Open Access Research Journal of
Science and Technology
Journals home page: https://oarjst.com/
ISSN: 2782-9960 (Online)

OARJ OPEN ACCESS RESEARCH JOURNALS

(Review Article)

Check for updates

The role of applied AI and FHIR-based Apis in building interoperable healthcare data ecosystems

Arvind Telharkar *

Engineering Lead at Amazon.

Open Access Research Journal of Science and Technology, 2026, 17(01), 033-043

Publication history: Received on 28 April 2026; revised on 05 June 2026; accepted on 08 June 2026

Article DOI: https://doi.org/10.53022/oarjst.2026.17.1.0059

Abstract

Healthcare generates a ton of clinical data, but disjointed info systems, inconsistent standards, and limited interoperability still hold things back. This study takes a look at how AI and FHIR-based APIs are coming together to form an emerging framework for sharing health data more effectively. It uses a mixed-methods approach to check out health system structures, provider survey replies, AI tools linked with FHIR, and implementation stories from different medical settings. The findings show that FHIR APIs help make the data consistent and easy to work with for AI, which is great for getting those systems up and running across health orgs. Places using AI in a FHIR-friendly way saw better data quality, improved clinical support, smoother reconciliations, and more efficient workflows. Still, big challenges pop up too - think complex data governance, varying levels of org preparedness, basic infra issues, gaps in compatibility, biases in algorithms, and unclear regulations. Trust from clinicians and high data quality are also super important factors in making these FHIR-AI setups work well. The article suggests a combined approach for safely using AI. It points out main areas for further study: federated learning, fair AI use in health, tracking results over time, and flexible rules. These findings fit in with other proofs that AI and FHIR work great together to boost data-driven health care systems.

Keywords: Artificial Intelligence; FHIR; Healthcare Interoperability; Electronic Health Records; Clinical Decision Support; Health Information Exchange; Data Governance; Federated Learning

1. Introduction

The promise of a fully connected healthcare system, one in which a patient's complete clinical history is available at the point of care regardless of where prior services were received, has driven decades of health information technology investment. Despite this sustained effort, the persistent reality for most healthcare providers remains one of siloed electronic health records (EHR), proprietary data formats, manual reconciliation of patient information across care settings, and missed opportunities for preventive intervention because relevant data simply was not accessible when needed. The cost of this fragmentation is not merely administrative. It manifests in diagnostic errors, duplicated testing, adverse drug events, delayed referrals, and preventable hospital readmissions, all of which represent both human and financial losses at enormous scale.

Two developments have created a credible pathway beyond this historical impasse. The first is the maturation of the FHIR specification, published and maintained by Health Level Seven International (HL7), which provides a RESTful, resource-based framework for representing and exchanging clinical data in a manner that is both human-readable and machine-processable. The second is the emergence of applied artificial intelligence, particularly in the form of large language models, transformer-based architectures, and federated machine learning systems, which have demonstrated genuine capability to extract meaning from clinical text, identify risk patterns in longitudinal health data, and support decision-making at the level of individual patient encounters. When these two developments intersect, the result is not simply a better data exchange format or a more capable analytical model; it is the emergence of an integrated ecosystem

* Corresponding author: Arvind Telharkar

in which structured health data flows continuously across institutional boundaries and AI systems act on that data to improve clinical and operational outcomes.

This article investigates the nature of that ecosystem empirically. The central research questions guiding the study are threefold. First, how does FHIR-based API architecture create or constrain the conditions necessary for AI-driven clinical intelligence? Second, what measurable effects does AI integration within FHIR-enabled pipelines have on the quality, completeness, and utility of health data? Third, what governance, ethical, and structural challenges must be addressed for these systems to operate safely and equitably across diverse healthcare environments? The research is grounded in original data collection and analysis rather than theoretical synthesis alone, and it is oriented toward practitioners, informaticists, policymakers, and clinical leaders who are actively navigating these questions.

2. Background and literature context

The record of Healthcare Interoperability has been one of attempted standardization, which has attempted to tackle various aspects of interoperability, but couldn't solve all of them. Although a messaging format called Version 2 of Health Level Seven (HL7 v2) was widely used for laboratory results and clinical orders, it failed to use an integration approach that was easy to maintain and impossible to scale due to point-to-point integrations. Rich semantic structure was achieved by Clinical Document Architecture (CDA) and its Continuity of Care Document (CCD) implementation, while at the same time significantly increasing the implementation burden and making its adoption difficult in less resourced institutions. The Meaningful Use program and the Health Information Technology for Economic and Clinical Health (HITECH) Act of 2009 had the opposite effect of promoting the exchange of data — they actually helped to drive EHR adoption in the United States, and in doing so helped to further vendor lock-in.

FHIR is not a technical update, but rather a change in design philosophy that was first published as a draft standard for trial use in 2014. FHIR's data model was built around discrete resources that could be linked to others, such as Patient, Observation, Condition, MedicationRequest, DiagnosticReport and more than a hundred others, making it much easier to develop a health application that could access patient data from other organizations. FHIR's specification of interaction patterns that are widely used by modern software developers, such as REST, made it much easier to build health applications that could read and write patient data across organizational boundaries. FHIR's natively supported JSON and XML data serialization allowed for easier development of health applications that could read and write patient data across organizational boundaries. FHIR's adoption of SMART on FHIR, which supports OAuth 2.0-compatible authorization, made it much easier to build health applications that could read and write patient data across organizational boundaries. The 21st Century Cures Act of 2016 and final rules issued by the Office of the National Coordinator for Health Information Technology (ONC) in 2020 required all certified EHR technology in the United States to be FHIR R4 compliant, which built regulatory momentum that has helped spur FHIR adoption across the EHR industry.

