



Architecting Intelligence: A Review of Artificial Intelligence's Role in the Design and Development of Information Systems

Ben-Izuagie, O. R.

Department of Computer Science, College of Education, Warri, Delta State, Nigeria

*Corresponding author email: michaelpaulps@gmail.com

Abstract

The integration of Artificial Intelligence (AI) into Information Technology, particularly within Information Systems (IS), represents a paradigm shift from systems that simply store and retrieve data to intelligent, adaptive, and predictive infrastructures. This paper reviews research published between 2019 and 2025 to examine AI's role in modern information systems design and development. We analyse 97 peer-reviewed articles drawn from systematic syntheses across accounting information systems, geographic information systems, digital assistant technologies, and general IS architecture research. Our findings identify four thematic clusters in the literature: (1) foundational AI technologies (machine learning, natural language processing, generative models, and computer vision) being integrated into IS architectures; (2) architectural transformations including intelligent automation, multi-agent systems, and retrieval-augmented generation frameworks; (3) governance evolution from traditional data governance to comprehensive AI governance frameworks addressing explainability, bias detection, and drift monitoring; and (4) implementation challenges spanning algorithmic bias, data governance complexity, technical preparedness gaps, and ethical deployment concerns. The review shows that AI-driven IS significantly enhance operational efficiency, data processing accuracy, and real-time decision-making, yet current research still leaves persistent challenges unaddressed. We conclude by proposing a strategic framework for designing resilient and ethical AI-augmented IS that synthesises insights from six pillars of AI governance architecture, emphasising the necessity of human oversight and interdisciplinary collaboration. This paper contributes a synthesised overview of the field, offering a roadmap for researchers and practitioners navigating the convergence of AI and information technology.

Keywords: Artificial Intelligence, Information Systems, Systematic Review, Machine Learning, System Architecture.

Introduction

The digital transformation era is characterised by the convergence of Information Technology (IT) and Artificial Intelligence (AI). Traditionally, information systems served as structured processes for data storage, retrieval, and management to support organisational operations. The advent of AI has fundamentally altered this dynamic, endowing these systems with the ability to learn from data, automate complex workflows, and generate predictive insights, thereby transforming them from passive repositories into active, autonomous agents in decision-making (Mohammed et al., 2025). This integration is not merely a technological upgrade but a strategic imperative across industries. From predictive maintenance in manufacturing and fraud detection in finance to personalised customer experiences in retail, AI-driven IS are becoming central to competitive advantage. Organizations employing AI-augmented systems report significant improvements in speed, accuracy, scalability, and the ability to process unstructured data—capabilities difficult to attain using traditional IS architectures (Krčadinac et al., 2025). However, the rapid adoption of these technologies raises critical questions about robustness, fairness, and governance. Algorithmic biases, data privacy concerns, and the "black box" nature of some AI models pose significant risks. Real-world failures illustrate these dangers: Amazon's hiring algorithm discriminated against women by learning from biased historical data; the UK's exam grading algorithm unfairly disadvantaged students from lower-performing schools; and the Dutch child benefits scandal saw thousands of families wrongfully penalized through algorithmic racial profiling, ultimately leading to the government's resignation (GoodData, 2025). These cases demonstrate how poorly governed AI erodes trust and

destroys value faster than it creates. Despite a growing body of literature, there is a need for consolidated understanding of how AI is being integrated into the core fabric of IT systems and what governance mechanisms are emerging to address associated risks. This paper addresses this gap by conducting a systematic review of recent literature. We aim to answer the following research questions:

1. What are the primary AI technologies being integrated into contemporary information systems across different application domains?
2. How do these technologies reshape information systems design, architecture, and development processes?
3. What governance frameworks, ethical safeguards, and implementation challenges characterize the current landscape of AI-driven information systems?

