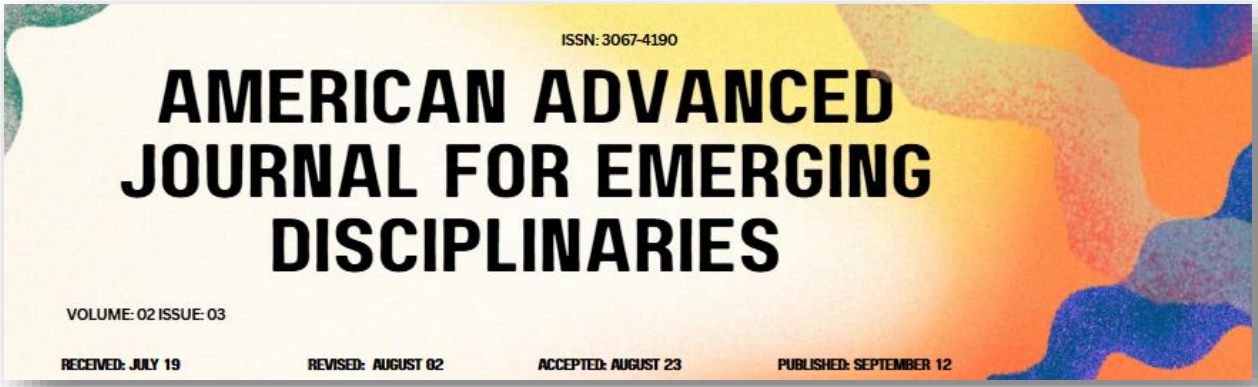


Trust-Embedded Agent Orchestration for Enterprise Service Management

Benjamin Clark
Asst Professor

Abstract

With enterprises increasingly adopting generative AI models, there is an escalating need for technology governance. Agentic generative AI promises increased process automation but requires intricate ownership and accountability mechanisms. This architectural design addresses these issues when RAG pipelines automate IT service management (ITSM) or human resource service delivery (HRSD) use cases. It proposes a governance-native agentic AI architecture, embedding governance in every part of the design, from process selection to retrieval-augmented generation (RAG) pipeline specification and technological implementation.

Every enterprise's AI strategy and governance framework should determine their agentic AI governance principles. Here, ownership of the agentic generative AI is repositioned from the enabler to the providing organization's governance risk compliance team. All internal IT knowledge sources, including policies, process, procedures, service catalogs, and product catalogs, require deep dive ingestion and curation. External knowledge sources, including resumes, job descriptions, and industry-specific suppliers, also need similar preparation. Therefore, three RAG pipelines—management of IT operations incidents and turn-up requests and deploying non-major service change requests—are augmented with a governance layer that enforces audit trails, supervises process usage and model alignments, and drives awareness and training.

Keywords: Governance-Native Agentic AI, Generative Artificial Intelligence Governance, Retrieval-Augmented Generation Pipelines, Enterprise Technology Governance, Agentic Process Automation, Ownership And Accountability Models, Governance Risk Compliance, IT Service Management Automation, Human Resource Service Delivery, Embedded AI Governance, Audit Trail Enforcement, Policy-Aware AI Systems, Knowledge Ingestion And Curation, Internal Enterprise Knowledge Sources, External Knowledge Integration, Model Alignment Supervision, Process Usage Monitoring, RAG Pipeline Governance, AI Awareness And Training, Enterprise-Scale Agentic Architectures.

1. Introduction

The Governance-Native Agentic AI Architecture for Enterprise ITSM and HRSD Automation via Orchestrated RAG Pipelines is presented in a formal, precise, and professional style, emphasizing governance, accountability, and architecture. Definitions, objectives, and implications for enterprise ITSM and HRSD are all mapped out clearly. Automation and augmentation of Information Technology Service Management (ITSM) and Human Resource Service Delivery (HRSD) processes using Generative Artificial Intelligence (GAI) promise significant operational efficiencies and improved end-user experience. These exciting developments are tempered, however, by deep concerns around trust and accountability. The consolidated generation of enterprise-grade GAI services must therefore be designed following governance-native principles. Audit records of all agentic GAI actions must be preserved to ensure clear accountability. Governance and version control processes must also be established for all foundation models and other data and algorithms driving agentic GAI behavior.

The Governance-Native Agentic AI Architecture for Enterprise ITSM and HRSD Automation Using Orchestrated RAG Pipelines brings together these concepts through the design of governance-native, audit-capable,

ISSN: 3067-4190

AMERICAN ADVANCED JOURNAL FOR EMERGING DISCIPLINARIES

VOLUME: 02 ISSUE: 03

RECEIVED: JULY 19 REVISED: AUGUST 02 ACCEPTED: AUGUST 23 PUBLISHED: SEPTEMBER 12

agentic GAI services for enterprise ITSM and HRSD. The design is validated through the architecting of two RAG pipeline use cases—ServiceNow Incident and Request Management and Azure OpenAI Change Enablement and Asset Management—where generative LLMs are orchestrated in the roles of retriever and synthesiser.

1.1. Overview of Content Structure

The Governance-Native Agentic AI Architecture for Enterprise ITSM and HRSD Automation Using Orchestrated RAG Pipelines is presented in nine main sections. Following this introduction, the research significance and objectives are outlined. The system architecture and design principles are described before details of the ITSM and HRSD automation use cases examined. A concise account of RAG architecture delineates data ingestion, retrieval-augmented reasoning, and retrieval-augmented generation, elaborating the function of each in generating end-user-facing content. Six concluding sections discuss trust, accountability, and explainability; auditability and governance logs; model governance and versioning; integration of third-party RAG-based agentic capabilities; orchestration of multiple RAG-based agents; and future work to extend the current design principles into alternative orchestration framework designs. The examination of security, privacy, and ethical impacts of deploying AGI-based agents remains key priorities within ongoing research-enabling follow-up architecture-based implementations across the realisation of Academia Group brand products and Impact Instrumentation strategies.

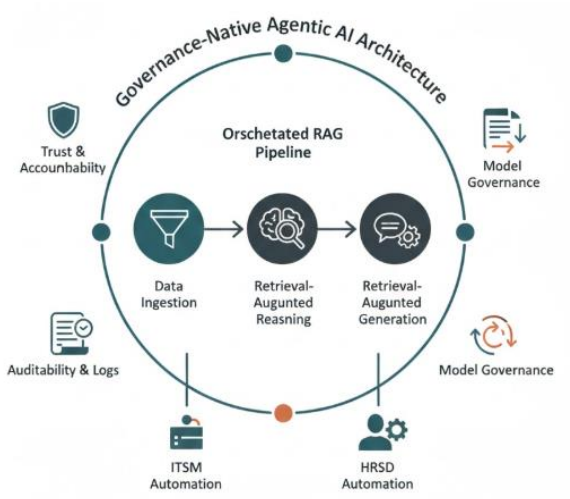
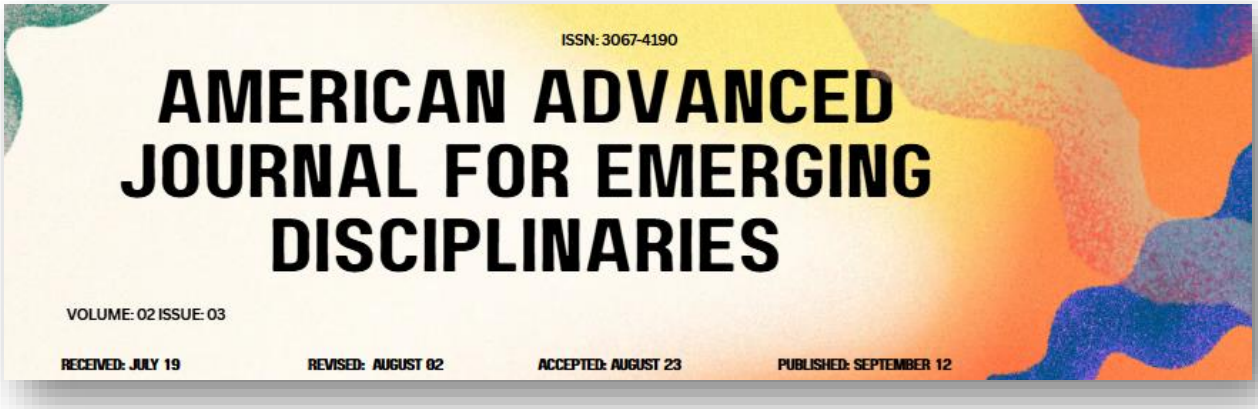


