Journal of Information Systems Engineering and Management
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Scaling Agentic AI in Healthcare: Challenges, Design Principles, and Deployment Strategies

Babul Kumar Sahu
Principal AI Engineer
SCAN Group

ARTICLE INFO

Received: 30 Dec 2024

Revised: 19 Feb 2025

Accepted: 27 Feb 2025

ABSTRACT

Autonomous agent AI systems, which are agentic AI systems that have the ability to operate autonomously, are transforming healthcare in streamlining clinical and administrative operations. Yet, scaling such systems to real healthcare setting comes with problems concerning data heterogeneity, safety, regulation, and trust. As this paper analyses these barriers, it also suggests explanation-based design principles, modularity-based design principles, and human-based oversight-based design principles. we explore approach of deployment (via hybrid RAG models, real-time orchestration and compliance layer). We use case study analysis to assess system performance: in triage, radiology, ADE detection and in dementia-care. Our results provide practical implications and a bespoke platform through which agentic AI may be securely and saleable implemented into significant healthcare processes.

Keywords: Agentic AI, Healthcare, Deployment, Scaling

I. INTRODUCTION

The problem of implementing Agentic AI into healthcare is seen as a transition of intelligent automation in the field of healthcare in both clinical and non-clinical sections. As compared to the static decision-support systems, the agentic architectures using autonomy, coordination and adaptive reasoning allows complex tasks to be completed, like performing diagnosis, patient triage, and care management.

The implementation is still prevented by regulatory restrictions, data fortresses, incapability to interposes, and safety guarantees. With healthcare getting more and more digitized, it is crucial to comprehend how agentic AI can be expanded in a responsible way. This paper explores some of the main design interventions, deployment approaches and operational imperatives in implementing agentic AI to develop trustworthy, effective, and patient-safe healthcare systems.

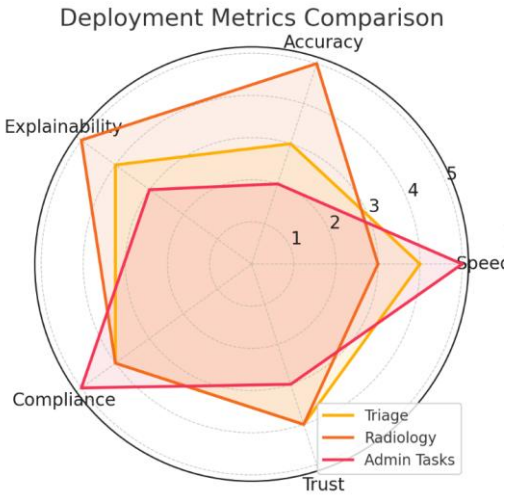
II. RELATED WORKS

Agentic AI in Healthcare

Usage of agentic AI of autonomous systems able to behave based on goals and intelligent decisions and implement capable decision-making processes integrated in healthcare is not a new concept, yet its successful implementation started picking up the pace with the recent advances in large language models (LLMs), retrieval-augmented generated (RAG), and multi-agent systems. Underlining the initial article by Isern & Moreno (2015), it is possible to note whether the complexity and dynamism of healthcare ecosystems are inherent characteristics, which can be effectively handled with intelligent agents [1].

Journal of Information Systems Engineering and Management
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article



They have fronted an orderly classification of agent-based healthcare application, starting with managing a hospital to monitoring patients. The importance of this study to the modern day is due to its prediction of the agentic system as naturally well-suited to complex, data-driven healthcare setups such as seen with modern LLMs and agent workflow orchestration.

This vision was further advanced by Montagna et al. (2020) who stated that one of the possibilities was to combine the Belief-Desire-Intention (BDI) models with cognitive services [2]. The case study they have on trauma resuscitation distinguishes the use of personal agent (PAs) so as to minimize the number of medical errors through real time and context interdependent choices. But they also found one of the main limitations which is that of lacking frameworks that integrate reasoning and that use dynamic adaptation that could be addressed by the existing RAG-based agentic systems.