The structure

The study focused on peer-reviewed journal articles and conference proceedings published in English between January 2019 and December 2025 to capture the most recent developments following the widespread adoption of transformer-based architectures and generative AI. Following the approach of Hastuti and Widuri (2025), who analysed 97 international scientific articles from the Scopus database, we prioritised studies employing systematic review methodologies, empirical investigations, and theoretical frameworks that made clear contributions to understanding AI-IS integration. Initial searches yielded over 300 papers. After screening titles and abstracts for relevance to the research questions and applying quality assessment criteria, we selected a final corpus of 97 articles for full-text analysis, building upon and synthesising existing reviews in the field (Krčadinac et al., 2025; Mohammed et al., 2025; Preiß & Westner, 2025). Screening was conducted in two stages: title/abstract screening followed by full-text assessment against the inclusion criteria above. Quality was assessed using a checklist adapted from standard review guidance, evaluating each study's methodological transparency, sample/corpus size, and peer-review status. Screening was conducted by the author; a second reviewer or inter-rater step was not used. A thematic analysis of the selected papers using an iterative coding process was conducted. Key findings were extracted and categorised into themes related to AI techniques, system design impacts, application domains, governance approaches, and implementation challenges. We employed cross-disciplinary synthesis to identify patterns across different subfields, including accounting information systems (Hastuti & Widuri, 2025), geographic information systems (Hayati, 2024), digital assistant technologies (Preiß & Westner, 2025), and general IS architecture research (De Luzi et al., 2025; Kulkarni & Ramanathan, 2025).

Literature Review

The literature reveals a rapidly maturing field characterised by increasing specialisation, methodological diversity, and growing attention to governance and ethical dimensions. We organise our review around four thematic clusters that emerged from the analysis.

Core AI Technologies Reshaping Information Systems

The literature consistently highlights four primary families of AI technologies driving change in information systems. Machine Learning and Deep Learning. Machine learning algorithms, particularly deep learning architectures, form the computational foundation of modern intelligent systems. Krčadinac et al. (2025) identify ML as enabling systems to identify patterns, support predictive analytics (e.g., forecasting demand, predicting system failures), and power recommendation engines. In accounting information systems, ML enhances automation and analytical capabilities while internal controls remain crucial for ensuring reliability and accountability (Hastuti & Widuri, 2025). Roy et al. (as cited in Hayati, 2024) demonstrate how ML empowers geographic information systems by advancing spatial analysis and decision-making capabilities across environmental monitoring, disaster management, and urban planning applications.

Natural Language Processing. NLP technologies are revolutionizing user-system interaction and information extraction from unstructured data. They are central to the development of intelligent chatbots, voice-activated assistants, and systems capable of extracting meaning from vast amounts of text, thereby enhancing information analysis processes (Fakhlina et al., 2025). The systematic review of digital assistant technologies by Preiß & Westner, (2025) identifies five main areas of impact: productivity and efficiency, business development, resource optimization, quality enhancement, and the promotion of learning and creativity. These technologies, including intelligent agents,

chatbots, and voice assistants, hold significant transformative potential for large enterprises by automating routine tasks and enriching user interactions.

Generative AI and Large Language Models. More recent literature points to the transformative role of generative models, particularly large language models, in IS engineering. De Luzi, Monti, and Mecella (2025) provide a comprehensive framework for exploiting LLMs in information systems engineering, covering prompt engineering, retrieval augmented generation, and multi-agent systems. Their work demonstrates how LLMs offer unprecedented opportunities to support IS design, development, maintenance, and reengineering across the entire system lifecycle. Technical advances in LLM applications extend to specific IS tasks. Gupta et al. (2025) propose a system that translates natural language descriptions into executable ETL workflows using a Classifier-Augmented Generation approach, demonstrating improved accuracy and efficiency while substantially reducing token usage. Their work represents the first system with detailed evaluation across stage prediction, edge layout, and property generation for natural-language-driven ETL authoring. Similarly, their Schema and NL-Aware In-context Learning framework addresses GraphQL query generation, showing up to 20% performance improvement for certain LLMs—a critical contribution as Gartner predicts over 60% of enterprises will use GraphQL in production by 2027 (Gupta et al., 2025). **Computer Vision and Multi-modal AI.** In cross-disciplinary applications, computer vision enhances spatial analysis capabilities within web-based geographic information systems. Hayati (2024) reviews how AI technologies including computer vision improve automation, scalability, and user interactivity across sectors such as environmental monitoring, smart agriculture, and public health. The integration of multiple AI modalities enables more sophisticated system capabilities that transcend traditional single-mode processing.

