International Journal of Advanced Research and Multidisciplinary Trends (IJARMT)
An International Open Access, Peer-Reviewed Refereed Journal
Impact Factor: 7.2 Website: https://ijarmt.com ISSN No.: 3048-9458

Evidence-Grounded and Explainable Agentic Rag for Reliable Decision Support in Indian E-Governance

Devesh Agarwal
Scientist D and District Informatics Officer
National Informatics Centre, Ministry of Electronics and Information Technology,District Centre, Churu, Rajasthan, India

Abstract
Government departments in India now publish a great deal of information online: schemes, eligibility rules, required documents, laws and administrative procedures. Citizens need answers they can trust and check, and AI tools are increasingly used to provide them. This study compares three ways of answering such queries: a standard large language model (LLM), conventional retrieval-augmented generation (RAG), and an evidence-grounded, explainable Agentic RAG framework. We tested all three on 250 Indian e-governance queries spread over five areas: government schemes and benefits, citizen services and documentation, laws and regulations, public welfare and eligibility, and administrative procedures. Answers were scored on accuracy, relevance of evidence, source attribution, completeness of context, hallucination, traceability, explanation and overall reliability. The Agentic RAG framework came out ahead of both alternatives, with 94.6% answer accuracy, 93.8% evidence relevance, 96.2% source attribution and 94.1% overall reliability. It also scored 93.4% on evidence traceability, 94.3% on evidence-to-claim consistency, 95.8% on hallucination-free responses and 94.0% on overall hallucination mitigation. Taken together, the results suggest that retrieving evidence in several passes, checking it, attributing sources and validating the final answer makes AI-generated decision support for e-governance more accurate, more transparent and easier to audit.
Keywords: Agentic RAG; E-Governance; Explainable AI; Hallucination Mitigation; Evidence Grounding; Decision Support

INTRODUCTION

Public administration is moving online, and with it the need for systems that give citizens correct, current and dependable information. Indian e-governance portals handle a wide mix of material, including government schemes, citizen services, eligibility conditions, laws, regulations and administrative procedures. Large language models (LLMs) could make this material easier to reach through plain conversation. The difficulty is that an LLM sometimes gives answers that are wrong or invented, particularly when the question depends on recent changes, specialised domain knowledge or policy detail. Wrong answers erode public trust, and they carry real risk when people act on them.

Retrieval-Augmented Generation (RAG) is one response to this problem, since it fetches relevant external evidence before the model writes its answer. Conventional RAG still has weaknesses, though. It can bring back irrelevant passages, fail to check whether the evidence actually fits, or leave claims without a source. Agentic RAG goes further: the system can

International Journal of Advanced Research and Multidisciplinary Trends (IJARMT)
An International Open Access, Peer-Reviewed Refereed Journal
Impact Factor: 7.2 Website: https://ijarmt.com ISSN No.: 3048-9458

retrieve more than once, assess what it has found, verify sources and validate the response before returning it. Add explainability and explicit evidence grounding, and AI-supported e-governance services should become more precise, easier to trace and easier to trust.

This paper proposes an Evidence-Grounded and Explainable Agentic RAG framework and tests it on Indian e-governance decision support, against conventional RAG and a standard LLM. We look at three things in particular: accuracy, traceability and hallucination mitigation. We also examine evidence relevance, source attribution, explainable reasoning and reliability across the main categories of e-governance query. The aim is to find out whether an agentic architecture that verifies its evidence can give citizens answers that are more reliable, more transparent and open to audit.

1. LITERATURE REVIEW

Pandey and Singh (2026) worked on explainable rating prediction and sentiment-segmented summarisation for e-governance applications. By combining rating prediction with sentiment-based summaries, they made automated feedback analysis easier to interpret. Their study shows why explainable output matters when officials need to understand user feedback and make decisions openly, and it adds to the growing use of explainable AI in citizen-facing digital governance.