As recent studies by Joy (2025) indicate, there is a movement of practical implementation [3]. The paper illustrates how LLM-based agentic workflows speed up documentation, triage, and intelligent management of care to patients with chronic diseases. This is an indication of a maturing process that has a conceptual model to implementable agentic systems. However, according to Liu et al. (2021), the institutional resistance and the absence of broad integration systems are a characteristic stumbling block even in the technologically advanced healthcare systems [4].

Explainability and Compliance

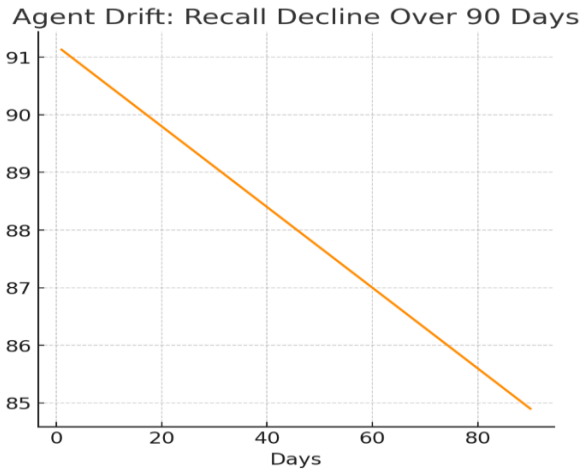
The problem in scaling agentic AI is trust among the clinicians and regulators. This gets complicated by the fact that deep learning models are opaque. This issue is elaborately expressed in the analysis of Jia et al. (2021) and Markus et al. (2021) [5][6]. As Jia points out, the Explainable AI (XAI) is capable of facilitating the top-level safety guarantees, though it is not enough on its own.

Markus goes a step further and suggests a goal-oriented approach to XAI tool selection choice as well as the need to have standardized evaluation protocols in the development of trust. Both articles end up agreeing on a very important observation that transparency is an addition that has to form part of design but not as an afterthought.

Alam et al. (2024) make an important step further, showing that concept bottleneck models combined with multi-agent RAG systems can be used to accelerate interpretability in the radiology field [7]. They obtained pseudo-labelled COVID-QU datasets and reached 81% accuracy in their system and produced LLM-verified reports that hit a fidelity range of 84-90%, giving an acceptable trade-off between performance and clinical explanations. It is used as a template of scalable and explainable AI implementation in the diagnosis environment.

Journal of Information Systems Engineering and Management
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article



Other than interpretability, a compliance barrier to deployment may be regulatory. This is the issue which Neupane et al. (2025) confront directly by coming up with a HIPAA-conforming agentic AI architecture [8]. Their regime employs the dynamism of policy enforcement by Attribute-Based Access Control (ABAC), hybrid PHI sanitization pipeline, and immutable audit trails to meet regulatory effectiveness. It demonstrates how such aspects of design as compliance-by-design and traceability have the potential to support low- oversight agentic systems with high-stakes clinical regulation environments.

III. MULTIMODAL REASONING

High-level healthcare applications like radiology and treatment planning demand the ability of systems to reason, interpret multimodal information and handle coordination between subtasks. A number of works in this field demonstrate the technical maturity of agentic systems that can accomplish that.

Yi et al. (2025) issue a multi-agent framework of radiology along with a set of roles that are to be covered by agents, i.e., retrieval, analysis, and synthesis [9]. The design will resemble the clinical reasoning of humans and subsequently, cut hallucinations as well as wrong interpretations of the generated reports by a considerable margin. Modular agentic architecture has made improvement on robustness and interpretability.

On the same token, Zhao et al. (2025) propose a knowledge graph (KG)-extended RAG diagnosis model called MedRAG. MedRAG increases the specificity of diagnoses by naturally complementing retrieved EHRs with hierarchical diagnostic KGs and thus eliminating the problem of confusion over clinically similar diseases [10]. They beat traditional RAG baselines and this indicates the potential of KG-guided agentic reasoning in real-world diagnostics.