Architectural Transformations in Information Systems Design

The integration of AI technologies fundamentally reshapes how information systems are designed, developed, and deployed. **Intelligent Automation and Process Optimization.** AI-driven automation extends beyond robotic process automation to "intelligent automation," where systems not only execute predefined tasks but adapt and optimize processes in real-time (Mohammed et al., 2025). Krčadinac et al. (2025) examine how AI technologies enhance requirement engineering, system modeling, and process optimization across various stages of system development. This transformation enables systems to handle complex, dynamic environments with minimal human intervention.

Multi-Agent Systems and Copilot Architectures. The evolution from simple agents to sophisticated "copilots" represents a significant architectural shift. Preiß & Westner, (2025) trace this trajectory in their systematic review of digital assistant technologies, noting the transition from basic chatbots to generative AI-driven copilots embedded in widely adopted productivity platforms such as Microsoft 365 and Google Workspace. These platforms dominate the productivity software market: as of February 2024, Google Workspace held over 44% of global market share, and Microsoft 365 approximately 30%, representing over 345 million active paid users (Preiß & Westner, 2025).

The embedding of AI capabilities into these ubiquitous platforms signals a mainstreaming of intelligent system architectures. De Luzi et al. (2025) provide technical depth on multi-agent systems, describing how multiple LLM-based agents can collaborate to solve complex information system tasks. Their framework encompasses agent communication protocols, task decomposition strategies, and coordination mechanisms that enable scalable intelligent system design. **Retrieval-Augmented Generation and Knowledge Integration.** Retrieval-augmented generation (RAG) architectures represent an important innovation for grounding AI systems in authoritative knowledge bases. By combining information retrieval with generative capabilities, RAG enables systems to produce outputs that are both contextually appropriate and factually grounded (De Luzi et al., 2025). This architecture is particularly relevant for enterprise information systems where accuracy and traceability are paramount. **Ethical Architecture and Guardrail Frameworks.** An emerging thread in the literature addresses how ethical considerations can be embedded into system architecture rather than added as an afterthought. Kulkarni and Ramanathan (2025) introduce the Guardrail Framework (GRAILS), a modular, reusable software framework that operationalizes responsible AI principles by decoupling ethical constraints from functional requirements. The framework enables ethical filtering of information based on data sensitivity (Low to High), user trust score (Low to High), and data granularity (Cell, Row, Column, or Table). Implemented and evaluated using Open Government Dataset, the framework demonstrates high cohesion, low coupling, and long-term maintainability in line with fundamental software engineering architectural design principles.

Governance Evolution: From Data Governance to AI Governance

A critical finding across the literature is the recognition that traditional governance frameworks are inadequate for AI-driven systems. The Governance Deficit: Mohammed et al. (2025) identify that while AI-driven IS enhances scaling and responsiveness, adoption processes face challenges including lack of technical preparedness, regulatory compliance issues, and scepticism toward AI in decision-making. The governance deficit is empirically documented: while 81% of companies have AI use cases in production, only 15% rate their AI governance as very effective (GoodData, 2025). Between 70% and 85% of GenAI deployment efforts fail to meet desired ROI, primarily due to governance gaps rather than technical limitations (GoodData, 2025). This finding aligns with Batool et al.'s (2025) systematic review, which similarly identifies a persistent gap between AI governance principles and their operational enforcement across organisations.

Why Traditional Governance Fails. Traditional governance frameworks break down when applied to AI systems because they fundamentally violate core assumptions underlying data governance approaches (GoodData, 2025):

- **Static Asset Management Breaks Down:** Traditional governance catalogs relatively stable datasets with predictable schemas. Machine learning models continuously evolve through training cycles, online learning, and adaptation mechanisms, making traditional asset management approaches obsolete.
- **Human Decision Loops Cannot Scale:** Traditional governance expects humans to interpret data and make decisions with clear accountability chains. AI systems make thousands of decisions per second without human intervention, yet accountability requirements intensify rather than disappear.
- **Binary Quality Standards Fail:** Data quality operates against fixed criteria. AI predictions exist on probability distributions rather than binary scales, and traditional quality frameworks cannot assess whether probabilistic outputs constitute acceptable performance without sophisticated contextual analysis.