Papageorgiou et al. (2025) evaluated the faithfulness of agentic RAG systems using LLM-based judging frameworks. They asked whether generated answers stay true to the evidence that was retrieved, and how far agentic RAG outputs can be trusted. Faithfulness matters a great deal in AI-assisted governance because unsupported and misleading responses do real harm. Their work is a useful baseline for judging hallucination reduction and evidence alignment in agentic RAG.

Chennupati et al. (2026) combined graph-based orchestration with emotion-aware LLM reasoning to explain customer feedback analysis. Their results show that pairing explainability with structured LLM reasoning gives analytical output that is more meaningful and easier to interpret. The application is customer feedback, not public services, but the same ideas of explainability and structured reasoning carry over to explainable decision-support systems.

2. RESEARCH METHODOLOGY

The study uses a structured comparative design to examine how well the proposed Agentic RAG framework supports reliable, explainable e-governance decisions. This section covers the research design, the query dataset, the approaches compared, the evaluation parameters, the evaluation procedure and the data analysis.

3.1 Research Design

We ran a comparative experiment to see how well an evidence-grounded, explainable Agentic Retrieval-Augmented Generation (RAG) framework supports decisions in Indian e-governance. The proposed framework was set against two baselines: a Standard Large Language Model (LLM) and Conventional RAG. The comparison covered answer accuracy, evidence relevance, source attribution, contextual completeness, hallucination mitigation, traceability, explainability and reliability.

International Journal of Advanced Research and Multidisciplinary Trends (IJARMT)
An International Open Access, Peer-Reviewed Refereed Journal
Impact Factor: 7.2 Website: https://ijarmt.com ISSN No.: 3048-9458

3.2 Dataset and Query Categories

In all, 250 representative Indian e-governance queries were evaluated. They fall into five categories: Government Schemes and Benefits (60), Citizen Services and Documentation (50), Laws, Rules and Regulations (45), Public Welfare and Eligibility (50) and Administrative Procedures (45). The queries were written to resemble the questions citizens typically ask when they need e-governance information to make a decision.

3.3 Comparative Approaches

All three approaches answered the same set of queries. The Standard LLM responded with no external retrieval. Conventional RAG retrieved material relevant to the question and then generated an answer from it. The proposed Agentic RAG framework selected evidence iteratively and verified it, attributed sources and validated the response, so the retrieved evidence was assessed before the final answer was written.

3.4 Evaluation Parameters

The evaluation indicators were fixed in advance and applied to all three approaches. Accuracy meant that a response was correct. Evidence relevance meant that the retrieved evidence was relevant to the question. Source attribution measured how well claims could be tied to a specific source, and contextual completeness checked whether a response covered enough ground to answer the question properly.

Hallucination mitigation was judged on four points: whether responses were free of hallucination, whether evidence was correctly aligned with claims, whether unsupported claims were avoided, and whether missing evidence was handled correctly. Explainability and traceability were judged on whether sources could be identified, whether direct supporting evidence existed, how complete the citations were, whether evidence matched claims, and whether the reasoning or justification was clear.

3.5 Evaluation Procedure

Every approach was run on each query under the same conditions, and each response was checked against the evidence and the predefined criteria. For the proposed Agentic RAG framework, the retrieved evidence was assessed for relevance and consistency before it went into the final answer. Cases of insufficient evidence or unsupported claims were recorded when hallucination mitigation was scored.

3.6 Data Analysis

The aggregated results were converted to percentages so the approaches could be compared directly. They are organised under four headings: comparative decision-support performance, hallucination mitigation, evidence traceability and explainability, and reliability by type of e-governance query. Tables and figures show how the proposed Agentic RAG framework compares with the other two approaches.

3. RESULTS

The results are reported for accuracy, hallucination mitigation, evidence traceability, explainability and reliability. The tables and figures below give the comparison for the proposed Agentic RAG framework. Table 1 compares the Standard LLM, Conventional RAG

International Journal of Advanced Research and Multidisciplinary Trends (IJARMT)
An International Open Access, Peer-Reviewed Refereed Journal
Impact Factor: 7.2 Website: https://ijarmt.com ISSN No.: 3048-9458

and the proposed Agentic RAG for e-governance decision support on five criteria: answer accuracy, evidence relevance, source attribution, contextual completeness and overall reliability. Figure 1 shows the same comparison graphically.