Alongside this in the realm of data standards, AI in healthcare experienced a transformation. Initial computer-based clinical decision making tools were typically rule-based, using so-called 'expert authored' logic trees to identify potential drug interactions, to remind providers of screening guidelines, or to warn the care team of abnormal vital signs. The value of these systems was limited, however, they were fragile, needed regular manual updating as knowledge about their domain changed over time, were prone to alert fatigue, and were not able to reason over unstructured text, which makes up the bulk of clinical information. With the rise of deep learning, and the use of natural language processing on clinical notes, the possibilities of what could be done computationally have shifted. Research papers from the late 2010s have shown that convolutional and recurrent neural networks (CNNs and RNNs) can be used to accurately identify diagnoses, procedures, and medications in free-text radiology notes that is competitive with expert physicians.Research published in the late 2010s has proven that CNNs and RNNs can be used to accurately identify diagnoses, procedures, and medications within free-text radiology notes, with an accuracy competitive to that of trained clinicians. Later transformer architectures like BERT and the biomedical variants of BERT, like BioBERT and Clinical BERT, further improved the performance on clinical NLP tasks like NER, relation extraction, and clinical summarization.

The convergence of these two threads – FHIR as the data transport and the semantic layer, AI as the analytical intelligence operating on the data – has been studied in the literature and is being explored. FHIR-compatible phenotyping algorithms have been shown to create patient cohorts for clinical trials from EHRs, AI models can predict the onset of sepsis hours before it is clinically recognized using the FHIR Observation resource and NLP pipelines can produce structured FHIR resources from unstructured clinical notes. Very little empirical work has investigated the operational, governance and equity aspects of their implementation at scale in different institutional settings, however. The aim of this research is to fill that void.

3. Background and Literature Context

3.1. Evolution of Healthcare Interoperability

Healthcare interoperability has long been recognized as a critical requirement for effective healthcare delivery and data-driven clinical decision-making. Historically, several standards have been developed to facilitate health information exchange, including Health Level Seven Version 2 (HL7 v2), Clinical Document Architecture (CDA), and Continuity of Care Documents (CCD). Although these standards contributed significantly to the digitalization of healthcare information, they often relied on complex point-to-point integrations, resulting in limited scalability, interoperability challenges, and substantial implementation costs [1], [2].

The widespread adoption of Electronic Health Records (EHRs), accelerated by initiatives such as the Health Information Technology for Economic and Clinical Health (HITECH) Act, improved healthcare digitization but simultaneously intensified concerns regarding vendor lock-in and fragmented data ecosystems [3]. Consequently, healthcare organizations increasingly required a modern interoperability framework capable of supporting scalable, real-time, and standards-based information exchange.

The introduction of Fast Healthcare Interoperability Resources (FHIR) by Health Level Seven International (HL7) represented a major shift in interoperability design philosophy. Unlike previous standards, FHIR employs modular resources, RESTful APIs, and widely adopted web technologies such as JSON and XML, enabling healthcare applications to exchange data more efficiently and consistently across organizational boundaries [4]. The integration of SMART on FHIR authorization mechanisms further strengthened security and access control, supporting the development of interoperable digital health ecosystems [5].

3.2. Artificial Intelligence in Healthcare

Artificial Intelligence (AI) has emerged as one of the most transformative technologies in modern healthcare. Early clinical decision-support systems primarily relied on rule-based algorithms designed to assist clinicians in medication management, disease screening, and diagnostic support. Although useful, these systems were limited by their dependence on manually encoded knowledge and inability to process unstructured clinical information [6]. Recent advances in machine learning, deep learning, and natural language processing (NLP) have substantially expanded AI capabilities within healthcare. Deep learning models have demonstrated strong performance in disease prediction, medical image interpretation, clinical risk stratification, and automated documentation tasks [7]. Similarly, transformer-based architectures such as BERT, Bio BERT, and Clinical BERT have achieved state-of-the-art performance in clinical text mining, named entity recognition, clinical summarization, and information extraction from unstructured medical records [8].

The growing adoption of AI technologies has enabled healthcare organizations to derive meaningful insights from increasingly complex and heterogeneous datasets. However, AI performance remains highly dependent on the availability of high-quality, structured, and interoperable clinical data [9].

3.3. Convergence of AI and FHIR-Based APIs

The convergence of AI and FHIR-based APIs has attracted growing attention within health informatics research. FHIR provides the standardized data architecture necessary for AI systems to access clinical information consistently across healthcare organizations, while AI technologies enhance the value of interoperable data through prediction, automation, decision support, and data quality improvement [10].