Fig 1: Governance-Native Agentic AI: Orchestrating RAG Pipelines for Integrated ITSM and HRSD Automation

2. Background and Motivation

OpenAI's ChatGPT, and similar solutions from other organizations, have sparked widespread interest in Generative AI (GenAI) technologies for automation within enterprises, particularly in Information Technology Service Management (ITSM) and Human Resource Services Delivery (HRSD). However, many implementations remain experimental, presenting business, legal, and reputational risks. Furthermore, apart from custom-built generative agents, implementations either do not focus on trust or accountability or treat these fleetingly, elevating their importance only in the future work section. In contrast, governance is a core principle of the framework being proposed.

A governance-native design for an agentic AI architecture suitable for enterprise ITSM and HRSD–automation use cases is presented. Such an architecture is particularly required when using retrieval–augmented generation (RAG) pipelines that reproduce this risk. Within governance, trust and accountability are prioritized, for



trust is essential for acceptance, and accountability is essential for governance in decision-making, especially for high-stakes automation.

Equation 1: Mathematical formalization of the paper's architecture (step-by-step)

Governance-native control plane supervising **orchestrated RAG pipelines**, including **audit logs**, **policy orchestration**, and **model governance/versioning**.

1.1 Sets, objects, and notation

Query & context

- User query: $q \in Q$
- Conversation context (history, user profile, ticket metadata): $c \in \mathcal{C}$
- Enterprise knowledge base (curated documents, service catalog, policies, etc.):
$$K = \{d_1, \dots, d_N\}$$

Policies (Policy Cloud)

- Active policy set: $\pi \in \Pi$
Policies define "allowable topics, exclusions, and information sources".
- Policy evaluation function:
$g_\pi(\cdot) \in$ {ALLOW,DENY,REDACT,ESCALATE}

2.1. Significance of Research and Objectives

In-post generative artificial intelligence (GenAI) can perform agentic tasks, such as interacting with users, executing complex requests, and integrating tools. These capabilities have been employed to design and deploy autonomous digital workers for enterprise IT service management (ITSM) incident and processing request fulfillment and human resources service delivery (HRSD) response handling. However, the rapid adoption of GenAI tools in sentiment- and language-focused use cases lacks governance. The increasing use of technology in business processes results in substantial risks that require monitoring, auditing, and governance. Customers want to know:

- Who trained these models?
- What data was used?
- How is the output being governed?
- What logging is in place?
- Who approved the various GenAI artifacts and models?

Additional use cases include change enablement and asset management, where GenAI augments the work of a human operator. An architectural framework is presented that emphasizes governance, accountability, and architectural design.

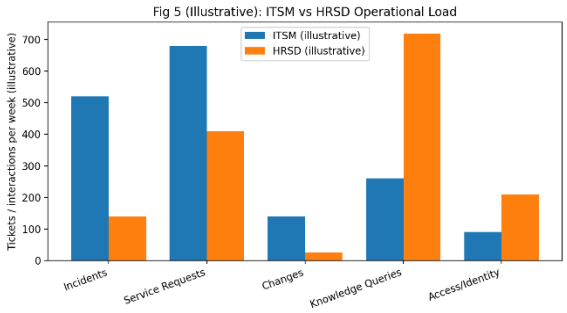
The dual-use nature of GenAI distribution implies that it can simplify both honest and malicious cyber activity. Any unscrupulous or malicious person can leverage GenAI for illegal cyber activity, including phishing scams, hacking, malicious social engineering, etc. These dangers have raised alarms for various stakeholders, and a number of storm warnings and advisory messages have been disseminated continuing an ongoing public dialogue. All stakeholders believe that legitimate safeguards and the required audit trail need to be established to ensure that responsible use becomes the standard rather than the aberration. A fundamental responsibility in any society comes from the accepted principle that use creates responsibility. For GenAI, society undertakes the responsibility of creating and establishing safeguards, conditions, and rules and regulation that helps maintain, uphold, and demand responsible use even when there is temptation to use it for malicious or dishonest purpose.

ISSN: 3067-4190

AMERICAN ADVANCED JOURNAL FOR EMERGING DISCIPLINARIES

VOLUME: 02 ISSUE: 03

RECEIVED: JULY 19 REVISED: AUGUST 02 ACCEPTED: AUGUST 23 PUBLISHED: SEPTEMBER 12



3. Architectural Principles

The Governance-Native Agentic IT Architecture for Enterprise ITSM and HRSD Automation Utilising Orchestrated RAG Pipelines is guided by five architectural principles. Foremost amongst them is the adoption of a governance-native design that embeds traceability, accountability, auditability, and transparency-gates at each step and transformation. This approach relies on co-ordination with external governance authorities across the full orchestration, transformation, and delivery chains to ensure that these native communication points, and the agentic AI's normal use of them to implement product and process domain specific governance and assurance decision chains, are themselves equipped with sufficient ethical, lawful, reliable, secure, and private AI capabilities.
These communication channels are therefore subject to the same level of risk assessment, governance and assurance as all of the AI acting on the organisation's behalf. This principle lays a foundation for a second architectural principle—an agent-native design that establishes RAG pipeline-driven COPs as the agents of choice for low-risk, high-volume, repetitive transactional execution supporting all aspects of enterprise mission outcomes, including IT service management (ITSM) and human resource service delivery (HRSD).