The Six Pillars of AI Governance Architecture. GoodData (2025) synthesizes enterprise implementations and regulatory analysis into a comprehensive framework comprising six interconnected pillars:

- **Accountability Framework:** Distributing accountability across technical, business, ethical, and operational roles, with the RACI-AI framework adding Adversarial and Impacted roles to ensure accountability connects to real-world outcomes.
- **Policy Infrastructure:** Operating at four hierarchical levels from board-approved principles to team-level work instructions, with policy automation using machine-readable formats like Open Policy Agent's Rego.
- **Risk Management and Compliance Operations:** Addressing technical risks (model degradation, adversarial attacks), operational risks (automation bias, skills gaps), ethical risks (discrimination, privacy violations), and strategic risks (regulatory sanctions, reputation damage).
- **AI-Ready Data Infrastructure:** Implementing domain-oriented decentralization through data products, federated governance computation, and data contracts specifying statistical properties, update frequencies, and quality guarantees.
- **Development and Deployment Governance:** Extending governance across the ML development lifecycle with specific gates from problem formulation through monitoring, using automated governance gates and testing pyramids.
- **Monitoring and Observability:** Spanning technical performance, statistical accuracy, business outcomes, and governance compliance, with advanced practices including canary deployments and shadow mode monitoring.

Implementation Challenges and Risk Dimensions

The literature documents persistent challenges that cut across application domains and system types.

Algorithmic Bias and Fairness. AI models trained on biased historical data can perpetuate and amplify societal inequalities. Mohammed et al. (2025) identify algorithmic bias as one of the most critical issues, noting that ensuring fairness in automated decisions remains a major hurdle. The Dutch child benefits scandal, where thousands of families were wrongfully penalised through algorithmic racial profiling, exemplifies the real-world consequences of inadequately addressed bias (GoodData, 2025).

Data Governance and Privacy. AI effectiveness is directly tied to data quality and quantity, creating pressure to collect more data and raising significant privacy concerns. Hastuti and Widuri (2025) emphasise the mediating role of information quality in linking AI, internal controls, and IS effectiveness, highlighting unresolved challenges in balancing innovation with robust governance. Organizations with mature data governance practices achieve more than twice the ROI compared to other companies (GoodData, 2025).

Explainability and the "Black Box" Problem. Many advanced AI models operate as "black boxes," making it difficult to understand why particular decisions were made. Krčadinac et al. (2025) identify explainable AI as an emerging trend essential for building trust in high-stakes domains like healthcare and finance. The lack of explainability creates barriers to adoption and complicates regulatory compliance.

Technical and Organizational Preparedness. Many organizations lack necessary technical expertise, infrastructure, and cultural readiness to successfully adopt AI systems. Mohammed et al. (2025) find that adoption processes are hindered by lack of technical preparedness, and that AI-generated outputs require human verification to ensure validity in specific contexts. The governance deficit is particularly acute: 47% of organizations have experienced at least one negative consequence from AI deployment (GoodData, 2025).

Cross-Domain Integration Challenges. Domain-specific reviews reveal additional challenges. In web-based geographic information systems, Hayati (2024) identifies limitations including ethical considerations, data standardization issues, and limited real-world implementation. In accounting information systems, Hastuti and Widuri (2025) emphasize the importance of balancing technological innovation with robust governance for managers, auditors, and policymakers.

Discussion

The findings paint a picture of a field in transition, moving from initial experimentation toward a focus on responsible and sustainable integration. Our synthesis of the literature reveals both convergence across domains and important domain-specific variations.

Thematic Integration

Four cross-cutting themes emerge from the literature with implications for theory and practice.