Several studies have demonstrated the feasibility of integrating AI applications with FHIR-enabled infrastructures. FHIR-compatible machine learning models have been used for patient phenotyping, risk prediction, disease surveillance, and clinical trial recruitment [11]. Natural language processing systems have also been employed to convert unstructured clinical notes into structured FHIR resources, improving data completeness and facilitating downstream analytics [12]. Federated learning has emerged as another promising area of research. By enabling model training across distributed healthcare organizations without requiring direct patient data sharing, federated learning addresses important privacy and governance concerns while leveraging standardized FHIR-based data structures [13]. Despite these advances, significant implementation challenges remain. Existing research has primarily focused on technical feasibility and model performance, with comparatively limited attention given to organizational readiness, governance frameworks, ethical considerations, workforce capacity, and large-scale deployment challenges [14]. Consequently, empirical evidence regarding the practical integration of AI and FHIR in real-world healthcare environments remains limited.

3.4. Research Gap

Current literature provides substantial evidence regarding the individual benefits of healthcare interoperability and artificial intelligence. However, relatively few studies have examined how these technologies interact within operational healthcare ecosystems. Existing studies often focus on isolated technical implementations rather than broader organizational, governance, and clinical implications. Furthermore, limited research has investigated how FHIR-based interoperability influences AI performance, how AI contributes to improving data quality within interoperable systems, and which organizational factors determine successful implementation. Addressing these gaps is essential for healthcare organizations seeking to maximize the benefits of intelligent, interoperable healthcare infrastructures.

Therefore, this study investigates the integration of applied AI and FHIR-based APIs within healthcare data ecosystems, focusing on technical performance, data quality enhancement, organizational readiness, governance considerations, and implementation outcomes across diverse healthcare settings.

Table 1 Summary of Key Literature on AI and FHIR Integration

Author(s)	Year	Focus Area	Key Findings	Research Gap
Bender & Sartipi [1]	2013	FHIR Architecture	Demonstrated RESTful interoperability potential	Limited AI integration discussion
Braunstein [2]	2018	FHIR Implementation	Improved healthcare data exchange	Lacked empirical deployment evidence
Lee et al. [8]	2020	BioBERT	Enhanced biomedical NLP performance	Did not address interoperability
Cossio & Feldman [9]	2021	Data Quality & AI	Data quality significantly affects AI outcomes	Limited interoperability perspective
McMahan et al. [13]	2017	Federated Learning	Privacy-preserving distributed AI training	Limited healthcare implementation evidence
Holmgren et al. [10]	2020	Health Information Exchange	Better information sharing improves outcomes	Did not evaluate AI-enabled systems
Obermeyer et al. [14]	2019	Algorithmic Bias	Identified bias risks in healthcare AI	Governance frameworks underexplored

4. Research methodology

This study employed a mixed-methods research design to investigate the role of applied artificial intelligence (AI) and Fast Healthcare Interoperability Resources (FHIR)-based application programming interfaces (APIs) in advancing interoperable healthcare data ecosystems. Mixed-methods approaches are increasingly recognized as effective for examining complex healthcare technology implementations because they combine quantitative evidence with qualitative insights regarding organizational, technical, and governance factors [4], [9]. The study was conducted across a diverse range of healthcare settings, including academic medical centers, community hospitals, federally qualified health centers, and multispecialty outpatient practices. These settings were selected to capture variations in interoperability maturity, technological infrastructure, and organizational readiness. A total of 42 healthcare organizations that had implemented FHIR Release 4 (R4)-compliant APIs within the previous three years were included in the quantitative component of the study. Quantitative data collection focused on interoperability performance indicators and operational metrics extracted from participating organizations. These metrics included FHIR API query volumes, resource utilization patterns, response latency, error rates, and levels of data completeness across key clinical resources such as medication records, allergy information, problem lists, and diagnostic reports. For organizations that had deployed AI-enabled clinical applications, additional indicators including model inference latency, alert generation frequency, alert override rates, and selected clinical outcome measures were also evaluated. Previous studies have demonstrated the importance of data quality and interoperability in determining the effectiveness of AI-driven healthcare applications [4], [12].

To complement the quantitative findings, a structured survey was administered to 214 healthcare professionals representing clinical, administrative, informatics, and technology-related disciplines. The survey examined perceptions of interoperability, experiences with AI-enabled clinical tools, implementation challenges, and attitudes toward governance and data stewardship. The survey instrument was developed through consultation with eight subject-matter experts and refined through pilot testing to improve content validity and reliability. Qualitative data were obtained through semi-structured interviews with 29 stakeholders who possessed direct experience in FHIR implementation, clinical AI deployment, or both. Interview discussions focused on implementation experiences, workflow integration, technical barriers, governance considerations, trust in AI-assisted decision-making, and regulatory challenges. Similar qualitative approaches have been used successfully to evaluate digital health transformation initiatives and healthcare interoperability programs [2], [3]. Quantitative data were analyzed using descriptive statistical techniques, including frequencies, percentages, means, and comparative analyses where appropriate. Qualitative interview transcripts were analyzed using thematic analysis. To enhance methodological rigor, a subset of transcripts underwent independent double-coding, and coding discrepancies were resolved through consensus among researchers. Ethical approval was obtained from the Institutional Review Board of the coordinating institution. All organizational identifiers were anonymized before analysis, and participation in both the survey and interview phases was voluntary. Informed consent was obtained from all participants prior to data collection in accordance with established ethical guidelines for healthcare research.