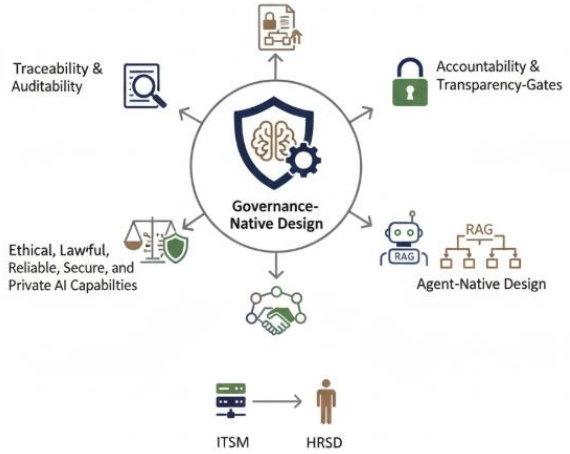


Fig 2: Architectural Principles for Governance-Native Agentic AI: Embedding Transparency Gates and Ethics-Aware RAG Pipelines in Enterprise Automation

3.1. Governance-Native Design

The architecture is explicitly designed for governance, facilitating accountability through auditable decision trails that are accessible to designated actors. No automated agentic actor executes any action without prior, specific governance endorsement. The action undertaken, and the rationale for the action in human-readable form, is explicitly and automatically captured as structured metadata and published to a governance support system. This supports the development, review, ratification/rejection, and archival of regular governance documentation. Temporal governance enables rapid response to urgent but time-varying situations without sacrificing any of the above properties.
Governance native design is distinct from safety. An equals sign between cause and effect is not assumed; models can generate dangerous outputs. Therefore, safety, distance, human supervision, and risk mitigation still matter. These need to be handled. It deals entirely with how the actions of agentic models get governed prior to execution. An underlying principle is that accountability cannot be achieved by adding some bureaucratic layer overhead, nor by punishing violations of rules. It is the interaction of

ISSN: 3067-4190
AMERICAN ADVANCED JOURNAL FOR EMERGING DISCIPLINARIES
VOLUME: 02 ISSUE: 03
RECEIVED: JULY 19 REVISED: AUGUST 02 ACCEPTED: AUGUST 23 PUBLISHED: SEPTEMBER 12

executive agents, governance, and record-keeping that focuses the system and separates the important parts from distorting noise. It can be difficult or impossible to create a training set for a shift in supervision model.

4. System Architecture

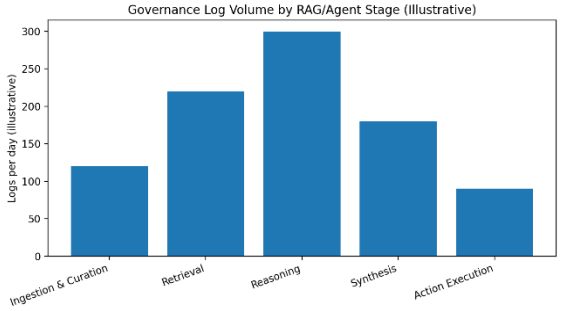
The architecture consists of a Governance-Native Agentic AI Control Plane and Orchestrated RAG Pipelines, designed to enable Trust and Accountability Framework Objectives.

Control Plane

The Control Plane implements the Governance-Native Agentic AI principles and acts as a liaison between the Policy Cloud and one or more Orchestrated RAG Pipelines for Enterprise ITSM/HRSD Automation. It contains the governance and accountability genetics for the Agentic AI ecosystem, conferring native capability for auditability, trust, policy orchestration, and model governance.—Auditability and Governance Logs—The Control Plane maintains dedicated audit logging for all agent prompts and responses to ensure that updates and changes executed through an Agent can be interrogated by stakeholders. All interactions are logged in a secure vault for query and retrieval. Each interaction is tagged for confidentiality, enabling the filtering of PII and sensitive data prior to external handling.

Policy Orchestration Layer—Policies in the Policy Cloud define the allowable topics, exclusions, and information sources for Enterprise ITSM and HRSD use cases. The policy extraction and orchestration layer works in tandem with the orchestration capabilities of the Orchestrated RAG Pipelines to deliver adaptive control over the prompt engineering and RAG architecture for each request.

Model Governance and Versioning—The Control Plane provides a layer of governance and versioning for the various models used in the Orchestrated RAG Pipelines. Model governance as a critical enabler of Trust and Accountability Framework Objectives.



Equation 2: Data ingestion & curation → curated KB

"Data ingestion and curation" as the first orchestration dimension.

Model this as a transformation pipeline:

1. **Raw sources** $S = \{s_1, \dots, s_m\}$ (internal + external)
2. **Ingestion function** $I(\cdot)$ parses/normalizes:

$$\tilde{d}_i = I(s_i)$$

3. **Curation & governance tagging** $C_\pi(\cdot)$ applies policy constraints and metadata (confidentiality/PII tags mentioned in audit logging).

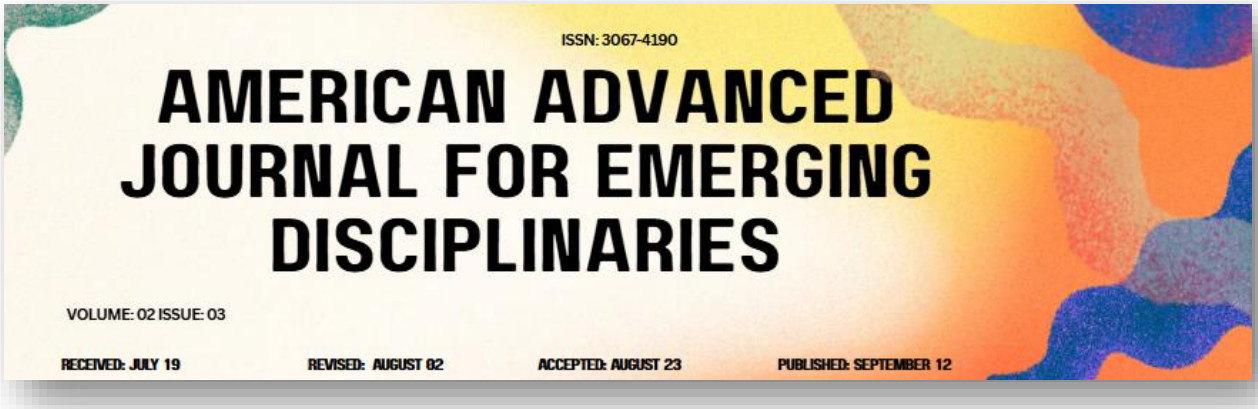
$$d_i = C_\pi(\tilde{d}_i) \quad \Rightarrow \quad K = \{d_i\}_{i=1}^{N}$$

So the curated KB is:

$$K = C_\pi(I(S))$$

4.1. Component Overview

The AI architecture comprises two primary components: a governance-native orchestrated RAG pipeline and a dynamic agent native to AI governance, accountability, and auditability. The orchestrated RAG pipeline offers enterprise ITSM and HRSD solutions with an emphasis on incident and request fulfilment, change enablement, and asset management. Complete pipelines address non-ITSM-



HRSD domains with a general-purpose pipeline supporting RAG operations outside current domain-specific scope. The interactive agent accesses the orchestrated RAG pipeline via major cloud framework APIs. Additional agentic functionality is being designed. Internal framework wrapper functionality allows the agent to serve as a governance-native, RAG-aware, intelligence-native, dynamic function independent of the RAG pipeline and oriented toward external governance-native products.

Category	ITSM (illustrative)	HRSD (illustrative)
Incidents	520	140
Service Requests	680	410
Changes	140	25
Knowledge Queries	260	720
Access/Identity	90	210

5. ITSM and HRSD Automation Use Cases

Governance-Native Agentic AI Architecture for Enterprise ITSM and HRSD Automation via Orchestrated Retrieval-Augmented Generation Pipelines
IT Service Management (ITSM) and Human Resource Service Delivery (HRSD) represent crucial spokes of the business-information-technology-service group of enterprise governance. Given their reliance on highly-scalable decisions of a wealth of routine but still-important business transactions, their governance-non-native automation depends on appropriately designed artificial intelligence technology. A Governance-Native Agentic AI Architecture enables the automation of representative use cases in ITSM and HRSD based on the current generation of Retrieval-Augmented Generation (RAG) pipelines. Governance-native design extends traditional database-centric architecture with a more-inclusive view that integrates trust, accountability, and explainability into the pipeline and model levels of the Agentic AI Architecture.

