bias, and interoperability limitations, which require coordinated regulatory frameworks and international cooperation. The authors concluded that achieving sustainable AI-driven healthcare transformation depends on strengthening health data governance, investing in interoperable systems, and fostering cross-sectoral collaboration among policymakers, IT specialists, and healthcare professionals.

4. Bayer, Mohammad & Eisawi, Amna. (2025). “Exploring the Landscape of Artificial Intelligence in Saudi Arabia’s Healthcare Sector: Current Trends and Challenges.” *Cureus*, 17(5).

Bayer and Eisawi (2025) conducted an exploratory analytical study that sought to examine the current state of artificial intelligence (AI) applications within Saudi Arabia’s healthcare sector, identifying both the prevailing trends and the major institutional and regulatory challenges impeding progress. The researchers adopted a documentary analytical methodology, drawing upon a diverse range of data sources, including official governmental policy papers, Ministry of Health strategic reports, and peer-reviewed academic publications. The findings of the study revealed a rapidly evolving but fragmented AI ecosystem in Saudi healthcare, characterized by multiple pilot projects implemented across various hospitals and health clusters without sufficient interoperability or unified data-sharing frameworks. The researchers emphasized that the absence of standardized systems for data exchange continues to pose one of the most significant barriers to advancing AI-driven digital transformation initiatives. This fragmentation limits the ability of healthcare providers to leverage predictive analytics and machine learning algorithms for integrated patient care. Furthermore, the study highlighted that the lack of a national open data infrastructure and insufficient cross-institutional coordination hinder the scalability and sustainability of AI projects.

Despite ongoing investments under Saudi Vision 2030's Health Sector Transformation Program, progress remains uneven due to disparities in digital maturity between public and private institutions.

Bayer and Eisawi also underscored that integrating AI technologies with robust health data management systems is essential for achieving operational efficiency, improving patient safety, and enhancing service quality. The authors argued that AI's success in the healthcare domain depends on establishing clear governance models, developing ethical frameworks for data use, and ensuring interoperable information architectures that facilitate real-time data flow among healthcare entities.

5. Alshararl, Fawaz Aied, et al. (2025). "Integrating Health Information Systems and Administrative Processes to Improve Patient Care Efficiency: A Comprehensive Review."

Alshararl et al. (2025) conducted a **comprehensive comparative review** examining the interrelationship between the integration of health information systems (HIS) and administrative processes in enhancing healthcare delivery efficiency. The study analyzed hospital information system models across the Gulf Cooperation Council (GCC) countries, aiming to identify the organizational and technical barriers that impede effective data exchange between clinical and administrative units. The findings revealed that the separation of administrative and clinical databases remains a major challenge within public healthcare institutions, resulting in limited coordination, duplication of efforts, and delayed decision-making. This structural fragmentation negatively affects the accuracy of clinical decisions and the overall efficiency of hospital operations. The authors highlighted that the absence of integrated data flows restricts hospitals' ability to implement evidence-based management and hinders the adoption of AI-powered systems that rely on comprehensive datasets. The study emphasized that unified data exchange frameworks are essential to ensure smooth interoperability across departments. By aligning administrative and clinical data systems, healthcare organizations can achieve more accurate performance monitoring, facilitate real-time resource allocation, and enhance the responsiveness of medical services. Furthermore, Alshararl et al. pointed out that the integration of HIS with AI-driven analytics can substantially improve predictive modeling for patient outcomes, reduce medical errors, and increase operational efficiency.

The authors concluded that achieving such integration requires both technological and managerial alignment, including standardized health data protocols, centralized governance models, and workforce training in data literacy. They also recommended that GCC health authorities adopt regional data interoperability

standards and invest in AI-supported performance dashboards to track institutional efficiency.

6. Almadani, Basem, et al. (2025). “A Systematic Survey of Distributed Decision Support Systems in Healthcare.” *Systems*, 13(3), 157.

Almadani et al. (2025) conducted a systematic survey focused on *Distributed Decision Support Systems (DDSS)* and their transformative potential in enhancing data integration within healthcare environments. The study comprehensively reviewed academic and technical literature published between 2018 and 2024, covering implementations of DDSS across hospitals, research centers, and public health institutions. Its main objective was to explore how distributed architectures and artificial intelligence (AI) technologies can overcome traditional data silos and facilitate real-time, evidence-based decision-making across healthcare networks.

Methodologically, the researchers employed a systematic literature review (SLR) framework following PRISMA guidelines to synthesize findings from over 70 peer-reviewed studies. The analysis considered the technological, structural, and ethical dimensions of distributed data systems, with particular emphasis on data interoperability, secure data sharing, and AI-driven analytics. The study’s findings highlighted that data silos remain a critical obstacle in healthcare decision support systems, leading to delays in diagnosis, inefficiencies in treatment coordination, and duplication of medical records. By contrast, DDSS especially those integrating AI and Internet of Things (IoT) technologies enable seamless, secure data exchange among multiple institutions while preserving patient privacy. Such systems allow for federated data processing, where information is analyzed locally and aggregated centrally to generate holistic insights without compromising confidentiality.

Furthermore, the authors demonstrated that combining DDSS with machine learning algorithms enhances predictive accuracy in clinical decision-making, enabling healthcare professionals to anticipate disease progression, optimize treatment pathways, and allocate resources more efficiently. Case studies from smart hospitals in the GCC and Europe showed that distributed systems reduced diagnostic delays by up to 30% and improved inter-institutional coordination.

7. Alquayt, Ala’A., et al. (2025). “AI-Driven Healthcare Innovations for Enhancing Clinical Services During Mass Gatherings (Hajj): Task Force Insights and Future Directions.” *BMC Health Services Research*, 25(1), 876.

Alquayt et al. (2025) conducted an in-depth case study examining the Kingdom of Saudi Arabia’s innovative use of artificial intelligence (AI) to enhance healthcare

service delivery during the annual Hajj pilgrimage — a uniquely challenging context characterized by high population density, diverse health conditions, and dynamic clinical demands. The study aimed to identify how AI-driven platforms could support real-time clinical decision-making, improve coordination across healthcare providers, and manage large-scale health data efficiently during mass gatherings.

The researchers adopted a qualitative case study methodology, focusing on data collected from the Ministry of Health's Hajj Health Task Force and several AI-enabled medical centers operating in Makkah and Madinah between 2022 and 2024. Multiple sources were triangulated, including official reports, patient records, and interviews with system engineers and clinical administrators.

Findings revealed that AI-powered integration of clinical and operational data significantly enhanced the capacity of healthcare teams to respond promptly to medical emergencies, optimize patient triage, and predict disease outbreaks. Machine learning algorithms were employed to analyze real-time patient inflow, detect early warning signals for infectious diseases, and support predictive resource allocation, such as ICU bed management and medical supply distribution.

One of the key contributions of the study lies in demonstrating how data silos can be effectively dismantled in complex, multi-institutional environments through the use of AI-supported data integration platforms. The study noted that prior to AI adoption, health data from different hospitals and mobile clinics were fragmented across unconnected databases, leading to delayed diagnoses and inefficiencies in patient tracking. The introduction of interoperable AI-driven systems enabled synchronized data sharing and facilitated a unified national health dashboard that improved situational awareness and decision-making across ministries and hospitals.

Furthermore, the study underscored the broader strategic implications of the Hajj experience, suggesting that the AI-based health data integration model can serve as a prototype for national-level healthcare transformation under Saudi Vision 2030. The authors emphasized the importance of expanding this model beyond Hajj-specific operations to encompass all public hospitals, thereby establishing a nationwide smart healthcare ecosystem that promotes interoperability, transparency, and patient safety.

2.3.2 Studies on Open Health Data and Interoperability

1. Lichtenauer, Norbert, et al. (2023). “A Scoping Review on Analysis of the Barriers and Support Factors of Open Data.” *Information*, 15(1), 5.

Lichtenauer et al. (2023) conducted a comprehensive scoping review exploring the barriers and enablers of open data implementation across multiple sectors, with particular attention to its applications in public service and healthcare. The study aimed to synthesize the existing body of knowledge on how open data initiatives can be effectively designed, governed, and sustained within institutional frameworks. Using the Arksey and O'Malley methodological framework for scoping reviews, the authors systematically analyzed 124 peer-reviewed studies published between 2010 and 2022, emphasizing cross-sectoral insights that can inform healthcare interoperability efforts. The review identified several key barriers to open data adoption, including a lack of standardized data-sharing protocols, institutional resistance to transparency, concerns about data privacy and intellectual property, and limited technical infrastructure. These barriers were found to be particularly pronounced in the healthcare domain, where data sensitivity and fragmented governance structures often hinder information exchange. The study also underscored that organizational culture and policy coherence play a central role in either facilitating or obstructing open data initiatives. On the enabling side, Lichtenauer et al. (2023) highlighted that clear regulatory framework, the establishment of data governance standards, and the implementation of open APIs (Application Programming Interfaces) are critical success factors for achieving interoperability. The review pointed to several successful models, such as the European Open Data Strategy and the UK's National Health Service (NHS) Digital Transformation Framework, as global benchmarks demonstrating how open data can drive **innovation, efficiency, and accountability** in healthcare systems.

2. Chen, Xieling, et al. (2024). “Artificial Intelligence and Multimodal Data Fusion for Smart Healthcare: Topic Modeling and Bibliometrics.” *Artificial Intelligence Review*, 57(4), 91.

Chen et al. (2024) conducted a **comprehensive bibliometric and topic modeling analysis** exploring the convergence of **artificial intelligence (AI)** and **multimodal data fusion** in the context of **smart healthcare systems**. The study sought to identify major research trends, thematic clusters, and emerging technologies that enable effective **integration of heterogeneous health data sources**, including electronic health records (EHRs), wearable sensor data, and medical imaging.

Using bibliometric mapping tools such as *VOSviewer* and *CiteSpace*, the authors analyzed over **2,500 peer-reviewed articles published between 2013 and 2023**, identifying key domains of innovation such as data interoperability, semantic harmonization, and context-aware analytics. The results revealed that **multimodal data fusion**—the process of integrating structured, semi-structured, and unstructured medical data—enhances diagnostic precision, patient monitoring, and personalized treatment.

The study emphasized that the **integration of open health data** with AI-driven analytics creates significant opportunities for predictive modeling and clinical decision support. However, the authors also highlighted persistent challenges related to **data silos, inconsistent metadata standards, and limited interoperability frameworks**. These findings underscore the importance of establishing **semantic ontologies and standardized health data models**, such as HL7 FHIR and SNOMED CT, to ensure the seamless flow of information across digital health systems.

Chen et al. (2024) concluded that **open and interoperable data ecosystems** are essential for the success of AI-based healthcare transformation. They proposed a conceptual framework for **multimodal data integration** grounded in machine learning, ontology mapping, and blockchain-enabled data validation. The study's implications extend to healthcare policymakers and system developers aiming to **leverage open data for smarter, more connected healthcare environments**, aligning with the digital transformation goals of Vision 2030.

3. Abdalzaher, Mohamed S., et al. (2025). “Quality-Focused Internet of Things Data Management: A Survey, Perspectives, Open Issues, and Challenges.” *IEEE Internet of Things Journal* (2025).

Abdalzaher et al. (2025) presented a systematic survey focusing on quality-driven Internet of Things (IoT) data management, with a particular emphasis on healthcare environments where sensor-based data collection plays a critical role. The study aimed to investigate the challenges of maintaining data quality, accessibility, and interoperability across IoT networks that feed into open health data platforms.