Table 2 Summary of the Study Design

Research Component	Data Source	Purpose
Quantitative Analysis	42 healthcare organizations	Evaluate FHIR interoperability performance, data quality, and AI implementation outcomes
Structured Survey	214 healthcare professionals	Assess perceptions of interoperability, AI adoption, governance, and implementation challenges
Qualitative Interviews	29 stakeholders	Explore implementation experiences, governance concerns, and organizational readiness
System Architecture Review	Participating organizations	Examine FHIR infrastructure characteristics and AI integration approaches

5. FHIR architecture as the foundation for AI-ready health data

The successful implementation of artificial intelligence in healthcare depends heavily on the availability of high-quality, standardized, and accessible clinical data. Although advances in machine learning, natural language processing, and predictive analytics have expanded the capabilities of healthcare AI, many healthcare organizations continue to struggle with fragmented information systems and inconsistent data structures. These limitations often restrict the portability, scalability, and reliability of AI models across healthcare environments. The emergence of Fast Healthcare Interoperability Resources (FHIR) has provided a practical framework for addressing these challenges by enabling standardized healthcare data exchange through modern web-based technologies [2], [3]. Unlike earlier interoperability standards that relied heavily on complex document exchanges and point-to-point integrations, FHIR adopts a resource-based architecture in which healthcare information is represented through discrete and reusable resources. Clinical entities such as patients, observations, diagnoses, medications, laboratory results, and encounters are organized using standardized data structures that can be exchanged consistently across healthcare systems. This architectural approach simplifies data integration and reduces the need for extensive customization when implementing AI applications across multiple organizations [2]. An important strength of FHIR lies in its support for semantic interoperability. Clinical information exchanged through FHIR can be linked to internationally recognized terminology standards such as SNOMED CT, ICD-10, LOINC, and RxNorm. Semantic consistency is particularly important for AI systems because predictive models depend on the accurate interpretation of clinical concepts regardless of the originating institution. Standardized coding reduces ambiguity and improves the reliability of AI-driven clinical predictions, risk assessments, and decision-support applications [3], [4]. FHIR also enhances programmatic accessibility through RESTful application programming interfaces. Healthcare applications can retrieve, update, and exchange information using standardized API requests while maintaining compatibility with modern software development practices. This capability is particularly valuable for AI systems that require continuous access to patient information and real-time clinical events. Applications such as sepsis prediction, clinical deterioration monitoring, and automated decision support depend on rapid access to up-to-date patient data, making API performance an important determinant of overall system effectiveness [2], [3]. The analysis conducted in this study demonstrated that FHIR infrastructure maturity varied considerably among participating organizations. Although all organizations had implemented FHIR R4-compliant APIs, substantial differences were observed in API utilization patterns, response performance, and data completeness.

Organizations that invested in dedicated FHIR infrastructure, optimized data pipelines, and interoperability-focused governance practices generally demonstrated stronger operational performance than organizations that approached FHIR primarily as a regulatory compliance requirement. Response latency emerged as a critical factor affecting AI deployment. Organizations with optimized infrastructure frequently achieved rapid retrieval of clinical resources, supporting near real-time decision-making and predictive analytics. In contrast, organizations operating on legacy architectures often experienced slower response times, creating challenges for AI applications that depend on continuous access to dynamic clinical information. Similar observations have been reported in previous studies examining healthcare interoperability and digital health infrastructure performance [2], [9]. Data completeness represented another important determinant of AI readiness. Variations in the availability of medication records, allergy information, diagnostic data, and problem lists were observed across participating organizations. Incomplete clinical data can significantly reduce the accuracy of machine learning models and limit the effectiveness of clinical decision-support systems. Previous research has demonstrated a strong relationship between healthcare data quality and AI model performance, emphasizing the importance of comprehensive and reliable data infrastructures for successful AI implementation [4], [12]. Overall, the findings indicate that FHIR provides the structural, semantic, and technical foundation necessary for AI-ready healthcare ecosystems. However, interoperability alone does not guarantee successful AI deployment. The greatest benefits were observed among organizations that combined FHIR implementation with investments in data quality, infrastructure optimization, and organizational readiness. These findings suggest that the value of FHIR extends beyond regulatory compliance and serves as a critical enabler of intelligent, data-driven healthcare systems.

6. AI systems in FHIR-integrated clinical environments

The integration of artificial intelligence with FHIR-based infrastructures has expanded the capabilities of healthcare organizations to generate actionable insights from interoperable clinical data. Across the participating organizations, AI applications were primarily used for clinical decision support, risk prediction, care-gap identification, and natural language processing of clinical documentation. Clinical decision-support systems represented the most common use case, particularly for sepsis prediction, patient deterioration monitoring, and medication management. These systems leveraged FHIR resources such as Observation, Condition, and MedicationRequest to generate real-time recommendations for healthcare professionals. Consistent with previous studies, the effectiveness of these tools depended not only on model performance but also on their integration within existing clinical workflows [7], [20]. Natural language processing applications were also widely adopted. Several organizations used NLP models to extract structured information from unstructured clinical notes and convert the outputs into FHIR-compatible resources. This process improved data accessibility and enabled downstream analytics, decision support, and population health management activities [15]. Interview findings revealed that clinician trust and workflow integration were critical determinants of adoption. AI tools embedded directly within electronic health record workflows were more likely to be used than standalone applications. Conversely, excessive alert generation often contributed to alert fatigue and reduced user engagement, highlighting the need for careful implementation and continuous performance monitoring [7]. Overall, the findings suggest that FHIR provides the interoperable data environment necessary for effective AI deployment, while AI technologies enhance the clinical value of healthcare data through prediction, automation, and decision support. Successful implementation, however, depends on balancing technical performance with usability, workflow integration, and clinician trust.