To highlight its implications in practice, incident and request management in ITSM and knowledge management in HRSD have been selected as immediate use cases for an Orchestrated RAG Pipeline. The first two dimensions of the use case logic — Data Ingestion and Curation, and Retrieval, Reasoning, and Synthesis — are addressed explicitly. The other dimensions of the design and execution logic anticipate future-scale, governance-native pipelines capable of integrating and orchestrating an expanded suite of Commercial Off-The-Shelf (COTS) RAG models, thereby unlocking their latent enterprise ITSM and HRSD automation capabilities.

5.1. Incident and Request Management

Correspondence and self-service management encompass IT Service Management (ITSM) automation responsibilities occurring outside the organization via email, websites, or APIs. In e-mail correspondences, the exchange of information occurs primarily in one language, while the automated generation of the first, second, and any subsequent reply can be in different languages, together with the automated translation of updates in ticket status.

ISSN: 3067-4190

AMERICAN ADVANCED JOURNAL FOR EMERGING DISCIPLINARIES

VOLUME: 02 ISSUE: 03

RECEIVED: JULY 19 REVISED: AUGUST 02 ACCEPTED: AUGUST 23 PUBLISHED: SEPTEMBER 12

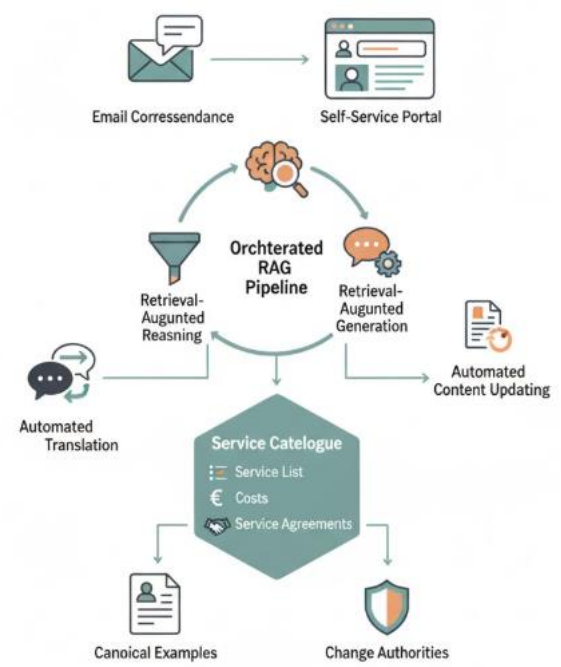


Fig 3: Automated Omnichannel Service Management: Multi-Language RAG Pipelines for Dynamic ITSM Correspondence and Catalog Governance

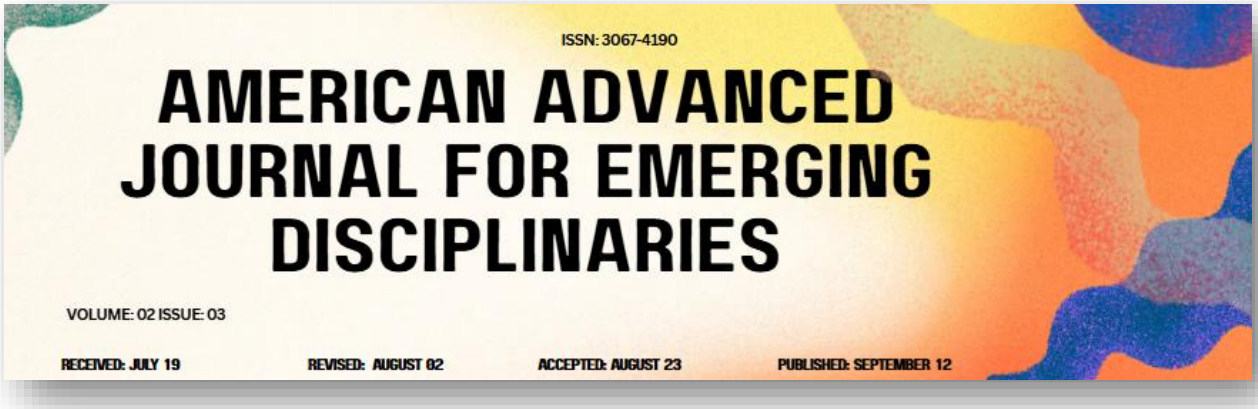
The service catalogue of a self-service portal requires accessible information up-to-date to encourage its use. It contains information about available services, service providers, prerequisites, costs, service description, terms of commitment, duration, or service agreements. Automated updating of this information can be through a large language model that ingests freely available information, such as News APIs or an aggregator service like the one present in WebSummarizer.ai, or specialised information, such as a website of Europe or a specific country that implements RAG pipelines. Automated reply to a request can be generated through canonical examples of previous solutions by the IT Support or changes approved by the Change Authorities.

5.2. Change Enablement and Asset Management

Designed for enterprise infrastructure, these use cases support common but challenging IT Service Management (ITSM) activities related to change enablement and asset management, which frequently require effective collaboration among different service delivery functions. The IT Service Management Forum not only provides guidelines for these ITSM activities but is also developing an ITSM ontology that enhances interoperation for enterprise Information Technology Service Management (ITSM). With the governance-native, agentic Agent Architecture for orchestration and automation of solicited and handled knowledge in the domain of ITSM—such as intelligent virtual agents (IVAs), governance-native agents (GNAs), agent-operating systems (AOSs), RAG pipelines, and so forth—the domain-specific RAG architecture can now be scaled and refined in a reproducible way.
Change enablement supports the planning and execution of changes with minimal disruption to IT services. The full-change enablement process encompasses the request for change, change assessment and authorisation, coordination, change build and test, change implementation, and post-implementation review. Change enablement operates a change advisory board (CAB) that assesses and authorises changes for implementation, as well as an emergency change advisory board (ECAB) that authorises urgent changes. The process draws on the enterprise asset, configuration and knowledge information, in addition to other service impact and risk information, to enable efficient assessment and planning of changes. Service continuity information is used to assess risk and impact of proposed changes on service continuity. Personalised macro templates and orchestration-type solutions enable efficient authoring and coordination of low-risk operational changes.

6. Orchestrated RAG Pipeline Design

A plaza is just a plaza unless piloted by pilgrimage. In an enterprise context, a pilgrimage is manifested as an orchestration of multiple retrieval-augmented generation (RAG) pipelines, with the plaza subsuming the governance



framework, the use case design and architecture, and the model governance, outcome validation, and pipeline supervision ancillary areas. Three delicate dimensions of orchestrating a RAG pipeline are identified: data ingestion and curation, retrieval-augmented reasoning, and reasoning-augmented synthesis, with agents assigned responsibilities within a dedicated governance framework. Overcoming a malaise of malapprop and malapod, redundancies are plumped and sneers ensured — as duplication is the price of delicacy. To bolster biceps and sorely needful elephants, a custodial orchestration of multiple RAG pipelines dramatises the plaza as pilgrimage. At the root refinery are three delicate domains, drawn from Stanza Process Server and Robotic4u Process Lifecycle555: denouement of series, the droplets of introduction, curation, and feeding into the pipelines; act two, landing at an ART centre, service and storage of arrival show; and returning to the custodian plaza.