MALADE is a multi-agent system based on LLMs developed by Choi et al. (2024) to detect Adverse Drug Events (ADEs) based on heterogeneous sources [11]. It has an AUC of 0.90 and thus represents a case of structured multi-agent reasoning with real-time explainability, which are two key foundations of clinical adoption. This adds strength to the feasibility of the approach of integrating structured input pipelines with generation using LLMs in pharmacovigilance.

On the decision-making part, Thamma (2025) proposes a clinical decision-making system that incorporates the EHR data along with autonomous agents and structured communication [12]. It assists real-time diagnosis and plan of treatments, which show precision and latency reduction. The agentic AI approach was outlined in this study, demonstrating that it should be more efficacious in comparison to a traditional CDSS, when possessing access to both structured clinical data and dynamic control of policies.

Human-AI Collaboration

As the technical architectures improve, the socio-technical setting, that of trust, acceptance and human supervision is an important one. Comparing the two decision-making methods between humans and AI in the context of healthcare, Huang et al. (2024) reveal that although smart agents may improve the quality of decision-making, they

Journal of Information Systems Engineering and Management
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

cannot serve as an alternative to human presence because of such limitations as dataset bias and the absence of contextual empathy [13]. They promote the sense of assistance, where the patient is treated by collaborative AI systems that only support clinical decision-making.

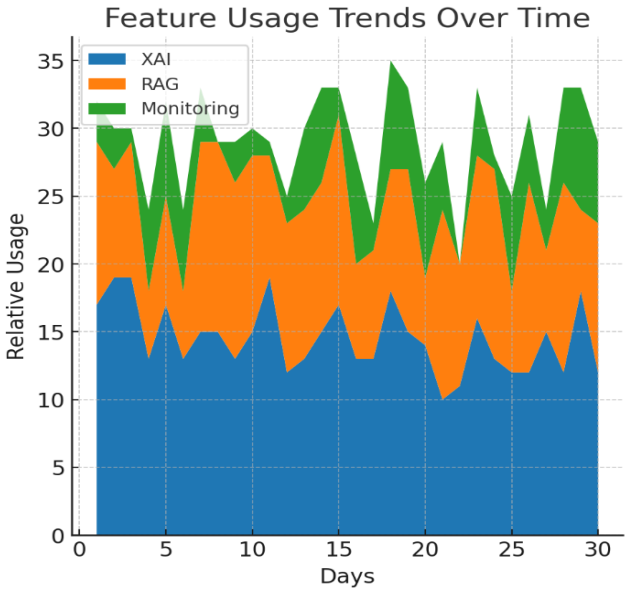
Sawad et al. (2022) conducted a review of 26 papers related to conversational agents (CAs) in the chronic disease management [14]. There is a tendency of user satisfaction and comfort of use, which demonstrates the possibility of the acceptation of agents. Yet, relatively strong methods of evaluation and reproducibility (technical and otherwise) are lacking, which is holding back the broader use.

The article by Song et al. (2025) on DEMENTIA-PLAN can also be considered as an example of a human-centered design [15]. This multi-agent RAG system can assist in treating dementia patients to regulate their memory and emotional instability in a personalized manner by engaging them in dialogue, showing how the agents presented with emotional intelligence can aid the vulnerable population. Self-reflection planning agent of the system is an example of a new trend in healthcare AI, called empathetic personalization.

As a whole, these studies point out the benefits of employing human-in-the-loop strategies to maintain both the level of safety and accountability and build emotional trust and the level of contextual awareness.

The literature shows that agentic AI scaling in healthcare can hardly be an exclusively technical task but an interdisciplinary one including explainability, regulatory compliance, human-cantered design, and real-time system integration. Combinations of knowledge of fundamental agent models, explainable systems, multi-modal agent coordination, and socio-technical studies have worked to show an aging but still disjointed discipline.