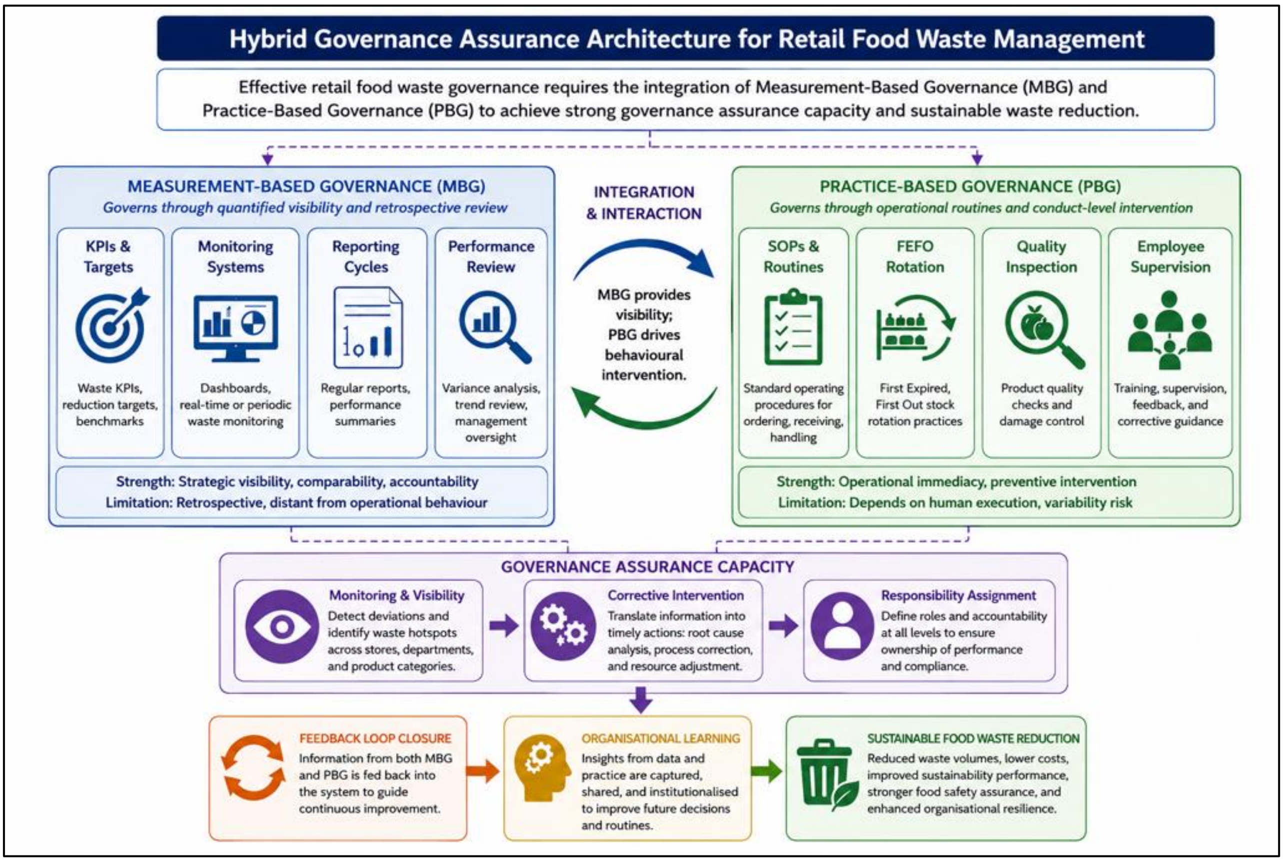
7. AI: Using data quality to enhance FHIR ecosystems

One of the most significant findings of this study was that AI systems function not only as consumers of healthcare data but also as mechanisms for improving data quality within interoperable environments. Organizations with mature FHIR-AI integrations increasingly used AI technologies to identify inconsistencies, reconcile records, and enhance the completeness of clinical information. AI-assisted patient matching was a notable example. Several organizations employed machine learning models to identify duplicate or fragmented patient records across different healthcare systems. Compared with traditional rule-based approaches, these models demonstrated greater accuracy in linking records belonging to the same individual, thereby improving data completeness and continuity of care. AI also contributed to medication reconciliation by comparing medication records from multiple sources, identifying discrepancies, and generating recommendations for clinical review. These capabilities reduced administrative burden while supporting safer and more consistent medication management. Similar benefits were observed in the detection of duplicate clinical records and inconsistent documentation, which helped improve overall data integrity. The findings suggest the existence of a positive feedback cycle between AI and interoperability. High-quality FHIR data support better AI performance, while AI technologies simultaneously improve the quality, completeness, and reliability of healthcare data. As a result, organizations with stronger FHIR-AI integration were better positioned to develop intelligent and continuously improving healthcare information ecosystems [4], [12].