Pipeline stage	Primary governance artifact
Ingestion & Curation	data_lineage_log
Retrieval	retrieval_provenance_log
Reasoning	policy_check_log
Synthesis	output_validation_log
Action Execution	approval_and_audit_commit

6.1. Data Ingestion and Curation

The RAG pipelines are designed for natural language understanding and generation-related tasks across diverse enterprise data sources. These tasks include modeling, coding, QA and Accountability retrieval, natural language understanding QA, FIPS, incident and service request management, knowledge curation, natural language understanding, governance logging, and notification message drafting. The task area-specific agents orchestrate processing pipelines and execute workflows involving manual involvement and approval steps. These agents accelerate agency tasks through access to multiple generative models or services and data sources.

Three types of models are needed for each task: a natural language generation model, a model that provides reasoning, judgment, and lookup capabilities, and a model that synthesizes the final output, combining the generation and reasoning model outputs. The pipeline orchestrator manages the execution and maintains a pool of models with the same capabilities or data. For data entry-level processing and operations, appropriate models are invoked along with services such as a enterprise data catalog, natural language understanding analysis, and Kubernetes-based application end user experience.

Equation 3: Retrieval as a constrained optimization

RAG retrieval selects top-k documents/passages relevant to q.

3. Embed query and docs:

$$\mathbf{e}_q = E(q), \quad \mathbf{e}_{d_i} = E(d_i)$$

2. Compute similarity score (e.g., cosine):

$$\text{sim}(q, d_i) = \frac{\mathbf{e}_q \cdot \mathbf{e}_{d_i}}{\| \mathbf{e}_q \| \| \mathbf{e}_{d_i} \|}$$

4. Apply **policy filter** (allowed sources/topics):

$$a_i = \mathbf{1}[g_\pi(d_i) = \text{ALLOW}]$$

4. Define policy-constrained score:

$$s_i = a_i \cdot \text{sim}(q, d_i)$$

5. Retrieve top-k:

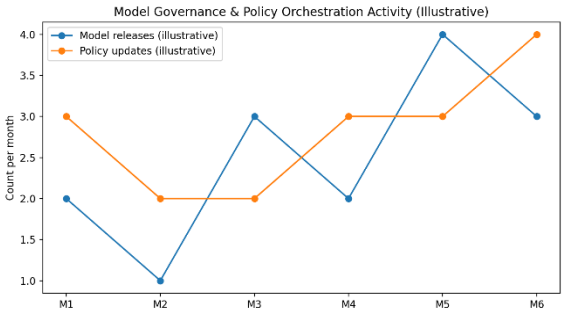
$$R_k(q, K, \pi) = \text{TopK}\left(\{(d_i, s_i)\}_{i=1}^{N}\right)$$

ISSN: 3067-4190

AMERICAN ADVANCED JOURNAL FOR EMERGING DISCIPLINARIES

VOLUME: 02 ISSUE: 03

RECEIVED: JULY 19 REVISED: AUGUST 02 ACCEPTED: AUGUST 23 PUBLISHED: SEPTEMBER 12



6.2. Retrieval, Reasoning, and Synthesis

The orchestrated RAG pipelines transform unstructured knowledge sources into formal models. Model templates define the attributes of scenario-specific and task-specific generative models for open-ended question answering and arbitrarily defined structured queries, respectively. Generation of the generative-model templates supports full, synthetic, knowledge-coverage, and union-based model subsets. The required knowledge sources are recursively acquired and converted from unstructured format into the generative-model definitions driven by the newly generated script-generating model; the result is a complete, fully defined open-ended question-answering process model. Compositions of RAG pipelines rewrite correspondences between task intents and interactions with task-specific structured generative models that formalize pre-structured task knowledge into result-response-generative models for structured query answering and task execution.

7. Trust, Accountability, and Explainability

Decisions made by systems should be trustworthy and auditable to the appropriate extent, being validatable by rooting all RAG-pipeline outputs back to credible sources. AI providers, model developers, and operational users are accountable for their workflows, models, and decisions, respectively. Operational end-users act as reviewers and validators of the generated outputs, ensuring that the intent of the portal-user query is preserved.

Recommended practice for external data ingestion into the LLMs is to create data subsets scoped to the domain areas of AI provider and user. Each domain-scope data subset should be curated, inhabited, and made accessible within LLMs in an incremental, controlled, and governance-driven fashion. Auditable and traceable governance decision logs should exist to prevent improper data-use claims during future audits, allow reconstruction of the LLM model-data-knowledge base trajectory over time, and enable what-if analyses for prospective diagnostic-contingency-support exploration. Governance-controlled model versioning is essential to support correct model deployments during risk and change deployment processes. Temporary and transitory-governance-controlled LLM-model and data-copy creations facilitate dry-runs, what-if explorations, and other necessary impact analyses without impacting the production-model integrity or authenticity.

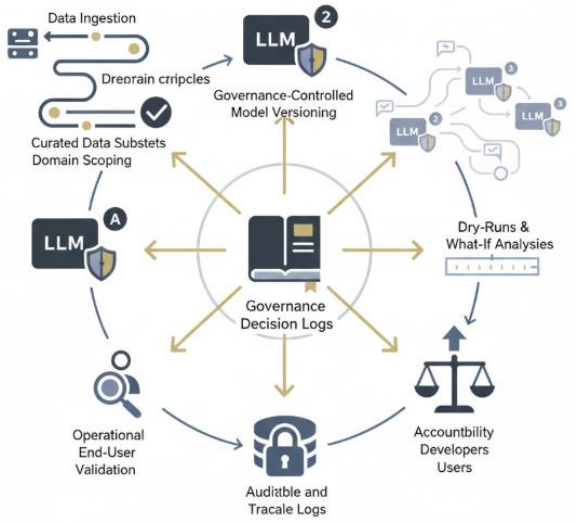
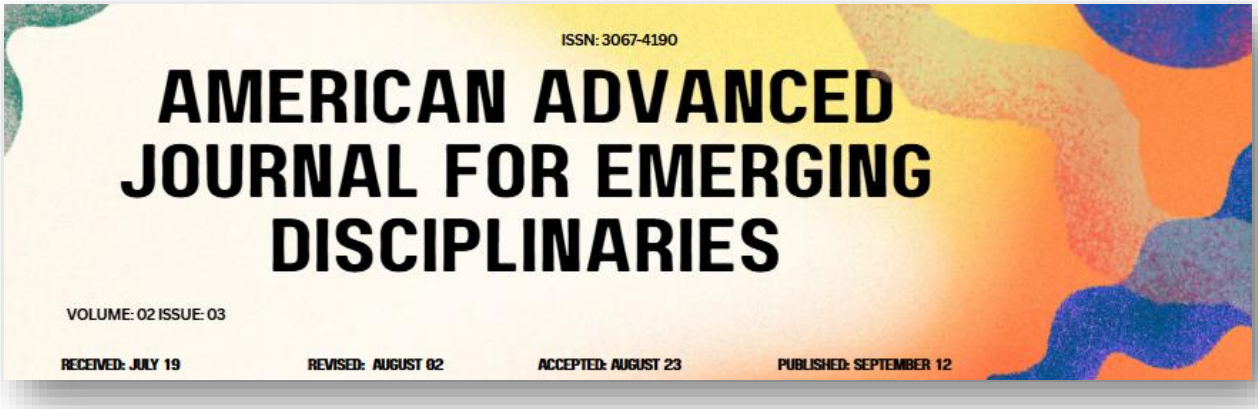


Fig 4: Rooted Trustworthiness in RAG Architectures: A Governance-Driven Framework for Auditable Data Lineage, Incremental Domain Curation, and Traceable Model Trajectories

7.1. Auditability and Governance Logs

Providing stakeholders with visibility into reasons for agentic AI actions and outputs requires a multi-layered



approach. One dimension of governance is the generation of immutable audit records—declaratives capturing activations of the agentic AI architecture during process execution. Factorised governance-oriented declaratives for each component facilitate governance logging requirements and alleviate overhead.