Drawing on over 180 published studies, the authors categorized IoT data management challenges into four dimensions: data acquisition, preprocessing, integration, and governance. The research found that heterogeneity of data sources, inconsistent timestamping, and lack of real-time synchronization often lead to data fragmentation and accuracy issues. In healthcare contexts, these shortcomings can severely undermine diagnostic reliability and delay clinical interventions.

The study further noted that adopting AI-powered data cleaning and integration algorithms, coupled with open data standards, can significantly improve the utility and reliability of IoT-generated health information. The authors emphasized that interoperability and transparency are key to enabling seamless integration between IoT systems and national health data infrastructures.

Ultimately, Abdalzaher et al. (2025) proposed a multi-layered data governance framework incorporating AI-based validation, distributed ledger technologies (blockchain), and standardized open APIs to ensure data integrity, traceability, and reusability. This framework represents a pivotal step toward realizing trustworthy open health data ecosystems that empower precision medicine and digital health transformation.

4. Nguyen, Tuan Anh (Ed.). (2025). *Blockchain and Digital Twin for Smart Hospitals*. Elsevier, 2025.

In this edited volume, Nguyen (2025) brings together contributions from experts in digital health, blockchain, and artificial intelligence to explore the integration of blockchain and digital twin technologies in smart hospitals. The book offers a multi-disciplinary perspective on how decentralized data management and real-time digital modeling can enhance interoperability, transparency, and operational efficiency in healthcare systems. The chapters collectively address the potential of blockchain-enabled open data frameworks to mitigate data silos, protect patient privacy, and facilitate secure information exchange across healthcare networks. Simultaneously, digital twin technology—which creates virtual replicas of physical hospital systems—enables continuous monitoring, simulation, and optimization of patient flows, equipment usage, and facility management.

Nguyen (2025) emphasizes that the synergy between blockchain and open health data supports both data sovereignty and cross-institutional collaboration, allowing healthcare organizations to maintain control over their data while contributing to national and global health databases. The volume also presents practical case studies from the UAE, Singapore, and Saudi Arabia, demonstrating how these technologies are being adopted to achieve smart, connected, and sustainable healthcare systems. The book concludes with a forward-looking framework suggesting that AI, blockchain, and digital twins together represent the technological foundation of the next generation of interoperable health ecosystems. This aligns directly with efforts to realize Vision 2030's objectives of digital transformation and healthcare innovation in the Middle East.

2.3.3 Integrated Frameworks and Comparative Insights for Healthcare Transformation

1. Awad, Wael A., & Amena Mahmoud. (2025). *The Intersection of AI, Cloud Computing, and Healthcare: A Comprehensive Overview*.

Awad and Mahmoud (2025) present a synthetic, integrative account of how **artificial intelligence and cloud computing** jointly enable scalable, resilient, and interoperable healthcare infrastructures. The chapter adopts a conceptual and evidence-synthesis approach, reviewing technological architectures, deployment models, and governance considerations across a range of health-system case studies. The authors argue that cloud-native AI platforms provide the necessary elasticity for processing multimodal health datasets and for deploying federated and containerized analytics at scale. Key conclusions emphasize that (a) cloud infrastructures reduce barriers to cross-institutional model deployment; (b) coupling AI with cloud APIs and standardized data services accelerates interoperability; and (c) strong governance layers (access control, audit trails, legal agreements) are required to translate technical integration into sustainable clinical value. The chapter thus provides a practical architectural template for integrated frameworks that combine AI, cloud services, and data governance to support national digital-health transformation.

2. Kumar, M. V., Patil, J., Shastry, K. A., Darshan, S., Sastry, N. K. B., Moonesar, I. A., Atalla, S., Almuraqab, N., & Rao, A. (2022/2023). *ICT Enabled Disease Diagnosis, Treatment and Management — A Holistic Cost-Effective Approach Through Data Management and Analysis*.

Kumar MV et al. synthesize empirical and applied work on ICT-enabled diagnostic and management pipelines, stressing governance mechanisms for data stewardship and AI lifecycle management. Methodologically, the work combines case-based evaluations from UAE/India pilot deployments with prescriptive governance guidelines. The authors highlight that successful integration of AI into clinical workflows depends on end-to-end governance: data provenance, metadata standards, model validation, continuous monitoring, and economic sustainability. Empirical evidence from pilot sites demonstrated improved diagnostic throughput and cost-effectiveness when ICT platforms enforced consistent data models and logging practices. The work contributes a pragmatic governance-oriented framework that links technical standards with institutional processes a model readily adaptable for comparative studies across healthcare systems.

3. Solaiman, Barry, et al. (2025). “A ‘True Lifecycle Approach’ towards Governing Healthcare AI with the GCC as a Global Governance Model.”

Solaiman and colleagues (2025) propose a lifecycle governance model for healthcare AI that spans design, validation, deployment, monitoring, and retirement. The paper uses policy analysis and stakeholder interviews across GCC health authorities to construct an operational governance framework emphasizing transparency, accountability, and stage-specific safeguards (e.g., pre-deployment fairness testing, real-world performance monitoring). Findings indicate that a lifecycle approach reduces regulatory blind spots and better aligns AI interventions with clinical safety and ethical norms. For GCC contexts, the study recommends harmonized regional standards, certification pathways for AI tools, and mandated data-sharing agreements to support cross-border learning and federated initiatives. The lifecycle model is particularly relevant when integrating AI into national interoperability strategies, since it prescribes both technical and institutional controls required for sustainable scale-up.

4. Abdelrahman Hussein, Selma, & Emna Chikhaoui. (2025). “The Silicon Oasis: IP-Enabled Generative AI Innovations for Saudi Arabia’s Environmental Vision.” In *Generative AI and Large Language Models* (pp. 263–310). Springer Nature.

Although focused on generative AI applications within environmental and industrial domains, Abdelrahman Hussein & Chikhaoui (2025) offer important strategic lessons for healthcare policymakers regarding innovation governance and intellectual-property (IP) frameworks. Through qualitative analysis and scenario modeling, the chapter illustrates how IP-sensitive governance can catalyze local innovation ecosystems while protecting public-data commons. Key insights applicable to health-sector transformation include the need for clear licensing models for trained AI assets, incentives for public–private collaboration, and mechanisms to ensure that generative models trained on open datasets respect confidentiality and attribution norms. The chapter recommends policy instruments such as shared data trusts and standardized model-licensing templates that can be adapted to the healthcare sector to support ethical, innovation-friendly ecosystems.

5. Kumar, Arun, et al. (2025). “Integrating 6G Technology in Smart Hospitals: Challenges and Opportunities for Enhanced Healthcare Services.” *Frontiers in Medicine*, 12, 1534551.

Kumar et al. (2025) explore forward-looking infrastructure that combines next-generation connectivity (6G) with AI-enabled clinical systems to enable ultra-low-latency data exchange, pervasive sensing, and real-time digital twins of hospital

operations. The study employs a mixed-methods design combining simulation experiments and stakeholder consultations. Results demonstrate potential benefits—real-time high-fidelity imaging transfer, seamless remote interventions, and ambient intelligence for patient monitoring—while also identifying integration challenges: legacy system heterogeneity, regulatory lag, and capital intensity. The work situates advanced communication technologies as enabling layers within an integrated framework, noting that interoperability and governance must be addressed in tandem for technical advances to yield clinical and organizational value.

6. Nguyen, Tuan Anh (Ed.). (2025). *Blockchain and Digital Twin for Smart Hospitals*. Elsevier.

Nguyen's edited volume compiles empirical and conceptual chapters on blockchain, digital twins, and their convergent roles in constructing resilient, auditable, and interoperable hospital ecosystems. Contribution's document piloted architectures where blockchain provides immutable provenance and consent management for shared records, while digital twins model hospital processes to optimize resource flows. Case studies in the book illustrate how these technologies support trustworthy open-data exchanges and simulation-driven policy testing prior to deployment. The principal implication for integrated frameworks is that decentralized ledgers and virtual-replica modeling can materially enhance both the technical integrity and the governance transparency of national health-data platforms, thereby reducing friction in cross-institutional data sharing and enabling more confident AI model orchestration.

7. Regin, R., & S. Suman Rajest. (2024). "Leveraging Data-Driven Intelligence for Remote Patient Monitoring in Digital Health." *Central Asian Journal of Medical and Natural Science*, 5(4), 798–808.

Regin & Rajest (2024) examine practical implementations of remote patient monitoring (RPM) programs that integrate wearable sensors, cloud analytics, and AI to support chronic-care pathways. Using program evaluation methods across several pilot deployments, the authors show that integrating open and semi-open datasets (EHR extracts + patient-generated data) with AI models improves early-deterioration detection and reduces avoidable hospitalizations. The study emphasizes governance enablers — consent mechanics, API-based interoperability, and feedback loops into clinician workflows — as prerequisites for RPM success. The findings reinforce the comparative argument that technical integration alone is insufficient: organizational redesign, clinician engagement, and defined data governance are required to translate integrated frameworks into measurable health outcomes.

2.4 summary of the Theoretical and Empirical Frameworks

1. Synthesis of Key Findings from Reviewed Studies

The reviewed body of literature collectively underscores that the digital transformation of healthcare is deeply shaped by the degree of data integration, openness, and the maturity of artificial intelligence (AI) ecosystems. Studies on data silos and AI integration (Kumar et al., 2025; Almadani et al., 2025; Alquayt et al., 2025) consistently highlight that the fragmentation of health data across institutions remains one of the most significant obstacles to deploying AI-driven solutions effectively. These studies converge on the notion that interoperability—both technical and semantic—is a prerequisite for realizing the full potential of predictive analytics, clinical decision support, and resource optimization.

The empirical studies on open health data and interoperability (Lichtenauer et al., 2023; Chen et al., 2024; Abdalzaher et al., 2025; Nguyen, 2025) have established that data openness fosters innovation by enabling secondary uses of health information, such as disease surveillance, personalized medicine, and policy evaluation. However, these studies also emphasize that achieving sustainable open data environments requires robust governance mechanisms, data quality assurance, and clear ethical frameworks to safeguard patient privacy. The theoretical discourse further situates open data within the paradigm of knowledge democratization, where data transparency serves as a catalyst for collaborative research and citizen participation in health systems.

Moreover, the literature on integrated frameworks for healthcare transformation (Awad & Mahmoud, 2025; Solaiman et al., 2025; Regin & Rajest, 2024) identifies the convergence of AI, cloud computing, and blockchain as the foundation for next-generation health systems. These frameworks suggest that digital transformation is not merely a technological transition but a holistic reform encompassing organizational learning, governance maturity, and cultural adaptation. Collectively, the reviewed studies illustrate a multi-layered theoretical evolution—from isolated digital initiatives toward ecosystem-based models of health innovation that are adaptive, interoperable, and data-driven.

2. Identified Strengths, Limitations, and Research Gaps

The reviewed theoretical and empirical contributions present several notable strengths. First, they provide robust conceptual clarity regarding the role of data integration and AI in driving healthcare efficiency and sustainability. Second, many studies have adopted methodologically rigorous approaches, including systematic literature reviews, cross-national comparisons, and mixed-methods analyses that enhance the generalizability of findings. Third, recent literature

(2024–2025) has advanced from descriptive analyses to more framework-oriented research, offering models that align digital transformation with governance, ethics, and policy-making dimensions.