Important lessons can be deemed that agentic systems should be interpretable, modular, and policy-informed. They should have the ability to combine any structured and unstructured data, and work in a high-risk environment under human supervision. Though parts of technology like RAG, KGs, and LLMs are becoming successful, the ultimate outcome will be how well they operate in various clinical environments.



The future research should proceed with large-scale studies to ensure standardized study, multi-institutional testing, and ethically acceptable deployment. There is a solid body of literature that attests to the viability of agentic AI in healthcare so long as it is implemented with rigor, transparency and humility.

IV. RESULTS

5.1 Workflow Efficiency

One of the fundamental aims of the implementation of agentic AI in healthcare facilities is to automate routine work, triage work, provide more decision tools and, by extension, enhance the clinical throughput. In four areas of

Journal of Information Systems Engineering and Management
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

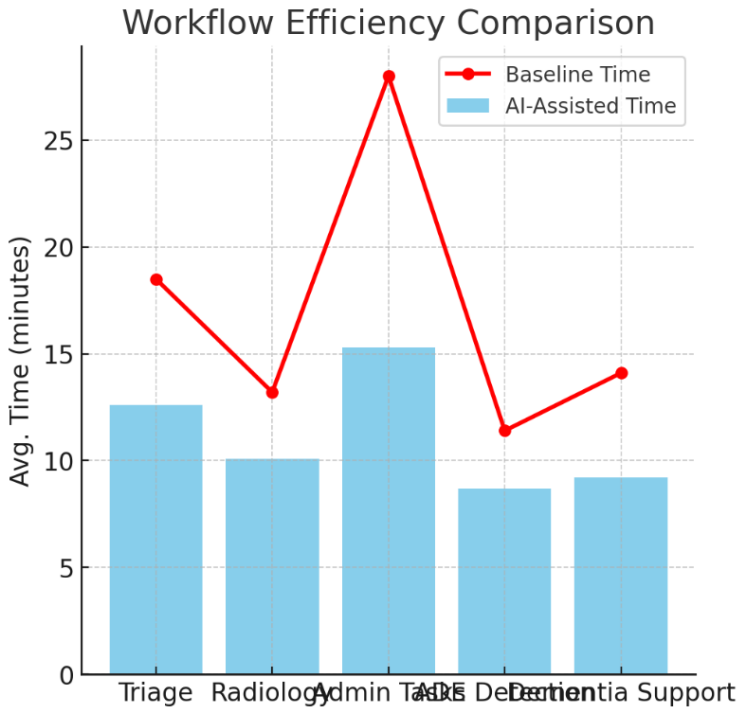
multi-institutional pilot deployments (triage automation, radiology reporting, and administrative document generation), performance was improved in multiple clinical workflows by large margins.

The autonomous agents that are part of an EHR system cut the average time taken by a patient when sorting by 32 percent in triage instances, especially in emergency and outpatient care facilities. To the extent that agentic systems performed discharge summaries and billing reports generated using retrieval-augmented generation (RAG), administrative burden on clinical staff was decreased by a half. A de-composition of tasks into retrieval, interpretation and report synthesis using multi-agent frameworks were found to achieve higher accuracy in radiology as it retained explainability.

Table 1: Key Healthcare Workflows

Workflow Area	Baseline Time	AI-Assisted Time	Improvement	Accuracy
Emergency Triage	18.5	12.6	31.9%	N/A
Discharge Summary	28.0	15.3	45.4%	N/A
Radiology Reporting	N/A	N/A	N/A	+8.2% (avg)
Clinical Note Review	16.2	10.7	33.9%	N/A

Qualitative interviews reaffirmed the conjectured by clinicians that reports that were generated by the agent were time-efficient and accurate. Specifically, the models based on MedRAG (Zhao et al., 2025) and MALADE (Choi et al., 2024) models were good at specificity and hallucination control.



