



Governed Agentic AI for Enterprise Automation

Benjamin Clark

Asst Professor

Abstract : In rapidly changing environments, enterprises must adapt at sufficient speed to survive. A popular experimental approach is to automate part of an organization by delegating the fulfilment of business functions to intelligence agents—software agents capable of performing work autonomously. An architecture is described where such agents are centered around an execution orchestrator and decision workflows. However, a key limitation of agentic systems is that speed of decision-making, required for throughput and reliability, is not always sufficiently fast, leading to delays.

The principal driver is a gap in the architecture. In enterprise environments, it is essential that the internal process of the agent decision-makers is harmonized by an orchestrator and that any decision pipeline supporting its reasoning must operate with sufficient speed. This gateway constitutes the core of secure knowledge-centric decision-making: the production of decisions, recommendations, and predictions with the quality and assurance properties required by the task domain. Secure knowledge-centric reasoning is thus a counterpart to the agentic foundation being automated, enabling speed, reliability, and trust in the choices that affect the execution of business functions.

Keywords: Governed Agentic AI,Enterprise Service Automation,Secure Knowledge Management,Knowledge-Centric Decision Intelligence,AI Agents,Retrieval-Augmented Generation (RAG),Enterprise Knowledge Graphs,Intelligent Workflow Orchestration,AI Governance and Compliance,Context-Aware Decision Support.

1. Introduction

An enterprise service automation framework keeps platforms and applications highly available, reliable, and performant. Ideally, these systems automate the enablement, assumptions, and support of services by maintaining, monitoring, and controlling service management, governance, and delivery platforms. Agentic artificial intelligence (AI) technology provides the means for highly integrated, autonomous automations that support enterprise service delivery and enterprise responsibility. Agentic enterprise service automations automate business processes, service management convolutions, and enterprise systems. These systems manage self-administering role-based-business-as-a-service components inside the enterprise ecosystem, addressing the service lifecycle management needs of enterprise responsibilities, including enterprise resource planning (ERP) and customer relationship management (CRM).

Agentic AI permits the automation of enterprise responsibilities, enabling Enterprise Service Automation with seamless functioning under all circumstances. However, as with any enterprise service, such Enterprise Agentic Automations must be governed, risk-managed, compliant, and assured. The proposed governance architecture offers a comprehensive policy, risk, compliance, and assurance framework that addresses enterprise needs. It establishes cohesive regulatory, risk, and compliance design principles



covering the life cycle of such enterprise services, extending beyond narrow Operational Technology and Information Technology Audit domains. Without a defined governance framework, organisations remain exposed to the risk of automation failing to serve the executive interest of the organisation leading to unnecessary service delivery failures and enterprise reputational damage whilst missing out on the benefits Enterprise Agentic Automations can deliver.

Table 1. GRC Layer Responsibilities and Governing Mechanisms

Layer	Primary Responsibility	Key Mechanism	Illustrative Metric
Policy	Defines design/operation rules for agentic AI and third-party actors	Policy-as-code, approval gates	Policy coverage ratio (A_p)
Risk	Aligns agent decisions with business objectives; mitigates adverse impact	Risk register, escalation triggers	Risk-mitigation coverage (M_r)
Compliance	Ensures adherence to privacy, security, sector regulation	Regulatory mapping, control testing	Compliance coverage (C_c)
Assurance	Independent verification of GRC evidence	Audit logging, third-party attestation	Audit pass rate (%)

Table 2. Agentic AI Execution Modules (Section 4)

Module	Function	Input	Output
Intent Understanding	Parses and disambiguates business request	Natural-language / API request	Structured intent object
Plan Generation	Builds task/decision workflow	Intent object, policy constraints	Executable task graph
Task Execution	Invokes enterprise resources via API wrappers	Task graph	Task results, logs
Monitoring	Tracks throughput, latency, faults	Execution telemetry	Performance metrics



Module	Function	Input	Output
Improvement (Runtime Learning)	Adjusts policy/plan based on outcomes	Metrics, escalation history	Updated plan heuristics

Table 3. KAI Trust Pillars (Section 5)

Pillar	Definition	Governance Question Answered
Explainability	Rationale for a decision is expressible in human-interpretable terms	"Why was this decision made?"
Traceability	Decision inputs and provenance can be reconstructed	"Where did the evidence come from?"
Coherency	Reasoning is internally consistent across the pipeline	"Does the answer logically follow?"
Auditability	Full decision trail is retrievable for independent review	"Can this be verified after the fact?"