Nevertheless, important limitations persist. Most existing studies are context-specific, often focusing on high-income or technologically advanced nations, leaving limited empirical evidence from developing or transitional healthcare systems—particularly in the Arab region. Furthermore, a significant portion of the literature treats AI adoption and open data as separate streams of inquiry, overlooking their mutual reinforcement in shaping smart healthcare ecosystems. In addition, there is an evident methodological bias toward technical evaluations, while fewer studies address organizational readiness, capacity building, and socio-cultural barriers that influence digital transformation outcomes.

From these limitations emerge several research gaps. There remains a lack of integrated analytical models that empirically test how AI integration interacts with open health data ecosystems to enhance healthcare innovation, equity, and resilience. Similarly, the governance dimension—covering data standards, accountability, and ethical use of AI—has not been sufficiently operationalized in empirical research. Therefore, the current study aims to fill this gap by developing and testing a conceptual model that links AI-driven data integration, open data governance, and healthcare transformation within the context of national digital health strategies.

3. Conceptual Model of the Current Research

Based on the synthesized theoretical insights and empirical evidence, the present study proposes a holistic conceptual model that explains the dynamic relationship between AI integration, open health data, and healthcare transformation. The model is anchored in three foundational constructs:

1. AI Integration Capacity – representing the technological, organizational, and human readiness to embed AI into health systems for data processing, prediction, and decision-making.
2. Open Health Data Ecosystem – reflecting the level of interoperability, accessibility, and transparency of health data, supported by governance frameworks that ensure privacy and ethical compliance.
3. Healthcare Transformation Outcomes – encompassing indicators such as service efficiency, quality of care, innovation capacity, and stakeholder satisfaction.

The model assumes that AI integration mediates the relationship between data openness and healthcare transformation: as open and interoperable data environments expand, AI systems gain access to richer, more diverse datasets that improve learning accuracy and innovation outcomes. Conversely, robust AI infrastructure can accelerate the standardization and utilization of open data by creating value through analytics and intelligent automation. Governance mechanisms spanning data stewardship, accountability, and capacity development serve as moderating factors that determine the extent to which this synergy translates into sustainable healthcare transformation.

2.5 Summary of Chapter Two

This chapter has provided a comprehensive review of the theoretical foundations and empirical evidence related to data silos, artificial intelligence (AI) integration, and open health data within the context of healthcare digital transformation. The theoretical framework was built upon three interrelated perspectives—the Systems Approach, Institutional Theory, and Data Governance Theory—which together explain the structural, organizational, and data-centric dimensions of digital health integration. These theories collectively support the assumption that overcoming data fragmentation and fostering interoperability are central to achieving intelligent, sustainable healthcare ecosystems aligned with national digital strategies such as Saudi Vision 2030.

The discussion of data silos in healthcare highlighted the persistence of fragmented information systems that hinder data accessibility, quality, and analytical utility. The reviewed literature revealed that such fragmentation not only reduces the effectiveness of AI applications but also limits evidence-based decision-making and patient-centered care. Similarly, the section on AI integration demonstrated the growing potential of machine learning, predictive analytics, and intelligent automation in enhancing diagnostic accuracy and operational efficiency—provided that data environments are sufficiently unified and standardized.

Furthermore, the review of open health data and interoperability underscored the dual role of data openness as both a technical enabler and a governance challenge. While open data initiatives can promote transparency, collaboration, and innovation, they also require strong institutional frameworks to manage ethical, privacy, and quality concerns. The relationship between AI and open data was conceptualized as mutually reinforcing: AI thrives on diverse and accessible datasets, while open data ecosystems benefit from AI-driven insights that enhance data utilization and value creation.

The section on previous studies synthesized recent empirical findings from global and Saudi contexts. The reviewed studies confirmed that integrating AI with open and interoperable data infrastructures has tangible impacts on healthcare outcomes—improving efficiency, accuracy, and public health management. However, they also revealed significant research gaps, particularly the lack of comprehensive models that empirically link AI, data openness, and healthcare transformation in developing or transitional systems.

In conclusion, this chapter established both the theoretical basis and empirical rationale for the present research. It emphasized the necessity of an integrated analytical model that connects AI-driven data integration and open data governance as dual enablers of healthcare digital transformation. The insights derived from this literature review directly inform the conceptual model and hypotheses presented in the next chapter, which will outline the research design, methodology, and analytical approach adopted to investigate these relationships within the Saudi healthcare system.

Chapter Three

Study Methodology

3.1 Introduction:

This chapter outlines the methodological framework adopted to achieve the objectives of the current study, which investigates the role of artificial intelligence (AI) in reducing data silos and enhancing the integration of open health data in Saudi public hospitals. The methodology serves as the foundation that guides the systematic collection, organization, and analysis of data to ensure the accuracy, validity, and reliability of the study's outcomes.

Given the complex and multidimensional nature of healthcare data management, this study employs a quantitative research design supported by empirical evidence obtained from healthcare professionals, IT specialists, and administrative staff working in Saudi public hospitals. The research design is structured to capture the perceptions, behaviors, and practices of participants in dealing with health data, focusing particularly on the technological, organizational, and administrative aspects that contribute to data fragmentation or integration.

This chapter begins by describing the research approach and justifying its suitability for the study. It then elaborates on the research population and sampling procedures, followed by the design and validation of the research instrument (questionnaire). Additionally, the data collection methods, measurement scales, and techniques of statistical analysis are discussed in detail to ensure methodological rigor and transparency.

The integration of multiple analytical dimensions technological, administrative, and regulatory within the methodological framework enables a holistic examination of the challenges and potential solutions for data interoperability. Furthermore, by aligning the methodological procedures with the principles of Vision 2030, the study ensures that the findings not only contribute academically but also provide actionable insights for policymakers and healthcare leaders in Saudi Arabia.

3.2 Research Design

This study adopts a descriptive–analytical research design supported by exploratory elements. The descriptive aspect aims to accurately identify and describe the current status of data silos within public hospitals in Saudi Arabia, as well as the extent to which open health data is available and utilized. The analytical dimension focuses on examining relationships between variables—particularly the impact of technological, administrative, and regulatory factors on data fragmentation and interoperability. The exploratory element, meanwhile, allows the researcher to investigate emerging technologies, such as artificial intelligence (AI), as potential tools for addressing existing data challenges and enhancing data

integration within healthcare systems. The selection of this design is justified by the complex and multidimensional nature of the research problem, which requires both factual description and empirical analysis to produce meaningful insights. By combining descriptive and analytical approaches, the study can bridge the gap between theory and practice offering a comprehensive understanding of how AI-driven solutions can mitigate data silos while supporting the objectives of Saudi Vision 2030 in healthcare digital transformation.

The research framework follows a quantitative methodology based on a structured questionnaire as the main data collection tool. This design ensures objectivity, replicability, and statistical validity in analyzing participants' perceptions, behaviors, and practices concerning health data management. The collected data will be statistically processed to test the study hypotheses and evaluate the relationships among variables within the proposed conceptual model.

3.3 Research Population and Sample

3.3.1 Study Population

The study population consists of employees working in public hospitals across various regions of Saudi Arabia, including healthcare practitioners, administrative personnel, and information technology specialists. This population was selected because it represents the primary group directly involved in data generation, storage, and utilization processes within the healthcare system. These individuals are most affected by data silos and are key stakeholders in implementing AI-driven and interoperable data systems. The inclusion of multiple functional categories (medical, administrative, and technical) ensures that the study captures a comprehensive and multi-dimensional view of data management practices and challenges across different levels of healthcare operations.

3.3.2 Sampling Method and Size

This study adopts a stratified random sampling technique to ensure balanced representation of participants from different hospital departments and professional roles. The stratification was based on three main categories: (1) medical and nursing staff, (2) administrative employees, and (3) IT and data management specialists.

A total of 350 questionnaires were distributed electronically to employees across ten major public hospitals located in Riyadh, Jeddah, Dammam, Medina, and other regions. Out of these, 310 valid responses were received and included in the final analysis, representing a response rate of approximately 88.6%.

The sample size was determined using Cochran's formula for large populations, ensuring an acceptable margin of error ($\pm 5\%$) and confidence level (95%). This sample is considered sufficient for conducting descriptive and inferential statistical analyses.

3.3.3 Demographic Profile

The demographic characteristics of the participants were analyzed to provide context for interpreting the study results. The key demographic variables included gender, age group, job title, educational qualification, years of professional experience, and department affiliation.

A summary of the demographic distribution is presented below:

Table (3.1): Demographic Characteristics of the Study Participants

Variable	Category	Frequency (n)	Percentage (%)
Gender	Male	182	58.7
	Female	128	41.3
Age Group	20–29 years	64	20.6
	30–39 years	132	42.6
	40–49 years	86	27.7
	50 years and above	28	9.0
Job Title	Medical Staff	138	44.5
	Administrative Staff	104	33.5
	IT Specialists	68	22.0
Qualification	Diploma	46	14.8
	Bachelor's Degree	164	52.9
	Master's Degree	72	23.2
	Doctorate	28	9.0
Experience (Years)	Less than 5 years	52	16.8
	5–10 years	116	37.4
	11–15 years	94	30.3
	More than 15 years	48	15.5
Department	Medical	142	45.8
	Administration	102	32.9
	IT Department	66	21.3

The demographic data presented in the table provide a comprehensive overview of the study participants and contribute to a better understanding of the sample composition. The results indicate that males constituted the majority of respondents (58.7%), compared to females (41.3%), suggesting a gender imbalance that may reflect the general workforce distribution in the studied sector.

In terms of age, the largest proportion of participants (42.6%) fell within the 30–39 years age group, followed by 27.7% aged 40–49 years, while younger participants aged 20–29 years represented 20.6%. This indicates that the majority of respondents are in the mid-career stage, with substantial professional

experience. Regarding job titles, medical staff accounted for the highest percentage (44.5%), followed by administrative staff (33.5%) and IT specialists (22%). This distribution shows that the sample included a balanced representation of key functional categories within the institutions under study.

As for educational qualifications, more than half of the participants (52.9%) held a bachelor's degree, 23.2% had a master's degree, and 9% held a doctorate, while only 14.8% possessed a diploma. This demonstrates a relatively high level of educational attainment among the respondents.

Concerning professional experience, 37.4% had between 5 and 10 years of experience, and 30.3% had between 11 and 15 years, suggesting that most participants had a moderate to high level of experience in their respective fields.

Finally, regarding departmental distribution, 45.8% of respondents were from the medical department, 32.9% from administration, and 21.3% from IT. This proportional representation ensures that the perspectives of different professional categories are well-reflected in the study findings.

3.4 Research Variables

This study investigates the role of AI-driven data integration in mitigating data silos and enhancing healthcare efficiency within Saudi public hospitals. The research framework is designed to explain the relationships between the independent and dependent variables through a multidimensional model.

3.4.1 Independent Variable: Data Silos and Open Health Data

The independent variable represents the extent and impact of data silos, fragmentation, and the accessibility of open health data. It is measured through three main dimensions:

1. **Technological Fragmentation:** refers to the lack of system interoperability, outdated IT infrastructure, and insufficient data linkage among hospital units.
2. **Organizational and Administrative Fragmentation:** reflects policy inconsistencies, weak coordination, and limited managerial support for data integration.
3. **Open Health Data Quality and Availability:** focuses on the accessibility, standardization, and usability of public health datasets for research and operational purposes.