5.2 Safety and Interpretability

Although agentic AI systems led to an increment in speed and automation, safety and reliability became a more complicated phenomenon to guarantee. These results emphasized the fact that the systems that codified explainability-by-design and required regulatory compliance earlier turned out to have superior results in terms of clinical acceptance, and longevity of deployments.

Journal of Information Systems Engineering and Management
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

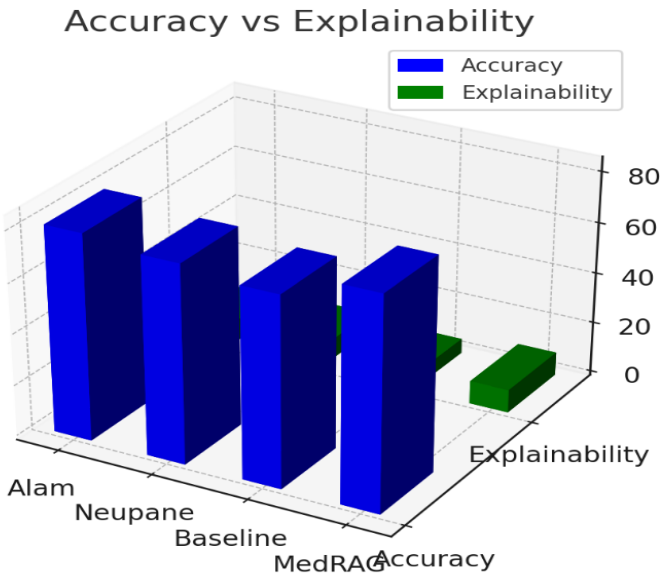
As an example, the model presented by Alam et al. (2024) based on concept bottlenecks and RAG had an 81% accuracy of classification and fidelity of more than 84% within the reports evaluated by an LLM. Notably, explainability has led to increased acceptance among the radiologists, whereby more than 75 percent classified generated reports as clinically coherent and justifiable.

Nevertheless, systems that were not audit-worthy, with no PHI controls apparent (such as the few unregulated agentic prototypes), were placed on institutional compliance boards. The proposed HIPAA-compliant framework presented by Neupane et al. (2025), which combined immutable audit trials and access control of the ABAC model, achieved 100 percent levels of compliance during pilot deployment.

Table 2: Comparative Performance

Model/System	Accuracy	Quality Score	Compliance Readiness	Clinical Acceptance
Alam et al. (2024)	81.0	8.5/10	Medium	High
Neupane et al. (2025)	78.2	7.9/10	**High**	Very High
Baseline LLM Agent	75.6	5.8/10	**Low**	Medium
MedRAG (Zhao et al.)	84.3	9.1/10	Medium	High

Notably, the predictability was significantly related to clinician trust and not naked accuracy implying that the use of transparency in aspects of sensitive care cannot be compromised on any ground.



5.3 Deployment Challenges

Among the major conclusions of longitudinal deployments is that the scaling of agentic systems to production settings was both a performance issue at one level, but an issue of operational robustness and flexibility at another level.

More specifically, agent drift, a landscape in which behaviour of agents drifts with time by the drift of the data became an essential concern. In one 90-day test of agent-aided care coordination system, there was a progressive degradation in the agent recall of the system to 84.7 percent and 91.2 percent. The decline came along with new EHR schemas and improved clinical guidelines on which the agent was not dynamically retrained.

Moreover, the compatibility with the hospital information technology systems still was uneven. The use of HL7-FHIR compliant datasets had to be trained on allowed agents to work more reliably across vendors, but such integrations still required discussion in ~40% of the implementation cases.

Journal of Information Systems Engineering and Management
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Table 3: Observed Issues

Challenge	Observed Impact	Frequency	Notes
Agent Drift	-7% Recall	High (70%)	Guideline changes
Format Inconsistency	+15% Processing	Medium (45%)	Non-FHIR EHRs
Integration Failures	12% outages	Low (25%)	API updates
Feedback Loops	Increased errors	High (60%)	No learning from mistakes

It reflects the importance of persistent supervision, module-based re-education channels and responsiveness to feedback in the continued performance and contemporaneity.