The current implementation of a governance-native architecture for enterprise ITSM and HRSD automation employs an immutable ledger based on the Hyperledger Fabric platform for audit trail generation. Creation of new logs in the ledger is accomplished by using chaincode that safely implements key transaction functions. The transaction invocation process is private, ensuring that only the governance-native architecture can activate it, while also providing self-sovereignty for the ledger.

7.2. Model Governance and Versioning

To ensure model governance and versioning, the design intrinsically integrates the act of development, retraining, and deployment of AI and machine-learning models for use in the orchestration. A simplified pipeline for the development and deployment of such models consists of four stages: (i) Refinement of the training data through a combination of human revision and interactive RAG pipelines that can be governed for accuracy and currency; (ii) Learning from telemetry and audit trails to support expected accuracy and other non-functionals of the models and pipelines; (iii) Learning from telemetry and audit trails during production for continuous improvement and nonredundancy to RAG pipelines; (iv) Solidity tests in a niche production environment for a subsegment of the user community for functional validation, in addition to a CI/CD environment for testing of non-cutting-edge code segments. The RAG model stays distinct but supported by the Telemetry Engine for accuracy and documentation. The Learning platform supports local and vendor models, guidance, and the sourcing of appropriately sized and skilled machine-learning teams.

8. Conclusion

In summary, the proposed governance-native agentic AI architecture enables natural-language automation for enterprise IT service management (ITSM) and human resources service delivery (HRSD). Support for diverse ITSM and HRSD use cases is established through the design of a Retrieval-Augmented Generation (RAG) pipeline that is subsequently orchestrated from data ingestion through augmentation to response synthesis in a sequence of stages: ingestion and curation of system operation and service desk resources; retrieval of relevant content within these resources in response to citizen queries; reasoning and elaboration of specific responses using large language models (LLM) within the Gov-ML-NG environment; and world knowledge augmentation using a continually updated context-specific Wikipedia subgraph. The architecture supports trust, accountability, and explainability by providing the auditability and governance logs required by decision-makers charged with overseeing LLM deployment in enterprise applications. The deployment of LLMs in ITSM and HRSD environments positions these technologies as viable enterprise agents. It is therefore essential to recognize the wider implications of agentic LLMs. Information governance—policy-driven design, stakeholder accountability, and effective model management—must be incorporated in future LLM applications to mitigate information stewardship risks and uncertainties. The first step in that direction is to consider information governance as a principal design constraint for agentic AI applications. Additional design principles for supporting governance by LLMs, either information-seeking or information-giving, are an area for future research.

ISSN: 3067-4190

AMERICAN ADVANCED JOURNAL FOR EMERGING DISCIPLINARIES

VOLUME: 02 ISSUE: 03

RECEIVED: JULY 19 REVISED: AUGUST 02 ACCEPTED: AUGUST 23 PUBLISHED: SEPTEMBER 12

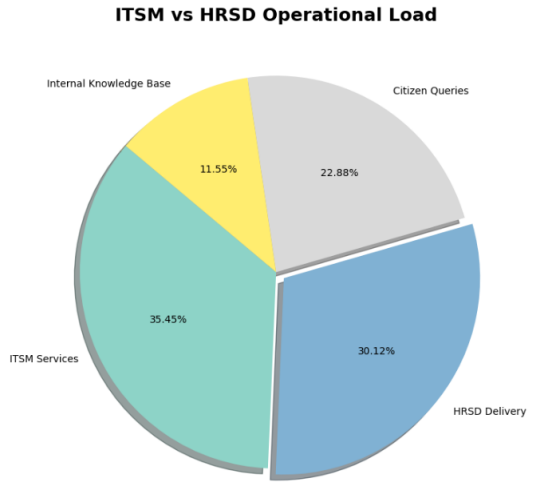


Fig 5: ITSM vs HRSD Operational Load

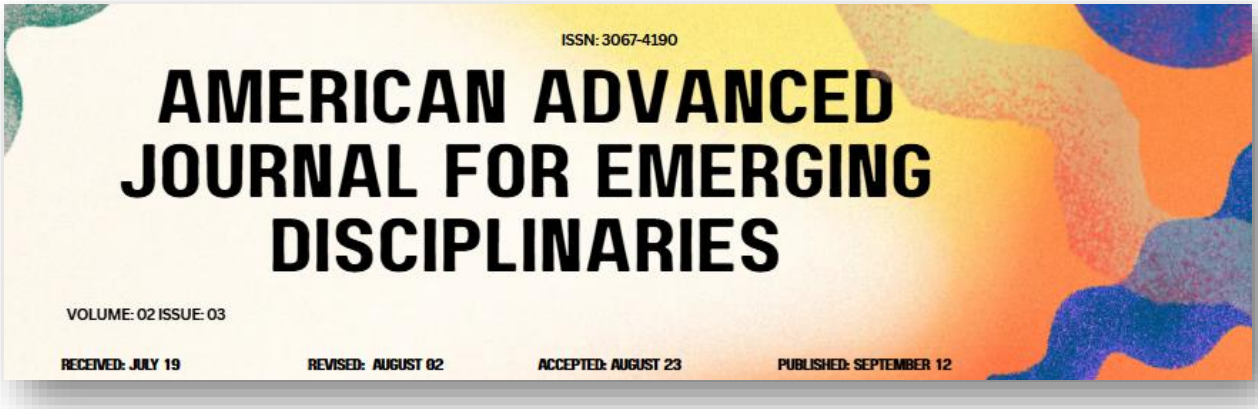
8.1. Summary of Key Insights and Future Directions

A state-of-the-art governance-native agentic AI architecture for automating enterprise IT service management and human resource service delivery using orchestrated retrieval-augmented-generation pipelines has been proposed. Enterprise ITSM and HRSD governance and assurance requirements have been articulated; distinct incident, service request, change enablement, knowledge management, and identity verification use cases have been defined; and agent orchestration support for the design and delivery of RAG pipelines augmented with governance and assurance facilities has been described.

Technical design components of the architecture have been specified, including the RAG pipeline stages required to endlessly absorb, curate, and enrich enterprise data sources and to support secure, explainable, and trusted Data-, Model-, and Governance-as-a-Service operations. A pipeline-centric agentic delivery approach has been articulated, recognizing the crucial dependency on data gravity, quality, and availability. The overall approach enables agents to invoke RAG pipelines tailored for different stakeholder needs, to fulfil user needs with trust, accountability, and governance embedded throughout the process.