3.4.2 Dependent Variable: Clinical and Administrative Efficiency

The dependent variable captures the level of efficiency achieved in both clinical and administrative domains as a result of improved data integration and AI utilization. It consists of two main dimensions:

1. **Operational Efficiency:** reflects improvements in workflow, speed of decision-making, and coordination across departments.
2. **Data Usage and Integration Practices:** measures the extent to which AI tools and interoperable platforms are adopted to support evidence-based management and healthcare delivery.
 - interoperability across public healthcare institutions.

3.5 Research Instrumentation

3.5.1 Data Collection Tool

The primary tool used for data collection in this study was an electronic questionnaire, designed to gather responses from employees working in public hospitals across Saudi Arabia. The questionnaire was developed based on an extensive review of previous studies and relevant literature in the fields of health data management, digital transformation, and the integration of artificial intelligence (AI) within healthcare systems. It consists of six main dimensions, aiming to assess participants' perceptions, behaviors, and practices related to health data management, with a particular focus on technological, organizational, and administrative aspects, as well as their impact on operational efficiency and data integration.

3.5.2 Tool Structure

The questionnaire is divided into two main sections:

- **Section One:** Demographic information, including gender, age, job title, educational qualification, years of experience, and type of hospital.
- **Section Two:** Study dimensions, which are summarized as follows:

Dimension No.	Dimension	Number of Items	Code
1	Technological Fragmentation	10	T1–T10
2	Organizational & Administrative Fragmentation	10	O1–O10

3	Open Health Data Quality & Availability	10	H1– H10
4	Operational Efficiency	10	E1– E10
5	Data Usage & Integration Practices	10	D1– D10
Total		50 items	

A **five-point Likert scale** was employed to measure participants’ responses, ranging from (1) “Strongly Disagree” to (5) “Strongly Agree.”

3.5.3 Validity and Reliability

First: Validity

To ensure the validity of the research instrument, two main approaches were adopted as follows:

1. **Content Validity:** The questionnaire items were developed based on established literature and previous empirical studies in the fields of health data management, digital transformation, and artificial intelligence integration in healthcare systems, such as those by Kumar et al. (2025); Moafa et al. (2024); Alquayt et al. (2025); Bayer & Eisawi (2025); and Alshararl et al. (2025). The instrument was further reviewed by a panel of academic experts to verify the clarity, precision, and appropriateness of each statement for the study dimensions, ensuring an accurate representation of the concepts of health data management and AI-driven institutional integration.
2. **Construct Validity:** Pearson correlation coefficients were calculated between each item and its corresponding dimension to assess the internal consistency of the instrument. The results showed that all correlation coefficients exceeded (0.70), indicating a high degree of internal consistency and construct validity. This finding aligns with recent research emphasizing the importance of statistical consistency verification when developing measurement tools in digital health studies (Almadani et al., 2025; Regin & Rajest, 2024; Nguyen, 2025; Wang et al., 2021).

Second: Reliability

The internal consistency of the questionnaire was assessed using Cronbach’s Alpha coefficient for each dimension, as shown in the table below:

Table 3.2Cronbach’s Alpha Coefficients and Reliability Levels for the Study Instrument Dimensions

Dimension No.	Dimension	Number of Items	Cronbach’s Alpha
1	Technological Fragmentation	10	0.89
2	Organizational & Administrative Fragmentation	10	0.87
3	Open Health Data Quality & Availability	10	0.91
4	Operational Efficiency	10	0.88
5	Data Usage & Integration Practices	10	0.90
Overall Instrument	All Dimensions Combined	50	0.93

These values indicate that the research instrument demonstrates excellent reliability and internal consistency, confirming its appropriateness for statistical analysis and hypothesis testing.

3.6 Data Collection Procedures

The process of data collection was carried out in several systematic stages to ensure the accuracy, validity, and reliability of the collected information. The steps are summarized as follows:

1. **Instrument Design:** The questionnaire was designed based on a comprehensive review of the literature related to health data management, digital transformation in healthcare, and the integration of artificial intelligence. The initial version included statements covering technological, organizational, and administrative aspects.
2. **Expert Review and Validation (Jury Evaluation):** The questionnaire was presented to a panel of academic experts and specialists in health informatics and research methodology to assess the clarity, relevance, and comprehensiveness of each item. Their comments and suggestions were taken into account to refine the instrument, ensuring content validity.

3. **Pilot Study:** A pilot test was conducted on a small sample (approximately 15 participants) from public hospitals to ensure that the questionnaire items were understandable and to identify any potential issues. Feedback from this stage helped to finalize the wording of the questions and the order of dimensions.
4. **Field Application:** The finalized questionnaire was distributed electronically to employees working in public hospitals across Saudi Arabia, including physicians, nurses, administrative staff, and IT specialists. The purpose and confidentiality of the responses were clearly explained to all participants before completion.
5. **Follow-up and Data Monitoring:** Follow-up reminders were sent to participants to increase the response rate. The researcher also monitored the progress of data collection to ensure a balanced representation across job categories and hospital types.
6. **Data Coding and Entry:** After data collection was completed, all responses were coded numerically according to the Likert scale (1–5) and entered into the Statistical Package for the Social Sciences (SPSS) for analysis. Data were checked for completeness and accuracy before beginning the statistical tests.

Through these procedures, the researcher ensured that the collected data were valid, reliable, and suitable for addressing the research objectives and testing the proposed hypotheses.

3.7 Data Analysis Techniques

The data collected from the questionnaire were analyzed using the Statistical Package for the Social Sciences (SPSS, Version 25). Several statistical methods were employed to examine the relationships between the study variables and to test the research hypotheses.

1. **Descriptive Analysis:** Descriptive statistics, including frequencies, percentages, means, and standard deviations, were used to describe the demographic characteristics of the participants and to summarize their responses to each dimension of the questionnaire (technological fragmentation, organizational and administrative fragmentation, open health data, and clinical efficiency).
2. **Validity and Reliability Testing:** The Pearson correlation test was used to verify the construct validity of the instrument by examining the correlation between each item and its respective dimension. All correlation coefficients

were greater than 0.70, indicating a strong relationship and confirming the internal consistency of the questionnaire. In addition, Cronbach's Alpha coefficients were calculated to assess the reliability of each dimension. All values exceeded 0.80, which reflects a high level of internal consistency.

3. **Inferential Analysis:** To test the research hypotheses, several inferential techniques were applied:
 - **Correlation Analysis** was used to identify the relationships between the independent and dependent variables.
 - **Multiple Regression Analysis** was applied to assess the effect of technological, organizational, and administrative fragmentation on data integration and efficiency in public hospitals.
 - **ANOVA tests** were also used to examine whether demographic characteristics (e.g., job title, years of experience) significantly influenced participants' perceptions of data management practices.
4. **Statistical Significance:** All tests were conducted at a significance level of 0.05 ($p \leq 0.05$) to determine whether the relationships observed were statistically significant.

These analytical procedures provided a comprehensive understanding of how technological, organizational, and administrative factors influence the effectiveness of data management and integration in Saudi public hospitals.

3.8 Study Boundaries

1. Spatial Boundaries This study is limited to public hospitals in the Kingdom of Saudi Arabia, focusing on analyzing the health data environment within these institutions. Private hospitals and non-governmental health centers are excluded due to differences in their digital infrastructure and the level of integration among their information systems compared to public hospitals.

2. Temporal Boundaries The study covers data, statistics, and information related to public hospitals during recent years, specifically from 2025, to align with recent digital transformations in the healthcare sector and the implementation phase of Saudi Vision 2030 initiatives.

3. Subject Boundaries The study focuses on the issue of data fragmentation (Data Silos) in public hospitals, examining technical, organizational, and legal causes of this challenge. It also assesses the role of artificial intelligence and open health data in enhancing data integration, accessibility, and analysis. Other aspects of healthcare, such as employment policies or healthcare financing, are only

addressed to the extent necessary to support understanding of the data-related problem.

4. Methodological Boundaries The study employs an analytical and applied approach, including a review of previous literature, examination of international models, and analysis of available data on integrative and open health information systems in Saudi Arabia. Direct laboratory experiments on hospital systems or development of new software are not included; the focus is on applied models and evidence-based strategic recommendations.

Chapter Four

Results and Recommendations

4.1 Introduction

This chapter presents the results of the field study conducted to examine the perceptions, behaviors, and practices of employees in public hospitals in Saudi Arabia regarding health data management, with a specific focus on technological, organizational, and administrative aspects. The chapter also explores how the integration of Artificial Intelligence (AI) tools contributes to improving data accessibility, interoperability, and decision-making efficiency within healthcare institutions. The analysis begins with a description of the data preparation and coding procedures, followed by a presentation of the demographic characteristics of the study participants. Subsequently, descriptive statistics are provided for each of the questionnaire's main dimensions to identify the general trends and attitudes of the respondents.

Furthermore, the chapter includes the results of hypothesis testing, which aim to determine the relationships between the main study variables and evaluate the impact of technological, organizational, and administrative factors on the effectiveness of health data management. The findings are then discussed in light of previous studies and the theoretical framework, emphasizing both the consistencies and discrepancies with prior research. Finally, the chapter concludes with a set of practical recommendations intended to enhance the integration of data systems, strengthen institutional coordination, and support the use of AI-based tools in the Saudi healthcare sector to achieve higher levels of efficiency, accuracy, and transparency in health data management.

4.2 Descriptive Statistics for Questionnaire Dimensions

This section presents the descriptive statistical analysis of the questionnaire dimensions used to assess the perceptions and practices of healthcare professionals regarding health data management in public hospitals in Saudi Arabia. The descriptive analysis aims to summarize the responses for each dimension, highlighting the central tendencies and dispersion measures such as means, standard deviations, and relative weights.

The questionnaire consisted of five major dimensions:

1. **Technological Fragmentation** – addressing issues related to system interoperability, data accessibility, and the technological infrastructure supporting data integration.
2. **Organizational and Administrative Fragmentation** – focusing on coordination, leadership support, policies, and communication within and across departments.

3. **Open Health Data Quality and Availability** – evaluating the accessibility, accuracy, standardization, and usability of open health data.
4. **Operational Efficiency** – measuring the extent to which hospitals achieve coordination, speed, and effectiveness in clinical and administrative operations.
5. **Data Usage and Integration Practices** – assessing how staff utilize integrated platforms, AI tools, and standardized procedures for effective data-driven decision-making.

For each dimension, mean scores and standard deviations were computed based on participants' responses using a five-point Likert scale ranging from (1 = Strongly Disagree) to (5 = Strongly Agree). A higher mean value indicates stronger agreement or a more favorable perception toward the corresponding construct.

The results presented in the following tables provide a detailed overview of respondents' attitudes and experiences across all dimensions, forming the basis for subsequent inferential analyses and hypothesis testing.

4.2.1 Technological Fragmentation

Table (4. 1): Descriptive Statistics for the Dimension “Technological Fragmentation”

No.	Statement	Mean	Std. Deviation	Rank	Level
T1	Clinical data is stored in systems separate from administrative records.	4.21	0.71	2	High
T2	Difficulty accessing patient data due to multiple databases.	4.09	0.77	4	High
T3	Current systems do not allow real-time data sharing between departments.	4.25	0.68	1	High
T4	Health data is not standardized across different hospitals.	3.98	0.83	6	Moderate–High
T5	Linking data from laboratories to emergency departments is difficult.	4.05	0.80	5	High
T6	The electronic system does not support integrated data analysis.	3.87	0.86	7	Moderate
T7	Stored data is often outdated or incomplete.	3.75	0.92	9	Moderate
T8	Current systems do not provide APIs for data exchange.	3.80	0.89	8	Moderate
T9	Middleware is needed to transfer data between different systems.	4.00	0.78	6	High
T10	The technology used is outdated and does not support smart integration.	4.11	0.75	3	High

In light of the results presented in Table (4.1), which displays the descriptive statistics for the dimension “Technological Fragmentation,” it is evident that there is a high level of technological challenges hindering data integration within hospitals. This finding highlights clear gaps in the technological infrastructure and health information systems.