Table 1: Comparative Performance of Different Approaches for E-Governance Decision Support

Performance Metric	Standard LLM (%)	Conventional RAG (%)	Proposed Agentic RAG (%)
Answer Accuracy	78.4	86.7	94.6
Evidence Relevance	72.8	84.5	93.8
Source Attribution	61.5	79.6	96.2
Contextual Completeness	74.3	85.1	92.7
Overall Reliability	71.9	83.9	94.1

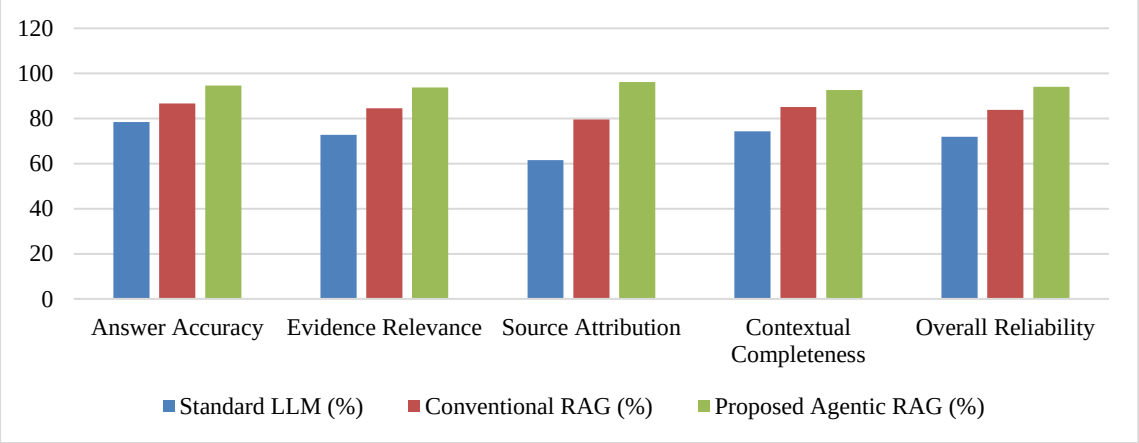


Figure 1: Graphical Representation of Comparative Performance of Different Approaches for E-Governance Decision Support

The proposed Agentic RAG scored highest on every parameter. Its overall accuracy reached 94.6%, against 86.7% for Conventional RAG and 78.4% for the Standard LLM. Evidence relevance was 93.8%. Source attribution improved the most, reaching 96.2%. Contextual completeness was 92.7% and overall reliability 94.1%. In short, the Agentic RAG gave more accurate, evidence-based decision support than the other methods, and its answers were easier to trace and more dependable.

Table 2 compares hallucination mitigation across the Standard LLM, Conventional RAG and Agentic RAG. It covers hallucination-free responses, evidence alignment, avoidance of unsupported claims, handling of missing evidence and overall hallucination mitigation. Figure 2 shows the same comparison graphically.

Table 2: Hallucination Mitigation Performance Across Different Approaches

Hallucination-Related Metric	Standard LLM (%)	Conventional RAG (%)	Proposed Agentic RAG (%)

International Journal of Advanced Research and Multidisciplinary Trends (IJARMT)
An International Open Access, Peer-Reviewed Refereed Journal
Impact Factor: 7.2 Website: https://ijarmt.com ISSN No.: 3048-9458

Hallucination-Free Responses	74.6	87.3	95.8
Correct Evidence Alignment	68.9	84.6	94.7
Unsupported Claim Avoidance	70.5	82.8	93.9
Correct Handling of Missing Evidence	63.7	78.4	91.6
Overall Hallucination Mitigation	69.4	83.3	94.0