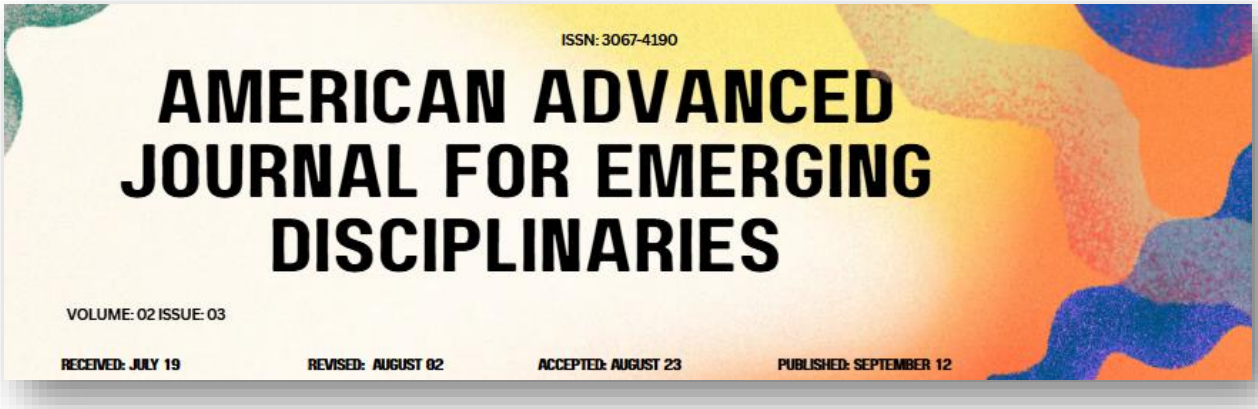
9. References

1. Amershi, S., et al. (2022). Guidelines for human-AI interaction. Communications of the ACM, 65(1), 84–92.
2. Baier, S., Gimpel, H., & Greeven, C. (2022). Enterprise AI governance: A systematic literature review. Business & Information Systems Engineering, 64(6), 739–758.
3. Reddy Segireddy, A. (2024). Federated Cloud Approaches for Multi-Regional Payment Messaging Systems. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 15(2), 442-450.
4. Bubeck, S., et al. (2023). Sparks of artificial general intelligence: Early experiments with GPT-4. arXiv Preprint.
5. Dwivedi, Y. K., et al. (2023). So what if ChatGPT wrote it? Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI. International Journal of Information Management, 71, 102642.
6. Mangalampalli, B. M. (2024). Transparent Intelligence Explainability Frameworks for AI-Driven Clinical Decision Support in Healthcare Business Intelligence. International Journal of Research Publications in Engineering, Technology and Management (IJRPETM), 7(3), 10566-10579.
7. European Union. (2024). Artificial Intelligence Act. Official Journal of the European Union.
8. Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. Minds and Machines, 30(4), 681–694.
9. Gunning, D., & Aha, D. W. (2021). DARPA's explainable artificial intelligence program. AI Magazine, 42(2), 44–58.
10. Guo, J., Cheng, J., & Cleland-Huang, J. (2021). Semantically enhanced service orchestration

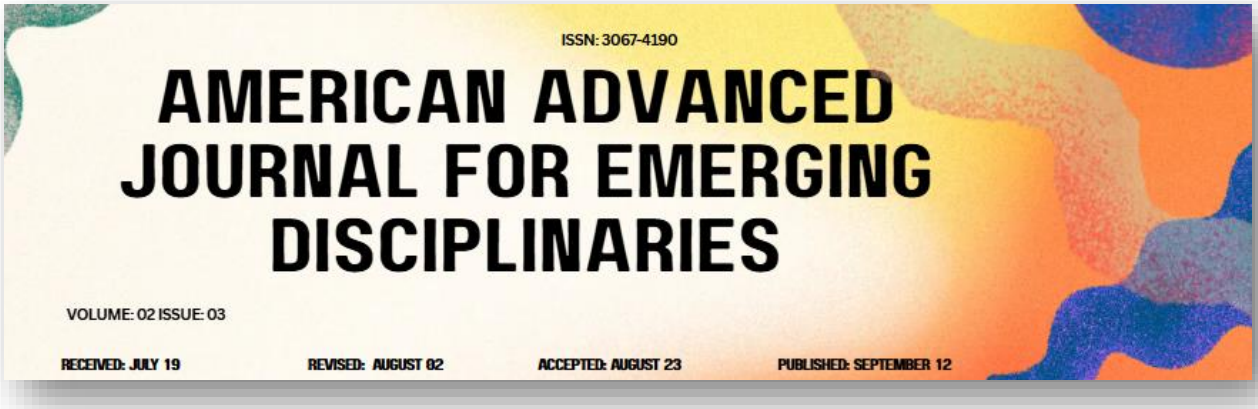


using artificial intelligence. IEEE Access, 9, 104132–104147.

11. Kolla, S. K. (2024). Clinical Knowledge Intelligence through Deep Learning and Natural Language Understanding in Healthcare Platforms. International Journal of Advanced Engineering Science and Information Technology (IJAESIT), 7(3), 14088.
12. Haarmann, M., & Röglinger, M. (2021). Enterprise service management: A research agenda. Business Process Management Journal, 27(7), 2065–2083.
13. Holzinger, A., et al. (2022). Towards trustworthy AI. Nature Machine Intelligence, 4(10), 832–842.
14. Hua, W., Yang, X., Jin, M., Li, Z., Cheng, W., Tang, R., & Zhang, Y. (2024). TrustAgent: Towards safe and trustworthy LLM-based agents. Findings of the Association for Computational Linguistics: EMNLP 2024.
15. Kolla, T. (2024). Intelligent Discovery and Governance of Healthcare Data Assets Through AI-Powered Catalog Architectures. International Journal of Emerging Trends in Engineering and Management Research, 9(4), 16083.
16. ISO/IEC. (2022). ISO/IEC 23894:2022—Information technology—Artificial intelligence—Risk management.
17. ISO/IEC. (2024). ISO/IEC 42001:2023—Artificial intelligence management system (AIMS).
18. Jarrahi, M. H., et al. (2023). Artificial intelligence and knowledge management: A systematic review. Journal of Knowledge Management, 27(5), 1295–1324.
19. Kolla, S. H., & Peddi, R. K. (2024). Designing Governance-Aligned GenAI Pipelines Using Small Language Models for Enterprise Workflow Intelligence. International Journal of Science, Research and Technology, 7(6), 13256-13268.
20. Jobin, A., Ienca, M., & Vayena, E. (2020). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 2(9), 389–399.
21. Kamar, E. (2020). Directions in hybrid intelligence. Proceedings of IJCAI 2020, 4792–4799.
22. Kaur, D., et al. (2022). Trustworthy artificial intelligence: A review. ACM Computing Surveys, 55(2), 1–38.
23. Kerner, S. M. (2023). Enterprise AI governance and responsible AI adoption. Computerworld.
24. Kölle, M., et al. (2023). AI-supported IT service management: Current state and future directions. Information Systems Frontiers, 25(6), 2147–2164.
25. Kolla, S. H. (2024). Retrieval-Augmented Enterprise Intelligence: Enhancing Accuracy, Trust, and Operational Decision-Making. International Journal of Future Innovative Science and Technology (IJFIST), 7(2), 12425.
26. LeCun, Y., Bengio, Y., & Hinton, G. (2021). Deep learning for AI systems: Recent developments and future challenges. Communications of the ACM, 64(7), 58–65.
27. Lewis, P., et al. (2021). Retrieval-augmented generation for knowledge-intensive NLP tasks. Advances in Neural Information Processing Systems, 34, 9459–9474.
28. Liao, Q. V., et al. (2020). Questioning the AI: Informing design practices for explainable AI. Proceedings of CHI 2020.
29. Lu, Y., & Xu, X. (2021). Resource virtualization and orchestration for intelligent manufacturing. Robotics and Computer-Integrated Manufacturing, 68, 102094.
30. Loganathan, R. (2021). Integrated Risk and Compliance Frameworks for Global Data Center Operations: A Governance-Centric Approach. Universal Journal of Computer Sciences and Communications, 1(1), 1-26.
31. Meske, C., Bunde, E., Schneider, J., & Gersch, M. (2022). Explainable AI in information systems research. Electronic Markets, 32(4), 2059–2078.