Item (T3), “Current systems do not allow real-time data sharing between departments,” recorded the highest mean (4.25) with a standard deviation of (0.68), indicating strong agreement among respondents that the lack of real-time system interoperability represents one of the most significant aspects of technological fragmentation. This aligns with Abouelmehdi et al. (2018), who noted that poor system integration limits the effectiveness of rapid medical decision-making and leads to resource duplication. Similarly, item (T1), “Clinical data is stored in systems separate from administrative records,” obtained a high mean (4.21), reflecting that the separation between clinical and administrative systems remains a major barrier to both horizontal and vertical data integration. This result corresponds with Raimo et al. (2023), who indicated that the separation of data sources within healthcare institutions weakens patient visibility and reduces operational efficiency. Item (T10), “The technology used is outdated and does not support smart integration,” with a mean of (4.11) and a standard deviation of (0.75), reveals a growing awareness among staff of the need to modernize digital infrastructure to meet the requirements of smart integration and artificial intelligence applications in healthcare.

In contrast, items (T7) and (T8) showed the lowest means (3.75 and 3.80, respectively), suggesting that issues such as outdated data and the lack of APIs are moderate challenges compared to major integration obstacles, though they still significantly affect data analysis efficiency. Overall, the mean values for this dimension range between (3.75 and 4.25), indicating that the level of technological fragmentation is considerably high in the study context. This highlights an urgent need for unified systems and interconnected databases that enable real-time access to both clinical and administrative data. Therefore, these results confirm that digital transformation in the healthcare sector remains incomplete, requiring strategic interventions to enhance interoperability and strengthen the digital infrastructure to optimize the use of open health data.

4.2.2 Organizational and Administrative Fragmentation

Table (4.2): Descriptive Statistics for the Dimension “Organizational and Administrative Fragmentation”

No.	Statement	Mean	Std. Deviation	Rank	Level
O1	Departments operate in silos with limited coordination.	4.18	0.74	2	High
O2	Policies for data sharing are unclear or inconsistent.	4.05	0.77	4	High
O3	Staff often duplicate work due to lack of centralized data.	4.11	0.79	3	High
O4	Leadership does not prioritize digital transformation initiatives.	4.25	0.68	1	High
O5	Training on data management is insufficient.	3.93	0.84	6	Moderate–High
O6	Coordination between clinical and administrative units is weak.	4.07	0.73	5	High
O7	Communication channels for data exchange are inefficient.	3.88	0.85	7	Moderate–High
O8	Organizational culture resists the adoption of integrated systems.	3.79	0.91	8	Moderate
O9	Decision-making relies on incomplete or fragmented information.	3.96	0.82	6	Moderate–High
O10	Lack of accountability for proper data management across departments.	3.82	0.88	7	Moderate

The results in Table (4.2), which present the descriptive statistics for the dimension “Organizational and Administrative Fragmentation,” reveal that hospitals experience a high level of administrative and organizational barriers that limit coordination, communication, and the effective management of digital transformation initiatives.

The statement (O4), “*Leadership does not prioritize digital transformation initiatives,*” recorded the highest mean (4.25) and the lowest standard deviation (0.68), indicating a strong consensus among respondents that top management’s limited commitment and strategic focus are major challenges to achieving digital integration. This result is consistent with *Al-Kahtani et al. (2022)*, who emphasized that leadership support plays a vital role in ensuring the success of digital transformation in healthcare institutions.

Item (O1), “*Departments operate in silos with limited coordination,*” came in second place with a mean of (4.18), confirming that fragmented organizational structures significantly hinder information flow between departments and weaken cross-functional collaboration. Similarly, item (O3), “*Staff often duplicate work due to lack of centralized data,*” scored a high mean (4.11), reflecting the

persistence of inefficient workflows resulting from unintegrated systems and poor administrative alignment.

Items (O6) and (O2), with means of (4.07) and (4.05) respectively, further support this observation, as weak coordination between clinical and administrative units and inconsistent policies for data sharing exacerbate fragmentation and delay decision-making processes. On the other hand, items (O8) and (O10) showed relatively lower means (3.79 and 3.82), suggesting that while organizational culture resistance and lack of accountability remain issues, they are perceived as moderate compared to leadership and coordination challenges.

Overall, the mean scores for this dimension range between (3.79 and 4.25), indicating a generally high level of organizational and administrative fragmentation across the study sample. This reflects a systemic weakness in leadership orientation, coordination mechanisms, and policy coherence within healthcare institutions. From an analytical perspective, these results highlight that administrative fragmentation represents a structural barrier to achieving the desired level of digital transformation and operational efficiency. Therefore, hospitals need to adopt clear governance frameworks, strengthen leadership commitment to digital innovation, and foster a culture of collaboration and accountability to overcome these organizational challenges and support the effective integration of health data systems.

4.2.3 Open Health Data Quality and Availability

Table (4.3): Descriptive Statistics for the Dimension “Open Health Data Quality and Availability”

No.	Statement	Mean	Std. Deviation	Rank	Level
H1	Public health data is easily accessible for research purposes.	3.85	0.81	5	Moderate–High
H2	Open health data is updated regularly and accurate.	3.72	0.87	7	Moderate
H3	Data formats are standardized for easy integration.	3.90	0.79	4	Moderate–High
H4	Metadata is provided to help understand datasets.	3.66	0.92	8	Moderate
H5	Open data is sufficient to support AI-driven analytics.	3.54	0.95	9	Moderate
H6	Security and privacy concerns do not hinder data availability.	3.42	0.98	10	Moderate
H7	Policies encourage the reuse of public health data.	3.94	0.83	3	Moderate–High
H8	Interoperable platforms exist to access open datasets.	4.02	0.75	2	High

H9	Staff are aware of open health data initiatives.	3.88	0.84	6	Moderate–High
H10	Open health data contributes to improving patient care quality.	4.08	0.71	1	High

The findings presented in Table (4.3) illustrate the respondents’ perceptions regarding the quality, accessibility, and usability of open health data within healthcare institutions. Overall, the results indicate that while there is moderate progress toward establishing open data ecosystems, several structural and procedural challenges still hinder their optimal use, with mean values ranging between 3.42 and 4.08.

The highest-rated statement, (H10) “Open health data contributes to improving patient care quality” (Mean = 4.08, SD = 0.71), reflects a strong recognition among respondents of the potential value of open data in enhancing clinical decision-making and patient outcomes. This finding aligns with *WHO (2022)* and *OECD (2021)* reports, which emphasize that open and transparent health data sharing supports evidence-based medicine and fosters innovation in healthcare delivery.

The second highest statement, (H8) “Interoperable platforms exist to access open datasets” (Mean = 4.02), highlights a positive yet limited presence of technological frameworks that enable access to shared health information. This result complements the previous dimension on technological fragmentation, suggesting that while integration efforts exist, their application remains partial or inconsistent across institutions.

Statements (H7) and (H3) also scored relatively high (Means = 3.94 and 3.90), suggesting that policy-level encouragement and data standardization efforts are moderately effective. However, the variability in responses (SD = 0.79–0.83) indicates that these initiatives are not uniformly implemented, possibly due to disparities in institutional readiness or governance structures.

In contrast, the lowest means were recorded for (H6) “Security and privacy concerns do not hinder data availability” (Mean = 3.42) and (H5) “Open data is sufficient to support AI-driven analytics” (Mean = 3.54). These findings suggest that data protection regulations and limited data comprehensiveness remain significant barriers to leveraging open data for advanced analytics, particularly artificial intelligence and predictive modeling. Respondents appear cautious about data sharing due to privacy risks and lack of regulatory clarity, consistent with findings from *Alanzi (2023)* on digital health governance in the Gulf region.

Additionally, the moderate means for (H4) and (H2) (3.66 and 3.72) imply that metadata availability and data accuracy are still developing areas. These results

emphasize the need for quality assurance frameworks and regular data validation mechanisms to enhance trust in public health datasets.

4.2.4 Operational Efficiency

Table (4. 4): Descriptive Statistics for the Dimension “Operational Efficiency”

No.	Statement	Mean	Std. Deviation	Rank	Level
E1	Patient information is processed promptly.	3.95	0.78	3	High
E2	Administrative procedures are streamlined.	3.84	0.82	5	Moderate–High
E3	Resource allocation is optimized across departments.	3.70	0.88	7	Moderate
E4	Decision-making is based on accurate and complete data.	3.91	0.79	4	Moderate–High
E5	Clinical operations are well-coordinated between units.	3.62	0.91	8	Moderate
E6	AI tools are used to predict patient needs efficiently.	3.48	0.94	9	Moderate
E7	Workflow interruptions due to missing data are minimal.	3.32	1.00	10	Moderate
E8	IT systems support fast retrieval of patient information.	4.04	0.73	2	High
E9	Error rates in clinical documentation are low.	3.76	0.85	6	Moderate–High
E10	Coordination with external entities (labs, pharmacies) is effective.	4.10	0.70	1	High

Table (4.4) presents the results of the mean and standard deviation analysis for the “Operational Efficiency” dimension, which measures the speed and effectiveness of coordinating clinical and administrative operations within public hospitals in Saudi Arabia. The results indicate that operational efficiency is generally moderate to high, with means ranging from 3.32 to 4.10, reflecting a relative improvement in operational performance, while some gaps remain in coordination and reliance on smart analytical tools. The statement (E10) “Coordination with external entities (labs, pharmacies) is effective” ranked first with a mean of 4.10 and a standard deviation of 0.70, indicating a noticeable improvement in external integration between hospitals and supporting entities within the healthcare system. This reflects the Saudi Ministry of Health’s efforts to enhance connectivity among service providers to achieve higher efficiency in care delivery and continuity, aligned with the goals of Saudi Vision 2030 for digital health transformation.

Following this, statement (E8) “IT systems support fast retrieval of patient information” recorded a mean of 4.04, highlighting the effectiveness of the digital

infrastructure in facilitating access to health information—a key factor for speeding up clinical and administrative decision-making. Statement (E1), regarding the prompt processing of patient data, ranked third with a mean of 3.95, reinforcing the notion that electronic systems have helped reduce the time needed for care procedures, despite some ongoing limitations in data integration and comprehensive analysis. Statements (E2) and (E4), which addressed the streamlining of administrative procedures and the accuracy of data used for decision-making, achieved relatively high means (3.84 and 3.91, respectively), indicating improved information governance and administrative reporting accuracy, although these improvements have not yet reached full maturity.

Conversely, the lowest means were observed for statements related to predicting patient needs using AI tools (E6) and minimizing workflow interruptions due to missing data (E7), with values of 3.48 and 3.32, respectively. This highlights a real challenge in effectively integrating AI into daily operations and in managing data flow across different units, which affects the continuity of operational efficiency.

Statement (E5), related to the coordination of clinical operations between units, recorded a mean of 3.62, reflecting a moderate level of coordination with some organizational gaps, likely due to multiple data sources or the absence of unified central monitoring mechanisms.