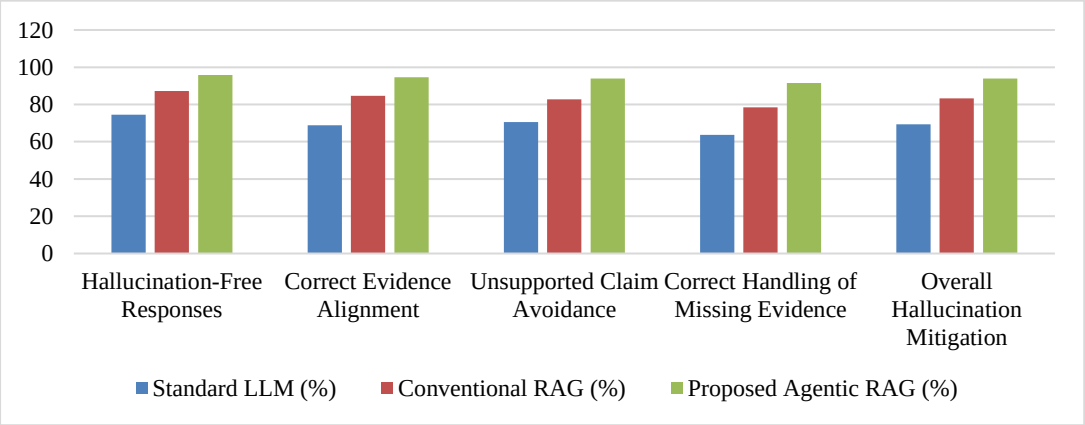


Figure 2: Graphical Representation of Hallucination Mitigation Performance Across Different Approaches

Agentic RAG did better than the other two approaches on every hallucination measure. Hallucination-free responses, correct evidence alignment and avoidance of unsupported claims came to 95.8%, 94.7% and 93.9%. Missing evidence was handled correctly in 91.6% of cases, and overall hallucination mitigation was 94.0%. Conventional RAG and the Standard LLM scored lower on every indicator. The proposed Agentic RAG therefore produced less unsupported information and kept its responses grounded in evidence.

Table 3 compares evidence traceability and explainability for the Standard LLM, Conventional RAG and Agentic RAG. The assessment looks at whether claims have identifiable sources, whether direct evidence supports the answer, how complete the citations are, whether evidence is consistent with claims, whether the reasoning can be explained, and overall traceability. Figure 3 shows the comparison graphically.

Table 3: Evidence Traceability and Explainability of Generated Responses

Traceability and Explainability Indicator	Standard LLM (%)	Conventional RAG (%)	Proposed Agentic RAG (%)
Identifiable Source for Claims	60.8	80.2	95.6
Direct Evidence Supporting Answer	58.9	78.7	93.5

International Journal of Advanced Research and Multidisciplinary Trends (IJARMT)
An International Open Access, Peer-Reviewed Refereed Journal
Impact Factor: 7.2 Website: https://ijarmt.com ISSN No.: 3048-9458

Citation Completeness	55.7	76.9	92.8
Evidence-to-Claim Consistency	63.4	82.5	94.3
Explainable Reasoning/Justification	57.6	75.8	90.7
Overall Traceability	59.3	78.8	93.4