First, the technological foundation is consolidating around a core set of AI capabilities—machine learning, NLP, generative models, and computer vision—that are being adapted to diverse IS contexts (Krčadinac et al., 2025; Hayati, 2024; De Luzi et al., 2025). This consolidation enables knowledge transfer across domains while raising questions about appropriate customization for specific applications. Second, architectural patterns are emerging that embed intelligence throughout the system lifecycle rather than treating AI as an add-on component. Multi-agent systems, copilot architectures, and retrieval-augmented generation represent distinct patterns with different trade-offs (Preiß & Westner, 2025; De Luzi et al., 2025). The Guardrail Framework demonstrates how ethical constraints can be architected as separable, reusable components (Kulkarni & Ramanathan, 2025). Third, governance is recognized as fundamental rather than optional, with frameworks evolving from data governance to comprehensive AI governance (GoodData, 2025; Mohammed et al., 2025). The six pillars framework provides an implementable blueprint, while empirical evidence documents the costs of governance failures. Fourth, implementation challenges are persistent and multidimensional, spanning technical, organizational, ethical, and regulatory dimensions (Mohammed et al., 2025; Hastuti & Widuri, 2025; Hayati, 2024). The literature suggests these challenges cannot be addressed sequentially but require integrated approaches from system inception.

A Proposed Framework for Ai-Augmented Information Systems

Synthesizing across the literature, we propose a multi-layered framework for developing future AI-augmented information systems that integrates insights from governance research (GoodData, 2025), ethical architecture (Kulkarni & Ramanathan, 2025), and system design (Krčadinac et al., 2025; De Luzi et al., 2025):

1. **Strategic Alignment Layer:** AI initiatives must be aligned with clear organizational goals and embedded within comprehensive IT governance frameworks. This includes risk-based governance scaling that classifies AI applications by risk tier (low, medium, high, critical) and applies proportionate governance requirements (GoodData, 2025).

2. **Ethical Architecture Layer:** Ethical considerations, including fairness, accountability, and transparency, must be embedded into system architecture from the outset. The Guardrail Framework demonstrates how ethical constraints can be operationalized as modular, reusable components decoupled from functional requirements (Kulkarni & Ramanathan, 2025).
3. **Human-in-the-Loop Layer:** Critical decisions should maintain meaningful human involvement to validate AI recommendations and override potentially flawed automated decisions. This addresses automation bias and ensures accountability (Mohammed et al., 2025; GoodData, 2025).
4. **Data Infrastructure Layer:** Investment in data quality, data lineage, and secure data storage is prerequisite for successful AI initiatives. Data contracts specifying statistical properties, quality guarantees, and compliance requirements enable reliable AI deployment (GoodData, 2025).
5. **Development and Deployment Layer:** Automated governance gates integrated into CI/CD pipelines ensure consistent application of standards across the ML lifecycle, from problem formulation through monitoring (GoodData, 2025).
6. **Monitoring and Observability Layer:** Continuous monitoring across technical, statistical, business, and governance dimensions enables detection of drift, bias emergence, and performance degradation (GoodData, 2025).
7. **Interdisciplinary Collaboration Layer:** Building effective and responsible AI systems requires collaboration between computer scientists, domain experts, ethicists, and policymakers. The literature consistently emphasizes this as essential for addressing the multidimensional challenges of AI integration (Mohammed et al., 2025; Hastuti & Widuri, 2025; Hayati, 2024).

Limitations and Research Gaps

Our review reveals several important gaps in the current literature. Longitudinal studies of AI system impacts remain limited (Mohammed et al., 2025), making it difficult to assess sustained benefits and unintended consequences over time. Research on the strategic use of AI in organizations and its long-term implications for work, skills, and organizational structures is underdeveloped. The effectiveness of different governance mechanisms in practice requires more empirical investigation, particularly as regulatory landscapes evolve. Cross-domain comparative studies would help identify which challenges are universal and which are domain-specific. Finally, research on AI in small and medium enterprises is limited compared to large enterprise contexts, despite the importance of SMEs in most economies. This review is also subject to methodological limitations. As a single-reviewer synthesis, it carries a risk of selection and interpretation bias that would be mitigated by independent double-screening. The search was restricted to English-language publications, which may exclude relevant non-English research, and reliance on indexed databases may underrepresent grey literature, potentially introducing publication bias.

Conclusion

This narrative literature review confirms that AI is profoundly reshaping information technology, offering unprecedented opportunities for efficiency, intelligence, and personalization across diverse application domains. We have identified the core technologies driving this change—machine learning, natural language processing, generative AI, and computer vision—and traced their integration into information systems architectures ranging from accounting systems to geographic information platforms. The review has highlighted the critical governance challenges that must be addressed to ensure these systems are not only powerful but also trustworthy and equitable, synthesizing emerging frameworks for AI governance and ethical system design. The proposed multi-layered framework offers a starting point for researchers and practitioners to approach AI integration in a more holistic and responsible manner, integrating insights from governance research, ethical architecture, and system design.