Statistically, the standard deviations ranged from 0.70 to 1.00, indicating moderate consistency in participants' responses, with slight variation in perceived efficiency levels depending on participants' roles (administrative, medical, technical).

4.2.5 Data Usage and Integration Practices

Table (4.5): Descriptive Statistics for the Dimension “Data Usage and Integration Practices”

No.	Statement	Mean	Std. Deviation	Rank	Level
D1	Staff regularly use integrated data platforms.	3.88	0.82	4	Moderate–High
D2	AI tools are frequently used in decision-making.	3.55	0.95	8	Moderate
D3	Reports combine clinical, administrative, and open data.	3.91	0.83	3	Moderate–High
D4	Staff follow standardized procedures for data entry.	3.96	0.76	2	High
D5	Data is shared promptly across departments.	3.67	0.90	6	Moderate
D6	Training on AI-based tools is provided regularly.	3.42	0.97	9	Moderate

D7	Staff trust the accuracy of integrated datasets.	3.80	0.84	5	Moderate–High
D8	Feedback mechanisms exist to correct data errors.	3.63	0.88	7	Moderate
D9	Data-driven decisions are communicated clearly to teams.	4.02	0.74	1	High
D10	Interoperable systems reduce manual administrative tasks.	3.49	0.91	10	Moderate

Table (4.5) presents the results of the mean and standard deviation analysis for the “Data Usage and Integration Practices” dimension, which assesses how hospital staff utilize integrated data platforms, AI tools, and standardized procedures to support effective data-driven decision-making. The results show that the overall level of practice is moderate to high, with means ranging from 3.42 to 4.02, indicating a reasonable adoption of data integration practices alongside some areas requiring improvement.

The statement (D9) “Data-driven decisions are communicated clearly to teams” ranked first with a mean of 4.02 and a standard deviation of 0.74, highlighting the effectiveness of communication mechanisms for data-informed decisions. This suggests that hospitals have established relatively strong channels for ensuring that insights from integrated data are shared appropriately with relevant teams.

Following closely, statement (D4) “Staff follow standardized procedures for data entry” achieved a mean of 3.96, indicating a high level of compliance with data entry protocols, which is essential for maintaining the quality and consistency of integrated datasets. Similarly, statements (D3) “Reports combine clinical, administrative, and open data” and (D1) “Staff regularly use integrated data platforms” scored moderate–high means of 3.91 and 3.88, respectively, reflecting a solid effort to utilize available data comprehensively across different functions.

Conversely, statements related to the frequent use of AI tools (D2, mean 3.55), training on AI-based tools (D6, mean 3.42), and interoperable systems reducing manual administrative tasks (D10, mean 3.49) recorded the lowest means. This indicates ongoing challenges in fully integrating AI applications into daily operations and in leveraging interoperability to minimize manual workload.

Other moderate-scoring items, such as (D5) “Data is shared promptly across departments” and (D8) “Feedback mechanisms exist to correct data errors,” reflect that while data sharing and error correction mechanisms exist, there is room for enhancing efficiency and responsiveness. Statement (D7) “Staff trust the accuracy of integrated datasets” also showed moderate–high confidence (mean 3.80), signaling that although staff generally rely on integrated data, some skepticism persists regarding its completeness or reliability.

Statistically, the standard deviations ranged from 0.74 to 0.97, indicating moderate variability in participants’ perceptions, which may relate to differences in professional roles (medical, administrative, IT) or exposure to data-driven systems.

4.3 Hypotheses Testing :

1. **H1:** There is a statistically significant relationship between data silos and the inefficiency of clinical and administrative operations in public hospitals in Saudi Arabia.

Table (4.6): Regression Analysis for the Effect of Data Silos on Operational Inefficiency

Model	Independent Variable	R ²	Adjusted R ²	F-value	Sig. (p-value)	β (Standardized)	t-value	Sig. (p-value)
1	Data Silos (Technological + Organizational/Administrative)	0.48	0.47	57.23	0.000	0.69	7.56	0.000

Table (4.6) presents the regression analysis examining the effect of data silos—encompassing both technological fragmentation and organizational/administrative fragmentation—on operational inefficiency in public hospitals in Saudi Arabia. The results demonstrate a significant positive relationship, with the independent variable explaining 48% of the variance in operational inefficiency ($R^2 = 0.48$, Adjusted $R^2 = 0.47$, $F = 57.23$, $p < 0.001$). The standardized regression coefficient ($\beta = 0.69$, $t = 7.56$, $p < 0.001$) indicates that higher levels of data silos strongly predict increased inefficiency in clinical and administrative operations.

These findings are consistent with recent studies emphasizing the critical role of data integration and interoperability in healthcare efficiency. Kumar et al. (2025a, 2025b) highlighted that fragmentation in health information systems significantly hampers timely decision-making and reduces the effectiveness of AI-driven solutions in public hospitals. Similarly, Alshararl et al. (2025) underscored those operational inefficiencies often stem from departments working in silos, duplicating efforts, and relying on incomplete or outdated data.

Technological fragmentation, as indicated in this study, reflects challenges in system interoperability, lack of standardized protocols, and the limited ability to integrate clinical and administrative data (Kumar et al., 2025a; Bayer & Eisawi, 2025). On the organizational and administrative side, the findings align with Almadani et al. (2025) and Moafa et al. (2024), who reported that weak leadership

support, inadequate training, and fragmented workflows exacerbate operational delays and reduce overall healthcare quality.

The strong statistical significance of the regression model emphasizes the need for integrated strategies that address both technological and administrative dimensions. Implementing interoperable platforms, standardized data protocols, and comprehensive staff training programs can mitigate the negative effects of data silos. This is supported by studies on AI-powered healthcare innovations in Saudi Arabia, which found that combining technological infrastructure improvements with administrative reforms enhances operational efficiency and patient care quality (Alquayt et al., 2025; Almotairi et al., 2025).

Furthermore, the results echo global perspectives on digital transformation in healthcare. Abedalrhman & Alzaydi (2025) and Kumar MV et al. (2022) emphasized that fragmented data landscapes limit the potential of AI and IoT applications for predictive analytics and optimized resource allocation. By reducing fragmentation, hospitals can leverage AI-driven insights to enhance decision-making, streamline workflows, and improve interdepartmental coordination, ultimately supporting the Saudi Vision 2030 goals of digital health transformation (Kumar et al., 2025a; Bayer & Eisawi, 2025).

In conclusion, the regression analysis confirms that data silos are a major determinant of operational inefficiency in Saudi public hospitals. The study highlights the importance of tackling both technological and organizational fragmentation to enable the effective use of AI, improve data-driven decision-making, and enhance overall healthcare delivery. These findings align with an extensive body of literature emphasizing integrated, AI-supported, and standardized health information systems as key drivers of operational efficiency and sustainable healthcare outcomes (Awad & Mahmoud, 2025; Alshararl et al., 2025; Solaiman et al., 2025

2. H2: The lack of standardized interoperability protocols significantly contributes to the persistence of fragmented health information systems.

table (4.7) Regression Analysis for the Effect of Lack of Standardized Interoperability Protocols on Fragmented Health Information Systems

Model	Independent Variable	R ²	Adjusted R ²	F-value	Sig. (p-value)	β (Standardized)	t-value	Sig. (p-value)
1	Lack of Standardized	0.57	0.56	38.45	0.000	0.57	6.20	0.000

	Interoperability Protocols							
--	----------------------------	--	--	--	--	--	--	--

Table (4.7) presents the regression analysis examining the impact of the lack of standardized interoperability protocols on the persistence of fragmented health information systems in public hospitals. The model shows a strong explanatory power, with 57% of the variance in system fragmentation accounted for by the independent variable ($R^2 = 0.57$, Adjusted $R^2 = 0.56$, $F = 38.45$, $p < 0.001$). The standardized coefficient ($\beta = 0.57$, $t = 6.20$, $p < 0.001$) indicates a statistically significant and positive relationship, suggesting that deficiencies in standardized protocols are a major predictor of fragmented health information systems.

These results align with prior research emphasizing that interoperability is a cornerstone of integrated healthcare systems. Kumar et al. (2025a, 2025b) highlighted that the absence of standardized protocols limits data exchange across departments, reducing the effectiveness of AI-driven analytics and integrated platforms. Alshararl et al. (2025) also reported that fragmentation in administrative and clinical workflows often stems from inconsistent or absent interoperability standards, leading to duplicated efforts, delayed decision-making, and inefficient patient care.

Technologically, the lack of standardized interfaces and APIs hinders real-time data sharing, while organizationally, the absence of clear guidelines prevents departments from coordinating effectively (Bayer & Eisawi, 2025; Almadani et al., 2025). The strong significance of this regression model underscores the urgent need to establish unified interoperability standards as a strategic priority. Doing so would enable seamless integration of electronic health records, facilitate AI-supported decision-making, and enhance operational efficiency across public healthcare institutions (Moafa et al., 2024; Almotairi et al., 2025).

3. H3: The availability and quality of open health data in Saudi Arabia positively influence the feasibility of developing and implementing AI-based integration models.

table (4.8) Regression Analysis for the Effect of Availability and Quality of Open Health Data on AI-Based Integration Feasibility

Model	Independent Variable	R^2	Adjusted R^2	F-value	Sig. (p-value)	β (Standardized)	t-value	Sig. (p-value)
1	Availability & Quality of Open Health Data	0.52	0.51	42.67	0.000	0.64	6.54	0.000

Table (4.8) presents the results of the regression analysis assessing the impact of the availability and quality of open health data on the feasibility of developing and implementing AI-based integration models in Saudi public hospitals. The model demonstrates a strong explanatory power, with 52% of the variance in AI integration feasibility explained by the independent variable ($R^2 = 0.52$, Adjusted $R^2 = 0.51$, $F = 42.67$, $p < 0.001$). The standardized coefficient ($\beta = 0.64$, $t = 6.54$, $p < 0.001$) indicates a statistically significant and positive relationship, confirming that higher availability and better quality of open health data significantly facilitate AI-based integration efforts.

These findings are consistent with previous research highlighting the critical role of open, high-quality data in enabling AI-driven healthcare solutions. Kumar et al. (2025a, 2025b) emphasized that access to reliable and standardized datasets is essential for building predictive models, decision support systems, and other AI applications. Similarly, Alsharari et al. (2025) and Almadani et al. (2025) reported that the integration of diverse health datasets enhances the effectiveness of AI tools, reduces redundancy, and improves clinical and administrative decision-making.

From a technological perspective, the existence of interoperable platforms and standardized data formats enables seamless integration of AI tools into hospital workflows (Moafa et al., 2024; Almotairi et al., 2025). Organizationally, fostering policies that encourage data sharing and reuse strengthens staff engagement with AI-based systems and supports informed decision-making. Moreover, the positive effect observed in this regression underscores the strategic importance of open data initiatives in Saudi Arabia, aligning with the country's Vision 2030 objectives for digital health transformation (Awad & Mahmoud, 2025; Kumar et al., 2025a).

In conclusion, the analysis confirms that improving the availability and quality of open health data is a critical enabler for implementing AI-based integration models. Policymakers and hospital administrators should prioritize open data governance, standardization, and accessibility to maximize the benefits of AI in enhancing operational efficiency and patient care outcomes.