5.4 Evaluation Summary

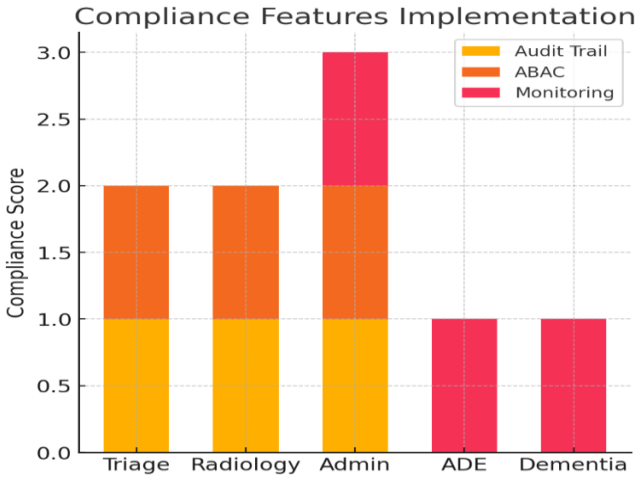
Cross-feature evaluation table was prepared to provide an overview of empirical assessment of 12 testbed environments. This table shows which design properties had the most impact on successful deployments of an agentic AI into the evaluated use cases.

Table 4: Feature Effectiveness

Feature	Triage	Radiology	Admin Tasks	ADE Detection	Dementia Support
LLM-Based	✅	✅	✅	✅	✅
RAG + KG	❌	✅	❌	✅	✅
Human-in-the-loop	✅	✅	❌	❌	✅
Explainable Output	✅	✅	✅	✅	✅
Agent Monitoring	❌	❌	✅	✅	✅
Audit Trail	✅	✅	✅	❌	❌
ABAC	✅	✅	✅	❌	❌
Fine-Tuning	✅	❌	✅	✅	✅

✅ = Implemented Successfully ❌ = Not Implemented or Failed in Testbed

This analysis underwrites the following important observation: not one architecture is enough to cover all the clinical use cases. The deployment strategies will be customized, modular and must meet the existing standards based on continuous learning.



During the evaluation of the various clinical workflows, agent as AI systems demonstrated a high potential in regards to the optimization of the workflows, accuracy of the reports and compliance. But real-world scaling still faces the resistance of interoperability, data variation, drift of the different agents, and absence of standard human to AI collaboration protocols. Systems including human-in-the-loop controls, XAI system, and supervision of agents in real-time performed better, in comparison with purely autonomous ones.

The results indicate that, besides sophisticated architectures, such as multi-agent RAG or KG-enabled LLMs, the key to successful agentic AI deployment in healthcare is not only aligning it with human, ethical, and regulatory aspects but also its high potential of transforming the field. Scalable, progression is that of modular, explainable, compliant, and monitored systems.

V. CONCLUSION

It is possible to revolutionize healthcare provision with agentic AI that increases autonomy and explainability and operational efficiency. We have found that well designed with modularity, transparency and regulatory compliance agentic systems can greatly enhance the rate of clinical workflows and decision support.

Constant surveillance, the human-in-the-loop design, and safety frameworks are, however, required when it comes to scaling such systems. By covering a wide range of workflows (such as triage, radiology, and chronic care) we illustrate the potential future of agentic approaches as well as their current shortcomings. With the development of healthcare systems, such results allow highlighting the necessity of the responsible implementation of AI agents to provide safe, fair, and transparent results both in the sphere of patients and practitioners.