Future research should focus on several key areas:

- **Longitudinal and Comparative Studies:** Conducting longitudinal studies to understand the long-term organizational impacts of AI-driven transformation (Mohammed et al., 2025) and comparative research across domains to identify universal versus context-specific challenges.
- **Governance Mechanism Effectiveness:** Developing and validating standardized benchmarks and methodologies for auditing AI systems for bias, fairness, and explainability, and empirically investigating the effectiveness of different governance mechanisms in practice (GoodData, 2025; Kulkarni & Ramanathan, 2025).

- Architectural Innovations: Investigating novel system architectures that can natively support explainability, transparency, and ethical constraints, building on emerging work in multi-agent systems and guardrail frameworks (De Luzi et al., 2025; Kulkarni & Ramanathan, 2025).
- Emerging Technology Implications: Exploring the implications of emerging trends like low-code AI platforms, federated learning, and edge AI on IS design and governance.
- SME and Public Sector Contexts: Expanding research to small and medium enterprises and public sector organizations, where resource constraints and regulatory requirements may differ significantly from large enterprise contexts.
- Human-AI Collaboration: Deepening understanding of how humans and AI systems can effectively collaborate, including studies of automation bias, trust calibration, and skill evolution in AI-augmented work.

As AI capabilities continue to advance and their integration into information systems deepens, the need for rigorous research that bridges technical, organizational, and ethical dimensions will only grow. This review provides a foundation for such research by synthesizing current knowledge and identifying priority directions for future inquiry.

References

- Batool, A., Zowghi, D., & Bano, M. (2025). AI governance: a systematic literature review. *AI and Ethics*, 5, 3265 - 3279. <https://doi.org/10.1007/s43681-024-00653-w>
- De Luzi, F., Monti, F., & Mecella, M. (Eds.). (2025). *Engineering information systems with large language models*. Springer.
- Fakhline, R. J., Rahmi, L., Zalmi, F. N. H., Fadzil, G. M., Yudisman, S. N., Hutagalung, M. W., & Trinova, R. (2025). Artificial intelligence as an information analysis process: A systematic literature review. *Record and Library Journal*, 11(2), 480–499. <https://doi.org/10.20473/rhj.V11-I2.2025.480-499>
- GoodData. (2025). From data governance to AI governance: The enterprise blueprint for responsible AI at scale [White paper].
- Gupta, N., Kesarwani, M., Ghosh, S., Mehta, S., Eberhardt, C., & Debrunner, D. (2025). Schema and natural language aware in-context learning for improved GraphQL query generation. In Proceedings of the 2025 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Vol. 3, Industry Track, pp. 1009–1015). Association for Computational Linguistics.
- Hastuti, D. A., & Widuri, R. (2025). Strengthening the effectiveness of accounting information systems: A systematic review of AI integration and internal control practices. *Edelweiss Applied Science and Technology*, 9(10), 427–438.
- Hayati, E. N. (2024). Artificial Intelligence in Web-Based Geographic Information Systems: A cross-disciplinary review. *International Journal of Research and Applied Technology (INJURATECH)*, 4(2), 309-316..
- Krčadinac, O., Stanković, Ž., Dudić, D., & Stošić, L. (2025). The role of artificial intelligence in modern information systems design: A systematic review. *Journal of Information Technology and Applications*, 15(1), 55–68.
- Kulkarni, A., & Ramanathan, C. (2025). GRAILS - A framework for embedding ethical safeguards in software applications for responsible AI. Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, 8(2), 1513–1523.
- Mohammed, Y. S., Akula, D. K., Syed, A., Haque, G. M. M., & Arafat, Y. (2025). The impact of artificial intelligence on information systems: Opportunities and challenges. *The American Journal of Engineering and Technology*, 7(8), 151–176.
- Preiß, N., & Westner, M. (2025, September). From Agents to Copilots: A systematic review of digital assistant technology adoption in proprietary productivity software. In *2025 20th Conference on Computer Science and Intelligence Systems (FedCSIS)* (pp. 565-576). IEEE.