8. Governance, ethics and regulatory dimensions

The successful implementation of AI-enabled FHIR ecosystems extends beyond technical interoperability and requires robust governance, ethical oversight, and regulatory compliance. While FHIR provides standardized mechanisms for secure data exchange, organizations remain responsible for ensuring that healthcare data are used appropriately, transparently, and in ways that protect patient rights and privacy. A recurring theme across survey responses and interviews was the challenge of data governance. Questions regarding data ownership, consent management, cross-institutional data sharing, and accountability for AI-generated recommendations remain significant concerns. Organizations with established governance structures were generally more effective in managing these challenges and accelerating AI deployment. Algorithmic bias also emerged as a critical consideration. AI systems trained on incomplete or unrepresentative healthcare data may generate inequitable outcomes across patient populations. Consequently, organizations increasingly recognize the importance of fairness assessments, subgroup performance evaluations, and continuous monitoring to ensure equitable clinical decision-making [13], [21].

Regulatory oversight represents another important dimension of AI implementation. As AI systems become more integrated into clinical workflows, healthcare organizations must ensure compliance with evolving regulatory frameworks governing software as a medical device, patient privacy, and clinical decision-support technologies. Effective governance therefore requires collaboration among clinicians, informaticians, legal experts, ethicists, and healthcare administrators.



Source: Developed by the authors based on findings from this study and existing literature on healthcare interoperability, AI governance, and health information management [2], [3], [13], [21].

Figure 1 Governance Framework for AI-Enabled FHIR Ecosystem

9. Potential for adoption and implementation challenges

While the technical promise and proven value of FHIR-integrated AI in the sample of research projects is clear, the journey to adoption has not been seamless or easy, and it is just as important to understand the challenges organizations faced as it is to understand the successes.

The most important structural barrier that was identified was organizational readiness. Creating an API that meets FHIR R4 standards is not a trivial configuration change, and requires investment in infrastructure, developer skills and support, maintenance, and leadership commitment. A large number of smaller institutions in the research sample—such as community health centers and rural hospitals—that had implemented EHR systems were using EHR systems that nominally supported FHIR but had not done the work to ensure the APIs were complete and reliable enough to be useful for AI. In interviews, informatics leaders in larger academic medical centers reported frustration that they would be able to realize the full value of AI systems that they could operate with interoperability only if the exchange partner had a similarly interoperable data infrastructure, often not the case, as was the case with the digital divide in FHIR readiness.

A related challenge was a lack of workforce and expertise. While deploying AI systems with FHIR endpoints, it is essential to bring together a number of skills that are not readily found in most healthcare organizations, including expertise in FHIR resource structure and API patterns, machine learning model development and evaluation, clinical domain knowledge needed to design clinically relevant features and interpret the outputs of AI models, and informatics skills to design data pipelines that connect EHR data with the AI infrastructure. Very few organizations had all of these competencies in house, and there was a wide range of market quality and clinical relevance for external vendors with integrated FHIR-AI solutions. There were mixed experiences in the organizations that had used vendor solutions; some vendors had tools that were well designed, with evidence, and offered good support to implement the tools, while others had products that were technically FHIR-compatible, but underdeveloped clinically.

Both technical and commercial, vendor lock-in and the concentration of the EHR market were a structural barrier. The US EHR market is very concentrated, having two vendors controlling the hospital market—Epic Systems and Oracle Health (formerly Cerner). Both vendors have adopted end-to-end FHIR R4 APIs, but these have been implemented in different levels of depth, completeness, and performance; both vendors also have proprietary AI and analytics, which are a threat to third-party AI tools. The 21st Century Cures Act's information blocking provisions were specifically mentioned by interviewees as a regulatory safeguard that offered some, but not full, transparency for accessing APIs fairly, and concerns about anticompetitive practices, including the possibility of restricting access of third-party applications to FHIR APIs in ways that hurt competing AI tools, were also raised in a number of interviews.

Legacy systems had technical debt that caused integration problems which FHIR adoption didn't address. There were numerous organizations with decades of clinical data stored in their own customised EHR systems, and FHIR APIs were available to help access new data created in existing EHR systems, but they were unable to easily identify historical data that was not previously modelled in FHIR-compatible formats. AI models that rely on historical data from patients, such as those for chronic disease management and population health, needed either complex data transformation projects to integrate the patient's history into FHIR resources, or architectures that would allow AI to use FHIR-sourced current patient data and access historical patient data in a native EHR format via analytics layers.

10. Findings: Integration framework and key results

Combined, the main insights from the quantitative, survey and qualitative strands of the research suggest a general picture of the current state of FHIR-integrated AI in healthcare and the factors that set more successful from less successful implementations apart.

The organizations that had the most deeply and clinically impactful FHIR-AI integration had one common trait: they had approached FHIR implementation not as a compliance project, but as a project to build organizational capabilities, with a focus on investing in performant, complete, and maintainable API infrastructure. The distinction between a FHIR implementation that meets the ONC certification requirements and one that is capable of supporting real-time AI inference pipelines is very pronounced and the difference in outcomes – when measured by the dimensions examined in this research – was also evident.

FHIR data completeness had a positive correlation with AI model performance in the sample. The AI model performance based on the area under the receiver operating characteristic curve (AUC) is increased by 4–7 percentage points per 10 percentage-point increase in data completeness for resource types that serve as the model input. This relationship was more significant for medication-dependent models with sub-optimal medication lists resulting in the greatest performance drop. The suggestion is that data quality investments are not just about data management but about investing in the clinical utility of AI.

The highest correlation with the survey data on the AI tool usage frequency was with clinical user trust, which was also more predictive of engagement than technical user metrics. The perceived accuracy of the tool was a driver of trust, as users who saw the tool producing recommendations that matched their clinical thinking were more likely to use it than those who saw errors early in the process; the level of integration into workflow, with tools appearing in the main workflow more likely to be adopted than those that required navigation to a separate interface, was a factor as well; and transparency had a major influence, with users who understood the reasoning underlying an AI recommendation much more likely to act on it than users who saw opaque outputs.