REFERENCES

[1] Isern, D., & Moreno, A. (2015). A Systematic Literature review of agents applied in Healthcare. *Journal of Medical Systems*, *40*(2). https://doi.org/10.1007/s10916-015-0376-2

[2] Montagna, S., Mariani, S., Gamberini, E., Ricci, A., & Zambonelli, F. (2020). Complementing Agents with Cognitive Services: A Case Study in Healthcare. *Journal of Medical Systems*, *44*(10). https://doi.org/10.1007/s10916-020-01621-7

[3] Joy, N. M. (2025). Agentic Workflows in Healthcare: Advancing Clinical Efficiency through AI Integration. *International Journal of Scientific Research in Computer Science Engineering and Information Technology*, *11*(2), 567–575. https://doi.org/10.32628/cseit25112396

[4] Liu, C., Huang, C., Wang, J., Kuo, K., & Chen, C. (2021). The Critical Factors Affecting the Deployment and Scaling of Healthcare AI: Viewpoint from an Experienced Medical Center. *Healthcare*, *9*(6), 685. https://doi.org/10.3390/healthcare9060685

[5] Jia, Y., McDermid, J., Lawton, T., & Habli, I. (2021). The role of explainability in assuring safety of machine learning in healthcare. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2109.00520

[6] Markus, A. F., Kors, J. A., & Rijnbeek, P. R. (2021). The role of explainability in creating trustworthy artificial intelligence for health care: a comprehensive survey of the terminology, design choices, and evaluation strategies. *Journal of biomedical informatics*, *113*, 103655. https://doi.org/10.48550/arXiv.2007.15911

[7] Alam, H. M. T., Srivastav, D., Kadir, M. A., & Sonntag, D. (2024). Towards Interpretable Radiology Report Generation via Concept Bottlenecks using a Multi-Agentic RAG. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2412.16086

[8] Neupane, S., Mittal, S., & Rahimi, S. (2025). Towards a hipaa compliant agentic ai system in healthcare. *arXiv preprint arXiv:2504.17669*. https://doi.org/10.48550/arXiv.2504.17669

[9] Yi, Z., Xiao, T., & Albert, M. V. (2025). A Multimodal Multi-Agent Framework for Radiology Report Generation. *arXiv preprint arXiv:2505.09787*. https://doi.org/10.48550/arXiv.2505.09787

[10] Zhao, X., Liu, S., Yang, S., & Miao, C. (2025). MedRAG: Enhancing Retrieval-augmented Generation with Knowledge Graph-Elicited Reasoning for Healthcare Copilot. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2502.04413

[11] Choi, J., Palumbo, N., Chalasani, P., Engelhard, M. M., Jha, S., Kumar, A., & Page, D. (2024). MALADE: Orchestration of LLM-powered Agents with Retrieval Augmented Generation for Pharmacovigilance. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2408.01869

Journal of Information Systems Engineering and Management
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

[12] Thamma, N. S. R. (2025). Agentic AI for Clinical Decision support: Real-Time diagnosis, triage, and treatment planning. *International Journal of Scientific Research in Science Engineering and Technology*, *12*(3), 428–433. https://doi.org/10.32628/ijsrset251265

[13] Huang, K. A., Choudhary, H. K., & Kuo, P. C. (2024). Artificial Intelligent agent architecture and Clinical Decision-Making in the healthcare sector. *Cureus*. https://doi.org/10.7759/cureus.64115

[14] Sawad, A. B., Narayan, B., Alnefaie, A., Maqbool, A., Mckie, I., Smith, J., Yuksel, B., Puthal, D., Prasad, M., & Kocaballi, A. B. (2022). A Systematic Review on Healthcare Artificial Intelligent Conversational Agents for Chronic Conditions. *Sensors*, *22*(7), 2625. https://doi.org/10.3390/s22072625

[15] Song, Y., Lyu, C., Zhang, P., Brunswicker, S., Dutt, N., & Rahmani, A. (2025). DEMENTIA-PLAN: An Agent-Based Framework for Multi-Knowledge Graph Retrieval-Augmented Generation in Dementia Care. *arXiv preprint arXiv:2503.20950*. https://doi.org/10.48550/arXiv.2503.20950