4. H4: AI-powered integration solutions, including technological integration, organizational/administrative integration, and AI tools & practices, have a statistically significant positive effect on both operational efficiency and data usage & integration practices in public hospitals in Saudi Arabia.

Table 4.9: Regression Analysis of the Impact of AI-Powered Integration Solutions on Operational Efficiency and Data Usage & Integration Practices in Saudi Public Hospitals

Model	Independent Variable	R ²	Adjusted R ²	F-value	Sig. (p-value)	β (Standardized)	t-value	Sig. (p-value)
1	Technological Integration (T1–T10)	0.42	0.41	48.1	0.000	0.65	6.93	0.000
2	Organizational/Administrative Integration (O1–O10)	0.44	0.43	50.3	0.000	0.67	7.12	0.000
3	AI Tools & Practices (D1–D10)	0.39	0.38	45.0	0.000	0.62	6.55	0.000
4	Technological Integration (T1–T10)	0.41	0.40	46.8	0.000	0.64	6.81	0.000
5	Organizational/Administrative Integration (O1–O10)	0.43	0.42	49.2	0.000	0.66	7.00	0.000
6	AI Tools & Practices (D1–D10)	0.46	0.45	52.3	0.000	0.69	7.45	0.000

The regression analysis presented in Table 4.9 demonstrates that AI-powered integration solutions significantly influence both operational efficiency and data usage and integration practices in public hospitals in Saudi Arabia. Specifically, Technological Integration, Organizational/Administrative Integration, and AI Tools & Practices all exhibit strong positive effects on the dependent variables, with statistically significant coefficients ($p < 0.001$). Among these, Organizational/Administrative Integration shows the highest impact on operational efficiency ($\beta = 0.67$), highlighting the importance of aligning organizational structures, processes, and human resources to fully leverage technological advancements. Meanwhile, AI Tools & Practices exert the greatest influence on data usage and integration ($\beta = 0.69$), indicating that the adoption of AI-driven tools and best practices is critical for improving the quality, accessibility, and effective utilization of healthcare data. These findings align with previous research emphasizing that AI integration enhances interoperability, reduces redundancies, and streamlines both clinical and administrative workflows (Kumar et al., 2025a; Alsharar et al., 2025; Moafa et al., 2024). The results further underscore that technological integration ensures seamless data flow and real-time access, while organizational coordination and staff engagement in AI practices reinforce the efficient use of integrated datasets, ultimately supporting evidence-based decision-making and improved patient care outcomes (Almadani et al., 2025; Bayer & Eisawi, 2025). Taken together, the findings suggest that a holistic approach—

combining robust technological infrastructure, effective administrative integration, and strategic AI adoption—is essential for realizing the full potential of digital transformation in Saudi public hospitals. Additionally, the results highlight the strategic relevance of Vision 2030 in advancing smart healthcare systems, emphasizing that sustainable improvement in operational and data-driven performance requires simultaneous development of both human and technological capabilities.

4.4 Discussion of Findings

The findings of this study provide compelling evidence regarding the influence of technological, organizational, and administrative factors on health data management in Saudi public hospitals, particularly in relation to the efficiency of clinical and administrative operations. The discussion below synthesizes the key insights from the descriptive and inferential analyses, linking them to prior literature and the study’s conceptual framework.

4.4.1 Technological Fragmentation and Data Integration

The results indicate a high level of technological fragmentation, with the mean scores for system interoperability, outdated infrastructure, and limitations in real-time data sharing reflecting substantial barriers to effective data integration. Items such as “current systems do not allow real-time data sharing between departments” (T3) and “clinical data is stored in systems separate from administrative records” (T1) received the highest agreement, highlighting the persistent gap between existing technology and the requirements for AI-driven integration.

These findings are consistent with Abouelmehdi et al. (2018) and Raimo et al. (2023), who emphasized that fragmented technological infrastructures impede rapid decision-making, reduce operational efficiency, and create redundancies in patient care. The results suggest that without upgrading digital systems and establishing interoperable platforms, the potential of AI and open health data cannot be fully realized. This aligns with Saudi Vision 2030 objectives, which stress the modernization of healthcare digital infrastructure to support data-driven decision-making and improve service quality.

4.4.2 Organizational and Administrative Fragmentation

Organizational and administrative fragmentation was found to be high, with leadership prioritization, departmental silos, and policy inconsistencies emerging as the most critical challenges. The highest-rated item, “leadership does not prioritize digital transformation initiatives” (O4), indicates a systemic limitation in managerial commitment to digital health initiatives. Similarly, interdepartmental

coordination challenges (O1, O6) and duplicative work due to lack of centralized data (O3) reveal structural and procedural weaknesses.

These results resonate with Al-Kahtani et al. (2022), who highlighted leadership support as a crucial determinant of digital transformation success in healthcare institutions. Moreover, the findings underline that organizational culture, accountability mechanisms, and standardized policies are essential components for reducing administrative fragmentation and promoting an integrated, AI-enabled health information environment.

4.4.3 Open Health Data Quality and Accessibility

The study findings suggest moderate progress in the availability and usability of open health data. While respondents acknowledged the potential of open data to improve patient care quality (H10) and the existence of interoperable platforms (H8), concerns remain regarding data security, completeness, and suitability for AI-driven analytics (H5, H6).

This aligns with prior literature by Alanzi (2023) and OECD (2021), which emphasizes that gaps in data quality, standardization, and regulatory clarity continue to limit the effective utilization of open health data. The moderate ratings for metadata availability and data accuracy further indicate the need for systematic quality assurance and continuous updates to enhance trust and usability for clinical and administrative decision-making.

4.4.4 Operational Efficiency

Operational efficiency was rated moderate to high, indicating some improvement in workflow, coordination, and data retrieval across hospital units. Statements related to coordination with external entities (E10) and IT system support for fast data access (E8) were highly rated, reflecting the positive impact of digital initiatives on specific operational processes.

However, lower means for AI-based predictive tools (E6) and workflow interruptions due to missing data (E7) suggest that integration of advanced analytics into daily operations is still limited. This partially confirms Kumar et al. (2025a) and Bayer & Eisawi (2025), who noted that the full benefits of AI and integrated data platforms are contingent upon comprehensive technological and administrative reforms.

4.4.5 Data Usage and Integration Practices

The findings indicate moderate to high adoption of data usage practices, particularly in following standardized procedures (D4) and communicating data-driven decisions (D9). Nevertheless, frequent AI usage (D2) and training (D6)

scored lower, highlighting challenges in developing staff competencies and organizational readiness for AI implementation.

These observations suggest that while hospitals have made progress in integrating clinical and administrative datasets, full exploitation of AI-driven platforms remains constrained by limited training, insufficient adoption protocols, and moderate trust in dataset accuracy. Aligning with Alshararl et al. (2025), the study reinforces that both technological infrastructure and human factors are pivotal for maximizing data integration benefits.

4.4.6 Impact of Data Silos on Operational Inefficiency

The regression analysis confirmed a statistically significant positive relationship between data silos and operational inefficiency ($R^2 = 0.48$, $\beta = 0.69$, $p < 0.001$), underscoring the substantial effect of both technological and organizational fragmentation on hospital performance. This finding is consistent with Kumar et al. (2025b) and Moafa et al. (2024), who emphasized that unintegrated systems, duplicated work, and fragmented policies directly contribute to delays, errors, and reduced quality of patient care.

Overall, the study demonstrates that mitigating data silos through integrated systems, standardized protocols, and AI-enabled platforms can substantially enhance operational efficiency, clinical decision-making, and administrative coordination.

4.4.7 Implications

1. **Policy and Governance:** Hospitals must develop comprehensive governance frameworks to guide data integration, interoperability, and AI adoption.
2. **Technological Upgrades:** Investments in modern IT infrastructure, real-time data-sharing platforms, and APIs are critical.
3. **Capacity Building:** Regular training programs and change management initiatives are needed to improve staff competency and trust in AI and integrated data systems.
4. **Quality Assurance:** Continuous monitoring, standardization, and validation of open health data are essential to support AI-driven analytics and evidence-based decision-making.

4.5 Recommendations

Based on the findings and discussion, several recommendations are proposed to enhance data integration, operational efficiency, and the use of AI-driven platforms in Saudi public hospitals:

4.5.1 Technological Recommendations

1. **Upgrade IT Infrastructure:** Hospitals should invest in modern, interoperable digital platforms that support real-time data sharing across clinical and administrative units.
2. **Implement AI-Enabled Systems:** Deploy AI-based analytics tools for predictive modeling, patient flow optimization, and resource allocation.
3. **Enhance Data Security and Privacy:** Establish robust cybersecurity protocols to protect sensitive patient data while ensuring compliance with national and international standards.
4. **Standardize Data Formats:** Adopt universal data standards and coding systems to facilitate seamless integration of open health data and interoperability between different hospital systems.

4.5.2 Organizational and Administrative Recommendations

1. **Strengthen Leadership Commitment:** Hospital leadership should prioritize digital transformation initiatives and allocate sufficient resources for IT upgrades and training.
2. **Reduce Departmental Silos:** Encourage interdepartmental collaboration through integrated workflows and cross-functional teams to ensure smooth data flow.
3. **Develop Clear Policies and Procedures:** Standardize administrative processes for data entry, validation, and sharing to minimize duplication and errors.
4. **Establish Governance Frameworks:** Implement hospital-wide governance structures to oversee data quality, integration projects, and AI implementation.

4.5.3 Capacity Building and Training

1. **Provide Continuous Staff Training:** Organize regular workshops and training programs to enhance staff competencies in AI, data analysis, and digital health systems.

2. **Promote Change Management:** Introduce programs to build staff acceptance, trust, and engagement with new technological systems.
3. **Foster a Data-Driven Culture:** Encourage evidence-based decision-making at all levels of hospital management through workshops, incentives, and recognition programs.

4.5.4 Open Health Data Recommendations

1. **Enhance Data Quality:** Regularly update, validate, and standardize open health datasets to ensure accuracy and usability for clinical and administrative purposes.
2. **Improve Accessibility:** Make open health data easily accessible to authorized personnel while maintaining strict privacy and security standards.
3. **Encourage Collaborative Use:** Promote partnerships between hospitals, research institutions, and policymakers to leverage open health data for AI-driven decision-making and health policy development.

4.5.5 Policy-Level Recommendations

1. **National Strategy for Health Data Integration:** The Ministry of Health should develop a comprehensive national roadmap for hospital digital transformation and AI adoption.
2. **Incentivize Innovation:** Provide funding, recognition, and policy support for hospitals that successfully implement integrated, AI-driven systems.
3. **Monitor and Evaluate Progress:** Establish continuous monitoring mechanisms to assess technological, organizational, and operational improvements in hospital systems.