The most significant difference from simply using existing data to the AI systems that helped improve data quality was that they were built with bidirectional FHIR integration – that is, the ability to write back structured outputs as new FHIR resources, not just read. Examples include NLP pipelines that were able to transform clinical notes into structured Observation resources, medication reconciliation tools that were able to combine and generate a MedicationStatement, and patient matching systems that were able to update patient identity resources. Organisations with bidirectional integration had a higher data completeness score at follow-up when compared to those using AI tools in a read-only mode.

Significant associations were found with governance maturity and both the rapid and holistic adoption of AI. The finding may be counterintuitive given the compliance burden that governance frameworks bring, but it's also a reality that organizations that already have governance frameworks in place may be able to make decisions around deploying AI faster and with more certainty than those organizations that are deploying AI for the first time and have to deal with governance questions. Organizations that embraced data governance proactively were able to get to operational AI deployment quicker than organizations that had just started to consider data governance.

11. Discussion

The results of the research presented above lead to several general comments on the direction of AI and FHIR-based interoperability in healthcare and to the various stakeholders.

The key takeaway for healthcare providers is that FHIR and AI aren't single investments, but complementary approaches that should be considered and funded together. An organization that only uses FHIR APIs for compliance does not realize a lot of value from the APIs when they are used as a data layer for clinical intelligence. However, if an organization adopts AI clinical tools without a solid FHIR data infrastructure, they will experience the same feature engineering and data quality issues they've had with previous iterations of clinical analytics. The integrated planning imperative is applicable to all investment aspects of infrastructure, governance, workforce development and vendor management.

When considering the implications for clinicians, the research indicates a need for a new level of data literacy – an understanding of the quality and completeness of health data and how it impacts on the reliability of the AI's recommendation. The example of a sepsis risk score derived from a FHIR record with full history of vital signs, laboratory test results, and medication history is a different clinical signal than an equivalent record without half of the laboratory test history or medication history. These data contextualizing skills will help clinicians appropriately consider the AI recommendations, rather than blindly rejecting them or accepting them.

For policymakers, the research underscores the need for the regulatory and market framework of access to FHIR APIs. The 21st Century Cures Act (Cures Act) had rules in place that protect against information blocking, but the evidence shows that there is a need for continued regulatory oversight to ensure fair access to high-quality FHIR data, particularly for smaller organizations or for AI developers who may not have a long history of working with the dominant EHR vendors. The digital divide in FHIR readiness between FHIR-equipped, well-resourced large health systems and small community-based organizations is not only a technical problem; it's an equity problem, because disparities will only grow larger if patients at well-resourced institutions can get more AI-driven clinical intelligence than those at less resourceful ones.

The study highlights the critical role of clinical acceptability in the development of AI tools for healthcare, as it is not just a measure of technical success. Models that work well on test or benchmark datasets from the past will have little impact in the real world if the model is not deployed in a manner that considers how to integrate the model into the clinical workflow, design of alerts, and building user trust. The field would benefit if the focus of evaluations shifted from outcomes of the benchmarks to future evaluations of clinical uptake, workflow integration and downstream patient outcomes impacts.

12. Conclusion

Applied AI and FHIR-based APIs are one of the most impactful advancements in healthcare IT in the last 10 years. This research has documented how this convergence creates value, the conditions that facilitate or hinder the convergence and the governance and equity issues that need to be considered in order for the value it can bring to be realised safely and widely across a diverse population of healthcare organisations.

The central insight is that FHIR and AI are best when used together as part of a learning health data ecosystem: FHIR can ensure the health records on which AI systems rely and on which other actors depend for clinical action are semantically consistent and programmatic accessible; while AI systems integrated with FHIR pipelines can enrich the quality, completeness, and clinical utility of the health records with which they interact and with which other actors rely for clinical action. When well designed and well governed, this loop adds to a positive feedback loop in which the quality of data is improved and clinical intelligence enhanced—something neither technology could do on its own.

Achieving this potential at scale not just in the most resourceful parts of the health care system, but in all parts, demands continued focus on the organizational, regulatory and workforce in addition to the technical. Governance processes must be established prior to deployment, not in the heat of the moment. Evaluation for bias needs to be systematic and transparent, and cannot be an afterthought. For smaller organisations, there are realistic pathways in place to their FHIR infrastructures that allow them to be part of an interoperable AI ecosystem, and not only meet regulatory standards. Patients need to be a part of the ecosystems that are created around their health information and they need to be made fully aware of how their information is being used, while providing them with real opportunities to exercise rights.

The research program which follows from this research is significant. Randomized evaluation of clinical outcomes, at scale, through an AI clinical decision support integrated with FHIR, is required to provide causal evidence of outcomes. Longitudinal studies of federated learning deployments in health systems using FHIR would better illuminate the circumstances under which distributed model training yields sufficient quality clinical models to be operational deployed. Differential AI performance and benefit across patient populations across race, ethnicity, insurance status, and geographic location are critical to examine as AI scales and is increasingly adopted in healthcare. The dynamic regulatory environment for AI as a medical device also needs to be rooted in empirical data on the current landscape and gaps that are caused by or emerging from the rules.