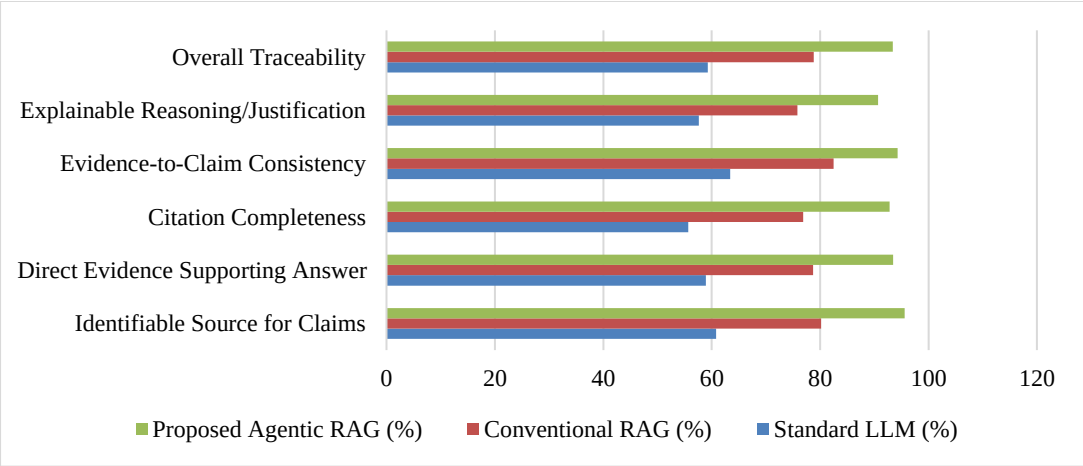


Figure 3: Graphical Representation of Evidence Traceability and Explainability of Generated Responses

Agentic RAG led on every traceability and explainability measure. Of its claims, 95.6% could be traced to an identified source, 93.5% had direct supporting evidence and 92.8% had complete citations. Evidence-to-claim consistency was 94.3% and explainable reasoning or justification was 90.7%. Overall traceability was 93.4%, well above Conventional RAG (78.8%) and the Standard LLM (59.3%). The responses were more transparent, easier to verify and easier to audit.

Table 4 shows how reliable the proposed Agentic RAG framework is across the main categories of Indian e-governance query: government schemes and benefits, citizen services and documentation, laws and regulations, public welfare and eligibility, and administrative procedures. Figure 4 shows the same results graphically.

Table 4: Reliability of Agentic RAG Across Major Indian E-Governance Query Categories

E-Governance Query Category	Number of Evaluated Queries	Accurate Responses (%)	Evidence-Verified Responses (%)	Hallucination-Free Responses (%)
Government Schemes & Benefits	60	95.0	94.2	95.8
Citizen Services & Documentation	50	94.0	93.6	94.0

International Journal of Advanced Research and Multidisciplinary Trends (IJARMT)
An International Open Access, Peer-Reviewed Refereed Journal
Impact Factor: 7.2 Website: https://ijarmt.com ISSN No.: 3048-9458

Laws, Rules & Regulations	45	93.3	92.9	93.8
Public Welfare & Eligibility	50	96.0	95.1	96.4
Administrative Procedures	45	94.4	93.8	95.0
Overall	250	94.6	93.9	95.0