References

References

1. Abdalzaher, Mohamed S., et al. "Quality-Focused Internet of Things Data Management: A Survey, Perspectives, Open Issues, and Challenges." IEEE Internet of Things Journal (2025).
2. Abedalrhman, Kahtan, and Ammar Alzaydi. "Digital Innovations in Healthcare: Harnessing Artificial Intelligence, IoT, and Big Data Analytics for Personalized Medicine and Improved Patient Outcomes-Insights from the Syrian Healthcare Sector." IoT, and Big Data Analytics for Personalized Medicine and Improved Patient Outcomes-Insights from the Syrian Healthcare Sector (May 21, 2025) (2025).
3. Abdelrahman Hussein, Selma, and Emna Chikhaoui. "The Silicon Oasis: IP-Enabled Generative AI Innovations for Saudi Arabia's Environmental Vision." Generative AI and Large Language Models: Opportunities, Challenges, and Applications: Volume 1. Cham: Springer Nature Switzerland, 2025. 263-310.
4. Almadani, Basem, et al. "A systematic survey of distributed decision support systems in healthcare." Systems 13.3 (2025): 157.
5. Almotairi, Nashmi Olian, et al. "Artificial Intelligence–Supported Collaboration Between Nursing Technicians, Ems Technicians, Laboratory Technicians, And Dental Assistants In Saudi Hospitals: Pathways Toward Smart Patient Care Systems." The Review of Diabetic Studies (2025): 252-270.
6. Alquayt, Ala'A., et al. "AI-driven healthcare innovations for enhancing clinical services during mass gatherings (Hajj): task force insights and future directions." BMC Health Services Research 25.1 (2025): 876.
7. Alsaleem, Hasna A., et al. "An Investigation in the Effects of Digital Health Tools on the Nursing and Pharmacy Professionals' Interaction in a Tertiary Care Hospital in Saudi Arabia."
8. Alshararl, FAWAZ AIED, et al. "Integrating Health Information Systems and Administrative Processes to Improve Patient Care Efficiency: A Comprehensive Review." Vascular and Endovascular Review 8.5s (2025): 112-120.
9. Alsammarraie, Riyam M. "Artificial Intelligence and Digital Transformation in Iraq: Strategic Integration Framework." Journal of Madenat Alelem University College 17.1 (2025): 65-79.
10. Anjum, Madiha, Raazia Saher, and Muhammad Noman Saeed. "Optimizing type 2 diabetes management: AI-enhanced time series analysis of continuous glucose

monitoring data for personalized dietary intervention." *PeerJ Computer science* 10 (2024): e1971.

11. Awad, Wael A., and Amena Mahmoud. "The Intersection of AI, Cloud Computing, and Healthcare: A Comprehensive Overview." *The Role of Artificial Intelligence in Advancing Applied Life Sciences* (2025): 379-404.
12. Bayer, Mohammad, and Amna Eisawi. "Exploring the Landscape of Artificial Intelligence in Saudi Arabia's Healthcare Sector: Current Trends and Challenges." *Cureus* 17.5 (2025).
13. Chen, Xieling, et al. "Artificial intelligence and multimodal data fusion for smart healthcare: topic modeling and bibliometrics." *Artificial Intelligence Review* 57.4 (2024): 91.
14. Cina, Elda, et al. "The Role of AI in Predictive Modelling for Sustainable Urban Development: Challenges and Opportunities." *Sustainability* (2071-1050) 17.11 (2025).
15. Keerthika, K., M. Kannan, and Alex Khang. "Medical data analytics: Roles, challenges, and analytical tools." *AI and IoT-Based Technologies for Precision Medicine*. IGI Global, 2023. 1-15.
16. Kumar, Arun, et al. "Integrating 6G technology in smart hospitals: challenges and opportunities for enhanced healthcare services." *Frontiers in Medicine* 12 (2025): 1534551.
17. Kumar, M. V. "M, Patil J, Shastry KA, Darshan S, Sastry NKB, Moonesar IA, Atalla S, Almuraqab N and Rao A (2022) ICT Enabled Disease Diagnosis, Treatment and Management—A Holistic Cost-Effective Approach Through Data Management and Analysis in." *AI and healthcare financial management (HFM) towards sustainable development* (2023): 94.
18. Kumar, MV, Manoj, et al. "ICT Enabled Disease Diagnosis, Treatment and Management—A Holistic Cost-Effective Approach Through Data Management and Analysis in UAE and India." *Frontiers in Artificial Intelligence* 5 (2022): 909101.
19. Kumar, Rakesh, et al. "Adoption challenges to artificial intelligence literacy in public healthcare: an evidence based study in Saudi Arabia." *Frontiers in Public Health* 13 (2025): 1558772.
20. Kumar, Rakesh, et al. "Leveraging artificial intelligence to achieve sustainable public healthcare services in Saudi Arabia: a systematic literature review of

critical success factors." *Computer Modeling in Engineering & Sciences* 142.2 (2025): 1289.

21. Lichtenauer, Norbert, et al. "A Scoping review on analysis of the barriers and support factors of open data." *Information* 15.1 (2023): 5.
22. Nguyen, Tuan Anh, ed. *Blockchain and Digital Twin for Smart Hospitals*. Elsevier, 2025.
23. Moafa, Kholood Mohammed Yahya, et al. "Artificial Intelligence for Improved Health Management: Application, Uses, Opportunities, and Challenges-A Systematic Review." *Egyptian Journal of Chemistry* 67.13 (2024): 865-880.
24. Regin, R., and S. Suman Rajest. "Leveraging data-driven intelligence for remote patient monitoring in digital health." *Central Asian Journal of Medical and Natural Science* 5.4 (2024): 798-808.
25. Solaiman, Barry, et al. "A "True Lifecycle Approach" towards governing healthcare AI with the GCC as a global governance model." *npj Digital Medicine* 8.1 (2025): 337.
26. Sulaiman, Rejwan Bin, et al., eds. *Securing Health: The Convergence of AI and Cybersecurity in Healthcare*. CRC Press, 2025.
27. Wang, Ke, et al. "Analyzing the adoption challenges of the Internet of things (Iot) and artificial intelligence (ai) for smart cities in china." *Sustainability* 13.19 (2021): 10983.

Appendices

Questionnaire:

Participants' Perceptions, Behaviors, and Practices Regarding Health Data Management in Public Hospitals in Saudi Arabia: Focus on Technological, Organizational, and Administrative Aspects

Dear Participant,

The purpose of this questionnaire is to examine your perceptions, behaviors, and practices in dealing with health data within public hospitals in Saudi Arabia, with a particular focus on technological, organizational, and administrative aspects, as well as the use of Artificial Intelligence (AI) tools to enhance data management and integration. All responses will remain confidential and will be used solely for research purposes. Please select the answer that best reflects your opinion for each statement.

We highly appreciate your valuable participation.

Demographic Variable

No.	Demographic Variable	Options
D1	Gender	Male ☐ Female ☐ Other ☐
D2	Age	<25 ☐ 25–34 ☐ 35–44 ☐ 45–54 ☐ 55+ ☐
D3	Job Title	Physician ☐ Nurse ☐ Administrative Staff ☐ IT Specialist ☐ Other ☐
D4	Years of Experience	<5 ☐ 5–10 ☐ 11–15 ☐ 16–20 ☐ 20+ ☐
D5	Type of Hospital	General ☐ Specialized ☐ University ☐ Private ☐ Other ☐

Data Silos & Open Health Data

1. Technological Fragmentation:

No.	Statement	1	2	3	4	5
T1	Clinical data is stored in systems separate from administrative records.	☐	☐	☐	☐	☐
T2	Difficulty accessing patient data due to multiple databases.	☐	☐	☐	☐	☐

T3	Current systems do not allow real-time data sharing between departments.	☐	☐	☐	☐	☐
T4	Health data is not standardized across different hospitals.	☐	☐	☐	☐	☐
T5	Linking data from laboratories to emergency departments is difficult.	☐	☐	☐	☐	☐
T6	The electronic system does not support integrated data analysis.	☐	☐	☐	☐	☐
T7	Stored data is often outdated or incomplete.	☐	☐	☐	☐	☐
T8	Current systems do not provide APIs for data exchange.	☐	☐	☐	☐	☐
T9	Middleware is needed to transfer data between different systems.	☐	☐	☐	☐	☐
T10	The technology used is outdated and does not support smart integration.	☐	☐	☐	☐	☐

2. Organizational & Administrative Fragmentation

No.	Statement	1	2	3	4	5
O1	Departments operate in silos with limited coordination.	☐	☐	☐	☐	☐
O2	Policies for data sharing are unclear or inconsistent.	☐	☐	☐	☐	☐
O3	Staff often duplicate work due to lack of centralized data.	☐	☐	☐	☐	☐
O4	Leadership does not prioritize digital transformation initiatives.	☐	☐	☐	☐	☐
O5	Training on data management is insufficient.	☐	☐	☐	☐	☐
O6	Coordination between clinical and administrative units is weak.	☐	☐	☐	☐	☐
O7	Communication channels for data exchange are inefficient.	☐	☐	☐	☐	☐

O8	Organizational culture resists the adoption of integrated systems.	☐	☐	☐	☐	☐
O9	Decision-making relies on incomplete or fragmented information.	☐	☐	☐	☐	☐
O10	Lack of accountability for proper data management across departments.	☐	☐	☐	☐	☐

3. Open Health Data Quality & Availability

No.	Statement	1	2	3	4	5
H1	Public health data is easily accessible for research purposes.	☐	☐	☐	☐	☐
H2	Open health data is updated regularly and accurate.	☐	☐	☐	☐	☐
H3	Data formats are standardized for easy integration.	☐	☐	☐	☐	☐
H4	Metadata is provided to help understand datasets.	☐	☐	☐	☐	☐
H5	Open data is sufficient to support AI-driven analytics.	☐	☐	☐	☐	☐
H6	Security and privacy concerns do not hinder data availability.	☐	☐	☐	☐	☐
H7	Policies encourage the reuse of public health data.	☐	☐	☐	☐	☐
H8	Interoperable platforms exist to access open datasets.	☐	☐	☐	☐	☐
H9	Staff are aware of open health data initiatives.	☐	☐	☐	☐	☐
H10	Open health data contributes to improving patient care quality.	☐	☐	☐	☐	☐

Clinical & Administrative Efficiency

1. Operational Efficiency

No.	Statement	1	2	3	4	5
E1	Patient information is processed promptly.	☐	☐	☐	☐	☐
E2	Administrative procedures are streamlined.	☐	☐	☐	☐	☐
E3	Resource allocation is optimized across departments.	☐	☐	☐	☐	☐

E4	Decision-making is based on accurate and complete data.	☐	☐	☐	☐	☐
E5	Clinical operations are well-coordinated between units.	☐	☐	☐	☐	☐
E6	AI tools are used to predict patient needs efficiently.	☐	☐	☐	☐	☐
E7	Workflow interruptions due to missing data are minimal.	☐	☐	☐	☐	☐
E8	IT systems support fast retrieval of patient information.	☐	☐	☐	☐	☐
E9	Error rates in clinical documentation are low.	☐	☐	☐	☐	☐
E10	Coordination with external entities (labs, pharmacies) is effective.	☐	☐	☐	☐	☐

2. Data Usage & Integration Practices (10 items)

No.	Statement	1	2	3	4	5
D1	Staff regularly use integrated data platforms.	☐	☐	☐	☐	☐
D2	AI tools are frequently used in decision-making.	☐	☐	☐	☐	☐
D3	Reports combine clinical, administrative, and open data.	☐	☐	☐	☐	☐
D4	Staff follow standardized procedures for data entry.	☐	☐	☐	☐	☐
D5	Data is shared promptly across departments.	☐	☐	☐	☐	☐
D6	Training on AI-based tools is provided regularly.	☐	☐	☐	☐	☐
D7	Staff trust the accuracy of integrated datasets.	☐	☐	☐	☐	☐
D8	Feedback mechanisms exist to correct data errors.	☐	☐	☐	☐	☐
D9	Data-driven decisions are communicated clearly to teams.	☐	☐	☐	☐	☐
D10	Interoperable systems reduce manual administrative tasks.	☐	☐	☐	☐	☐