2. Theoretical Foundations

Agentic AI concepts encompass the primary notion of autonomy in the formal definition of agents and explain agentic systems at organizational, operational, and functional actor types. The refinement of a decision workflow comprises the identification of actors and their collective decision participation as well as the delineation of intent-supporting data sources and structures for the analyst and author roles. The adoption of an actor-network theoretic perspective reveals how ensuing purpose designates agency and constructs agentic entities by conferring functional operating roles and responsibilities to functionally grouped actors.

Knowledge-centric decision intelligence involves a formalized computing in the knowledge domain complemented by skills and experience cognizance. Data-quality provenance provides a basis for decision confidence and impact assessments. Trust in automated decision and recommendation systems builds upon four complementary propositions: explainability, traceability, coherency, and auditability of operations. Such knowledge-centric decision rationale assists audit and regulation processes, ensuring enterprise legal responsibility, ethical mandates, and governance framework policy compliance.

Governance, risk, and compliance (GRC) embody a recognized organizing model that lends itself to enterprise IT and security investment planning. The three layers of policy, risk, and compliance subsume operations and assurance elements to encompass all facets



of an enterprise's use of GRC. The creation of enterprise governance structure design and operations facilities must be modelled or assisted effectively to ensure the undertaking's purpose conforms to policy and communication across concern domains. The integration of GRC with AI system design, operation supported, use ensured, and system-generated recommendations facilitates a design-and-deliver functioning commitment.

3. Governance and Compliance Architecture

An integrated governance architecture for agentic AI in enterprises encompasses policy, risk, compliance, and assurance elements, along with delineated roles, responsibilities, and accountability. Data governance controls safeguarding privacy, security, and compliance with regulatory regimes are specified, as are mechanisms for auditing, logging, verifiability, and assurance by third parties.

An overall governance framework underpins governance, risk, and compliance (GRC) considerations across policy, risk, and compliance layers. Policy governs the design, operation, and control of agentic AI systems and associated services across enterprise and third-party actors. Risk management aligns business objectives with the decisions of AI agents and decision-making systems to mitigate potential adverse impacts. Regulatory compliance ensures adherence to laws and regulations, including privacy, information security, and sector-specific regimes. Assurance collates compliance evidence and provides an independent assessment of GRC metrics through direct or indirect auditing and logging mechanisms. A formal account of responsibilities, spanning all stakeholders and governance bodies, is essential for these GRC elements.