32. Microsoft. (2024). Responsible AI transparency report.
33. Mikalef, P., Gupta, M., & Pappas, I. O. (2023). AI capability and organizational performance. Information & Management, 60(2), 103731.
34. NIST. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0).
35. OECD. (2021). OECD framework for the classification of AI systems.
36. Mattaparthi, R. (2022). Engineering Predictive Industrial Systems Through IoT-Driven Asset Monitoring and Machine Learning Prognostics. International Journal of Future Innovative Science and Technology (IJFIST), 5(1), 7790.
37. Papagni, G., de Pagter, J., Zafari, S., Filzmoser, M., & Koeszegi, S. T. (2023). Artificial agents' explainability to support trust: Considerations on timing and context. AI & Society, 38(3), 947–960.
38. Parasuraman, R., & Riley, V. (2020). Humans and automation: Trust in intelligent systems. Human Factors, 62(3), 369–391.
39. Park, J., et al. (2023). Generative agents: Interactive simulacra of human behavior. Proceedings of the ACM on Human-Computer Interaction, 7(CSCW1), 1–39.
40. Rahwan, I., et al. (2020). Machine behaviour. Nature, 568(7753), 477–486.
41. Mukesh, A., & Aitha, A. R. (2021). Insurance Risk Assessment Using Predictive Modeling Techniques. International Journal of Emerging Research in Engineering and Technology, 2(4), 68-79.
42. Raji, I. D., et al. (2020). Closing the AI accountability gap. Proceedings of FAT 2020.
43. Ribeiro, M. T., Singh, S., & Guestrin, C. (2020). Why should I trust you? Explaining the predictions of any classifier. Proceedings of KDD.
44. Russell, S. (2021). Human compatible: Artificial intelligence and the problem of control. Penguin Books.
45. Samek, W., Montavon, G., Lapuschkin, S., Anders, C. J., & Müller, K. R. (2021). Explaining deep neural networks and beyond. Proceedings of the IEEE, 109(3), 247–278.
46. Bandi, V. D. V. K. (2024). AI-Driven Predictive Risk Modeling Architectures for Financial Systems. International Journal Of Finance, 37(3), 54-78.
47. Sandhu, R., & Samarati, P. (2021). Access control in distributed enterprise systems. IEEE Security & Privacy, 19(3), 74–82.
48. Sarker, I. H. (2022). AI-based cybersecurity: Methods, applications, and challenges. Journal of Big Data, 9(1), 1–35.
49. Shneiderman, B. (2022). Human-centered AI. International Journal of Human–Computer Interaction, 38(6), 495–504.
50. Mangala, N. (2024). Leveraging Microsoft Fabric lakehouse as an AI-ready data platform for enterprise analytics. Journal of Information Systems Engineering and Management.
51. Stahl, B. C., et al. (2022). A systematic review of trustworthy AI. ACM Computing Surveys, 55(13), 1–38.
52. Sun, T. Q., & Medaglia, R. (2020). Mapping the challenges of AI in the public sector. Government Information Quarterly, 37(4), 101493.
53. Tangi, L., Janssen, M., Benedetti, M., & Noci, G. (2022). Digital government transformation and AI. Government Information Quarterly, 39(2), 101640.
54. Tchorzewski, J., et al. (2022). Enterprise orchestration for intelligent service ecosystems. Future Generation Computer Systems, 132, 95–108.
55. Reddy, V. A. R. (2022). Data-Driven Healthcare Operations: Architecting Unified Member, Provider, and Claims Intelligence Platforms. International Journal of Science, Research and Technology, 5(5), 8511-8521.



56. van der Aalst, W. M. P. (2021). Process mining and AI for business process management. Communications of the ACM, 64(8), 30–33.
57. Varshney, K. R. (2022). Trustworthy machine learning. Manning Publications.
58. Vinuesa, R., et al. (2020). The role of AI in achieving the Sustainable Development Goals. Nature Communications, 11, 233.
59. Wang, D., et al. (2022). Human-AI collaboration: State of the art and future directions. AI Magazine, 43(4), 392–405.
60. Weidinger, L., et al. (2022). Taxonomy of risks posed by language models. Proceedings of FAccT 2022.
61. Inala, R. Designing Scalable Technology Architectures for Customer Data in Group Insurance and Investment Platforms.
62. Wiener, M., et al. (2022). AI governance mechanisms in organizations. Business & Information Systems Engineering, 64(4), 425–440.
63. Xu, F., et al. (2021). Intelligent service orchestration in cloud computing. Future Generation Computer Systems, 115, 613–626.
64. Yang, G., et al. (2021). Explainable artificial intelligence for trustworthy decision support. Information Sciences, 578, 289–305.
65. Pamisetty, V., & Amistapuram, K. Smart Decision Support Systems For Dynamic Tax Policy Optimization Using Reinforcement Learning.
66. Yin, M., Wortman Vaughan, J., & Wallach, H. (2023). Understanding human trust in AI systems. Proceedings of CHI 2023.
67. Zhang, Y., et al. (2022). AI-enabled enterprise digital transformation. Information Systems Frontiers, 24(6), 1805–1821.
68. Zhao, X., et al. (2023). Multi-agent reinforcement learning: A survey. IEEE Transactions on Neural Networks and Learning Systems, 34(12), 9380–9402.
69. Yandamuri, U. S. AI-Driven Decision Support Systems for Operational Optimization in Hospitality Technology.
70. Zhou, L., et al. (2024). Large language model agents: A survey. arXiv Preprint.
71. Zhou, Y., et al. (2023). Autonomous agents powered by large language models: A survey. arXiv Preprint.
72. Zicari, R. V., et al. (2021). Z-inspection®: A process to assess trustworthy AI. IEEE Transactions on Technology and Society, 2(2), 83–97.
73. Davuluri, P. N. Integrating Artificial Intelligence into Event-Driven Financial Crime Compliance Platforms.
74. OpenAI. (2024). GPT-4 technical report. arXiv Preprint.
75. Google DeepMind. (2024). Frontier safety framework. Google DeepMind.
76. World Economic Forum. (2024). AI governance alliance: Responsible AI framework for enterprises.