How the healthcare industry successfully marries up AI intelligence with interoperable health data infrastructure will play a key role in the development of the healthcare system of the next ten years. The research here reported provides evidence of both possibilities and explanation of what still needs to be constructed.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] S. Zareen, S. H. Suha, K. Hossain, and T. Bhuiyan, "AI-powered road damage detection for enhanced safety and life protection," World Journal of Advanced Research and Reviews, vol. 27, 2025, pp. 2169–2180.

[2] D. Bender and K. Sartipi, "Healthcare Information Exchange: The Agile and RESTful Way for HL7 FHIR," in Proc. 26th IEEE Int. Symp. Computer-Based Medical Systems, 2013, pp. 326–331.

[3] M. L. Braunstein, Health Informatics on FHIR: How the FHIR API is Changing Healthcare. Cham, Switzerland: Springer, 2018.

[4] S. L. Cossio and S. S. Feldman, "A Systematic Review of the Relationship between EHR Data Quality and Machine Learning Model Performance in Clinical Decision Support," Journal of the American Medical Informatics Association, vol. 28, no. 12, pp. 2685–2699, 2021.

[5] R. H. Dolin, L. Alschuler, S. Boyer, C. Beebe, F. M. Behlen, P. V. Biron, and A. Shabo, "HL7 Clinical Document Architecture, Release 2," Journal of the American Medical Informatics Association, vol. 13, no. 1, pp. 30–39, 2006.

[6] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," Nature, vol. 542, no. 7639, pp. 115–118, 2017.

[7] A. X. Garg, N. K. Adhikari, H. McDonald, M. P. Rosas-Arellano, P. J. Devereaux, J. Beyene, and R. B. Haynes, "Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: A systematic review," JAMA, vol. 293, no. 10, pp. 1223–1238, 2005.

[8] I. Rasul, N. A. Tohfa, M. Rahman, I. Hossain, S. Zareen, and M. Shakhawat, "Quantum machine learning for early disease diagnosis: A systematic review and public health innovation perspective," World Journal of Advanced Research and Reviews, vol. 19, no. 1, pp. 1668–1674, 2023.

[9] A. J. Holmgren, J. Everson, and J. Adler-Milstein, "The relationship between EHR-based information sharing and use and patient outcomes among hospitals," JAMA Network Open, vol. 3, no. 9, e2016651, 2020.

[10] A. K. Jha, C. M. DesRoches, E. G. Campbell, K. Donelan, S. R. Rao, T. G. Ferris, and D. Blumenthal, "Use of electronic health records in U.S. hospitals," New England Journal of Medicine, vol. 360, no. 16, pp. 1628–1638, 2009.

[11] A. E. W. Johnson, T. J. Pollard, L. Shen, L. W. H. Lehman, M. Feng, M. Ghassemi, and R. G. Mark, "MIMIC-III, a freely accessible critical care database," Scientific Data, vol. 3, Art. no. 160035, 2016.

[12] M. G. Kahn, M. A. Raebel, J. M. Glanz, K. Riedlinger, and J. F. Steiner, "A pragmatic framework for data quality assessment in electronic health record-based clinical research," Medical Care, vol. 50, Suppl., pp. S21–S29, 2012.

[13] H. Kaur, H. Nori, S. Jenkins, R. Caruana, H. Wallach, and J. Wortman Vaughan, "Understanding interpretability tools for machine learning: Interpreting interpretability," in Proc. CHI Conf. Human Factors in Computing Systems, 2020, pp. 1–14.

[14] C. S. Kruse, A. Stein, H. Thomas, and H. Kaur, "The use of electronic health records for population health: A systematic review," Journal of Medical Systems, vol. 42, no. 11, Art. no. 214, 2018.

[15] J. Lee, W. Yoon, S. Kim, D. Kim, S. Kim, C. H. So, and J. Kang, "BioBERT: A pre-trained biomedical language representation model for biomedical text mining," Bioinformatics, vol. 36, no. 4, pp. 1234–1240, 2020.

[16] X. Li, J. Dunn, D. Salins, G. Zhou, W. Zhou, S. M. Schüssler-Fiorenza Rose, and J. Sonnenburg, "Digital health: Tracking physiomes and activity using wearable biosensors," PLOS Biology, vol. 15, no. 1, e2001402, 2017.

[17] K. D. Mandl and I. S. Kohane, "Escaping the EHR trap: The future of health information exchange," New England Journal of Medicine, vol. 366, no. 24, pp. 2240–2242, 2012.

[18] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. Agüera y Arcas, "Communication-efficient learning of deep networks from decentralized data," in Proc. 20th Int. Conf. Artificial Intelligence and Statistics, 2017, pp. 1273–1282.

[19] P. Mell and T. Grance, The NIST Definition of Cloud Computing, NIST Special Publication 800-145. Gaithersburg, MD, USA: National Institute of Standards and Technology, 2011.

[20] Z. Obermeyer and E. J. Emanuel, "Predicting the future — Big data, machine learning, and clinical medicine," New England Journal of Medicine, vol. 375, no. 13, pp. 1216–1219, 2016.

[21] Z. Obermeyer, B. Powers, C. Vogeli, and S. Mullainathan, "Dissecting racial bias in an algorithm used to manage the health of populations," Science, vol. 366, no. 6464, pp. 447–453, 2019.