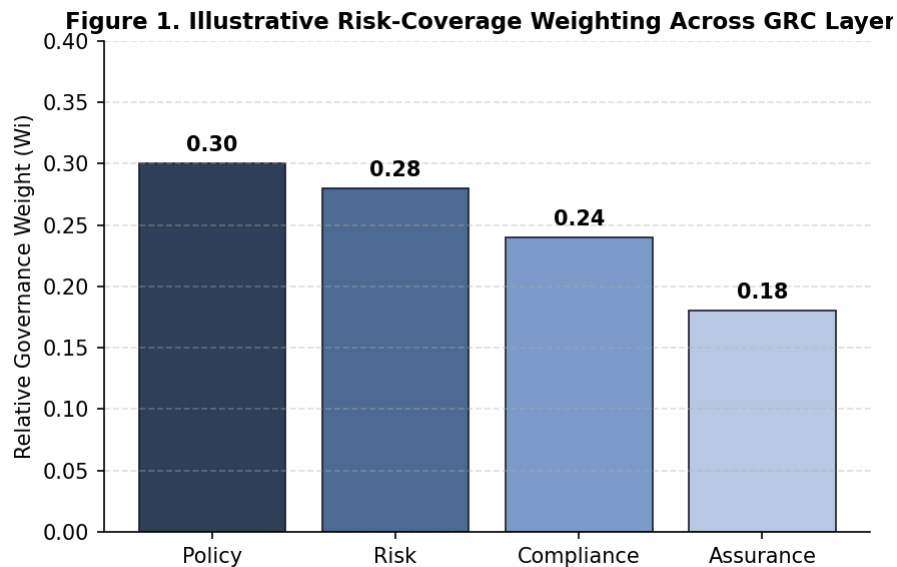


Figure 1. Illustrative risk-coverage weighting across the four GRC layers (Table 1), summing to unity.

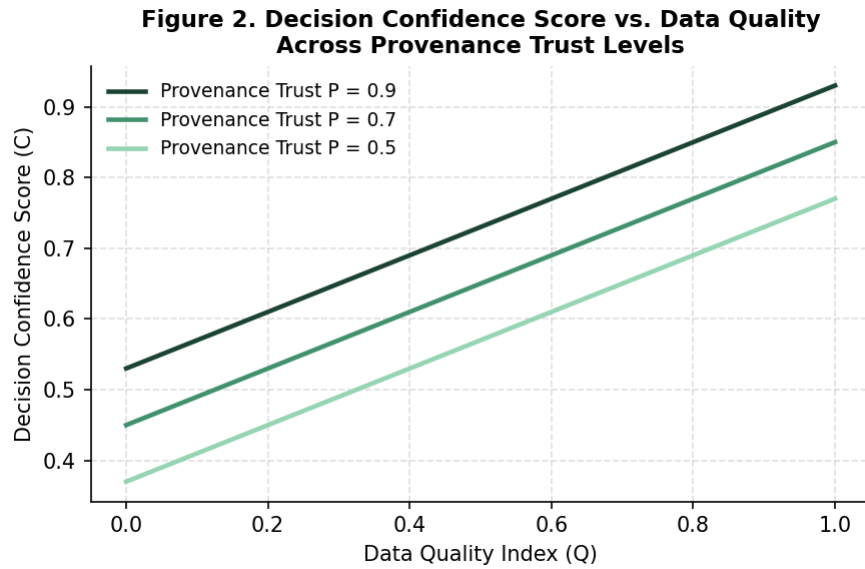


Figure 2. Decision confidence score (Eq. 1) as a function of data-quality index, plotted across three provenance-trust levels.

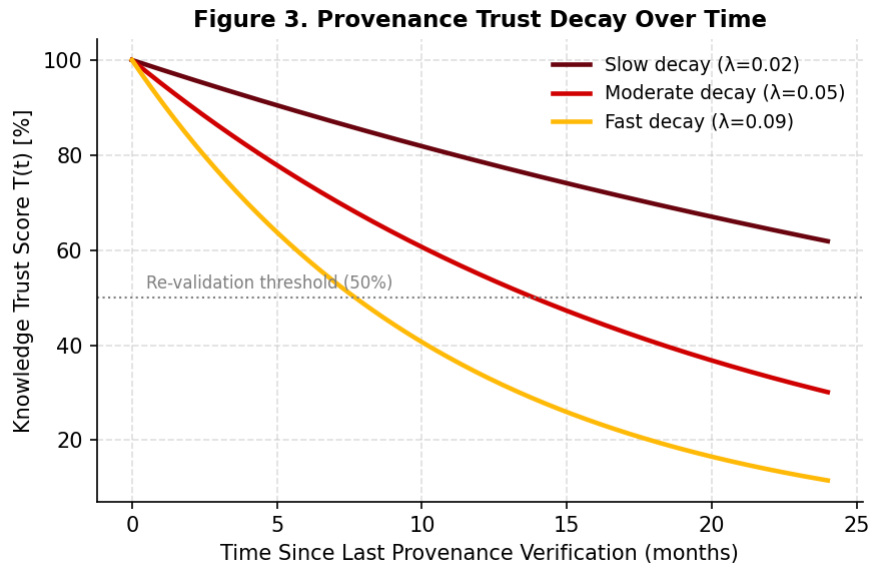


Figure 3. Exponential provenance trust decay (Eq. 3) under slow, moderate, and fast decay regimes, with re-validation threshold marked.