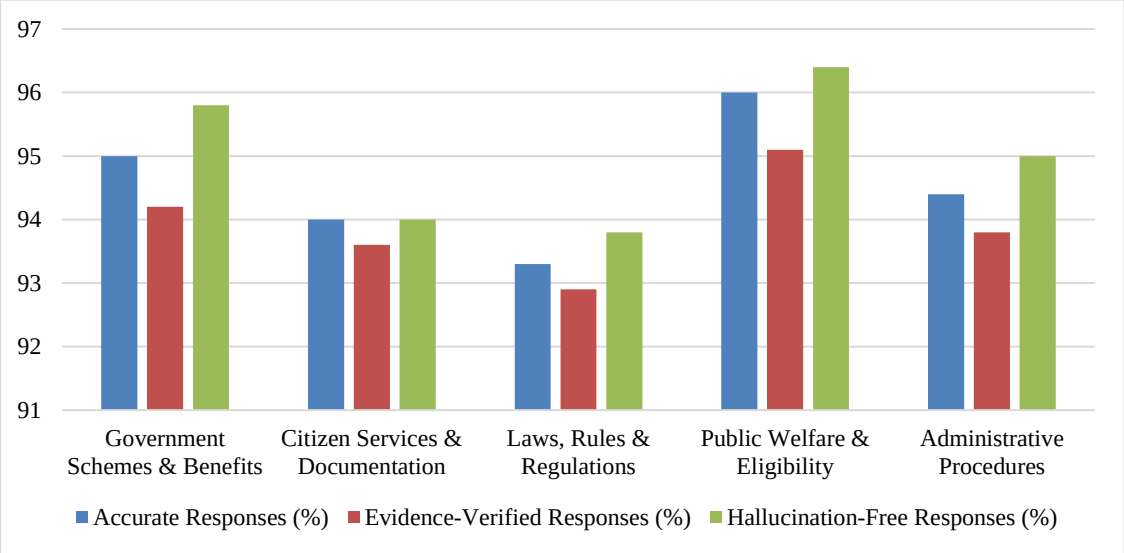


Figure 4: Graphical Representation of Reliability of Agentic RAG Across Major Indian E-Governance Query Categories

Performance was consistently high across every category. Accuracy was highest for Public Welfare and Eligibility (96.0%) and Government Schemes and Benefits (95.0%). Evidence-verified responses ranged from 92.9% to 95.1%, and hallucination-free responses from 93.8% to 96.4%. Overall, accuracy was 94.6%, evidence verification 93.9% and hallucination-free responses 95.0%, so the framework held up across the varied information needs of Indian e-governance.

4. DISCUSSION

The proposed Evidence-Grounded and Explainable Agentic RAG model was more accurate, retrieved more relevant evidence, attributed sources better and was more reliable than both the Standard LLM and Conventional RAG. With 94.6% accuracy, 93.8% evidence relevance and 96.2% source attribution, the results indicate that agentic retrieval and evidence verification make e-governance decision support more reliable and more transparent.

The framework also produced 95.8% hallucination-free responses and 94.0% overall hallucination mitigation, which shows it was effective at cutting down unsupported information. Its overall traceability of 93.4% and evidence-to-claim consistency of 94.3% point the same way, towards better explainability and auditability.

Performance stayed high in every e-governance category: overall accuracy was 94.6%, evidence verification 93.9% and hallucination-free responses 95.0%. Taken together, the results support Agentic RAG as a way to give reliable, evidence-grounded and explainable decision support in Indian e-governance.

5. CONCLUSION

The proposed Evidence-Grounded and Explainable Agentic RAG framework outperformed the Standard LLM and Conventional RAG as a source of reliable decision support in Indian e-governance. Its responses were more accurate, better supported by evidence and more dependable: 94.6% accuracy, 93.8% evidence relevance, 96.2% source attribution and 94.1% overall reliability. It reached 95.8% on hallucination-free responses and 94.0% on overall hallucination mitigation, so unsupported information was reduced. Overall traceability of 93.4% and evidence-to-claim consistency of 94.3% show better transparency, explainability and auditability. Accuracy was high in every e-governance category, with overall accuracy of 94.6%, evidence verification of 93.9% and hallucination-free responses of 95.0%. These results support combining agentic retrieval, evidence verification, source attribution and response validation to build AI decision-support systems for Indian e-governance that are accurate, transparent, reliable and explainable.

REFERENCES

1. Pandey, P., & Singh, J. P. (2026). Explainable Rating Prediction and Sentiment-Segmented Summarization for E-Governance Applications. SN Computer Science, 7(7), 715.
2. Papageorgiou, G., Sarlis, V., Maragoudakis, M., Magnisalis, I., & Tjortjis, C. (2025). Evaluating faithfulness in agentic RAG systems for e-governance applications using LLM-based judging frameworks. Big Data and Cognitive Computing, 9(12), 309.
3. Chennupati, P. K., Varanasi, V. S. V., Fakruddin, S., Yarlagadda, E. A., Devi, J. R., & Manasa, B. (2026, June). Explainable Customer Feedback Analysis Using Graph-Based Orchestration and Emotion-Aware LLM Reasoning. In 2026 5th OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 5.0 (pp. 1-6). IEEE.
4. Tavanxhiu, T., Godolja, M., Sevrani, K., & Naco, M. (2026). A technological blueprint for smart and AI-driven hospitality in emerging tourism markets: Evidence from Albania. Systems, 14(2), 188.
5. Hamza, A., Rizwan, M. B., Kayani, M., & Sheikh, S. A. A. (2026). Legalyze: An Open-Source Structure-Aware Agentic RAG Platform for High-Precision Legal Intelligence. Expert Systems, 43(10), e70415.
6. Ganesan, C. (2026). A Policy-Aligned Agentic RAG Framework for Risk-Aware Decision Support in Enterprise Customer Relationship Management. Journal of Intelligent Decision Making and Information Science, 3(8s), 1409-1443.
7. Prabhu, R. S., Manikandan, S. P., Poudel, A., Maity, D., & Kumar, M. S. (2026). Generative AI Technologies for Intelligent Decision-Making Applications.

International Journal of Artificial Intelligence and Machine Learning, 6(10s), 1530-1552.

8. Raj, A., Hareesh, B., Ramannavar, M., & Ankalkoti, P. (2026). Agent Breast-RAG: An Agentic Retrieval-Augmented Explainable Large Language Model Framework for Multimodal Breast Cancer Diagnosis and Clinical Decision Support. Computer and Decision Making: An International Journal, 3(3S), 01-13.
9. Wang, Z., Li, J., Guo, X., & Tsuchiya, T. (2026). From noisy feedback to evidence-aware issue specifications: an agent-governed retrieval-augmented generation approach. Automated Software Engineering, 33(3), 103.
10. Satpute, B., Rahane, W. P., Pawar, P., Vanjari, H., Kulkarni, R., Parhad, S. V., & Deshmukh, P. V. (2026). Multi agent retrieval validation and knowledge reasoning for enhanced retrieval augmented generation. Scientific Reports.
11. Gakhar, S., & Singh, R. K. (2026). Leveraging large language models (LLMs) for GeoAI-enabled digital agro-advisory. Frontiers in Remote Sensing, 7, 1839369.
12. Patil, S. S. (2026). Explainable Artificial Intelligence In Ayurvedic Clinical Decision Support For Prameha: Opportunities, Challenges And Future Directions. European Journal of Prosthodontics & Restorative Dentistry, 34, 148.
13. Arafin, P., Moniruzzaman, M., & Alkhateeb, A. (2026). Evidence-Grounded Clinical Pharmacogenomics Decision Support: Integrating Fine-Tuned Large Language Models with Hybrid Retrieval-Augmented Generation. Applied Sciences, 16(18), 9148.
14. Lakhani, T. L. (2026). From Static Recommendations to Autonomous Decisions: A Framework for Agentic Recommendation Systems with Memory, Reasoning and Tool Use. International Journal of Sustainability and Innovation in Engineering, 4(1).
15. Sablatnig, R. (2025). Retrieval-Augmented Generation for Enterprise Security Intelligence in Cloud-Based Decision Support and Knowledge Systems. International Journal of Research Publications in Engineering, Technology and Management (IJRPETM), 8(6), 13480-13488.
16. Wu, K. (2026). Trust-Calibrated Enterprise Service Copilots: Cost-Aware Agentic RAG for Customer-Support Knowledge Management. Stout in Business and Management, 2(1), 01-
17. Kafle, D. S., Talhi, E., & Coustaty, M. (2026, August). RAGXDoc: Structured Knowledge-Guided Retrieval and Explainable Re-ranking. In Document Analysis and Recognition–ICDAR 2026: 20th International Conference, Vienna, Austria, August 31–September 2, 2026, Proceedings, Part III (p. 70). Springer Nature.
18. Alghamdi, M. M., & Alamri, W. A. (2026). IKEA: Intelligent Knowledge Extraction with Reasonable and Agentic AI Agents. Applied Sciences, 16(15), 7704.
19. Huang, Y., Chen, Y., Zhang, H., Li, K., Zhou, H., Fang, M., ... & Wang, J. (2025). Deep research agents: A systematic examination and roadmap. arXiv preprint arXiv:2506.18096.