4. Agentic AI for Enterprise Service Automation

The Agentic AI architecture supports an agent-based approach to enterprise service automation. It delegates requests to appropriate resources, orchestrates and monitors execution, and employs runtime learning loops to enhance reliability and performance. Key modules address intent understanding, plan generation, task execution, monitoring, and improvement. The approach connects with existing enterprise resources and service platforms through API wrappers and other agents, facilitating diverse forms of automation and service management. Performance is judged by throughput, reliability, latency, fault tolerance, and accident safeguards.

AI systems are designed to support specific organizational objectives. These objectives inform a risk register that identifies important operational and technology-related risks, risk appetite, and modes of escalation. Efforts are made to ensure enterprise agent for enterprise service automation align with corporate aims. In practical implementations, failure to do so results in partial use of the service.

The enterprise environment provides a compelling business case for automating service delivery. However, current solutions are expensive, insecure, and non-compliant. Agentic AI eases the setup and lowers the cost of enterprise service automation, makes enterprise offerings available for free or at minimal expense, and embeds security controls and compliance in the operating model.

5. Secure Knowledge-Centric Decision Intelligence

Knowledge-centric Decision Intelligence delivers high-impact, reliable answers to critical queries, leveraging a structured pipeline for transforming raw data into decision-ready consumable knowledge. The pipeline encompasses data collection and ingestion; cleaning, augmentation, verification, and supervision (CAVS); provenance tracking; and preparation of a knowledge graph. Trusted data—protected from tampering and encapsulated in a decision knowledge graph—facilitates secure reasoning and answer generation. Security features include access control technologies, data encryption, privacy-by-design strategies, and secure multiparty computation (SMPC) for outsourced diagram construction without compromising privacy.

The KAI requires support for answer generation, level of confidence estimation for answers, and explanation of the underlying reasoning process. The generation mechanism evaluates the associated risk level and, based on defined thresholds and established policy-based controls, assigns trusted confidence levels to answers. For explainability, generating decisions, recommendations, or action plans employs an explainable sub-agent. Since decision systems can become targets for adversarial actors, resilience planning involves adversarial threat modelling, exploration of points of failure in KAI design and implementation, and development of incident response protocols. Grounded in uncertain theory, KAI performance measurement focuses on information trustworthiness and quality, action- or decision-impact analysis, and overall contribution to positive or negative enterprise outcome consequences, allowing lessons learned to form part of knowledge curation.



Mathematical Formulas:

Eq. (1) — Decision Confidence Score:

$$C = w1 \cdot Q + w2 \cdot P + w3 \cdot (1 - R), \text{ with } w1 + w2 + w3 = 1$$

where C is the decision confidence score assigned by the KAI generation mechanism; Q is the data-quality index (0–1) from the CAVS pipeline; P is the provenance trust level of the source knowledge graph segment; R is the residual risk exposure of the query domain; and w1, w2, w3 are policy-defined weighting coefficients.

Eq. (2) — Composite Governance Trust Score:

$$T = \frac{1}{4}(E + Tr + Co + Au)$$

where T is the composite trust score for an automated decision or recommendation, and E, Tr, Co, Au denote the four normalized (0–1) trust propositions: Explainability, Traceability, Coherency, and Auditability.

Eq. (3) — Provenance Trust Decay:

$$T(t) = T0 \cdot e^{(-\lambda t)}$$

where T(t) is the knowledge trust score at time t since last provenance verification, T0 is the initial trust score at t = 0, and λ is the domain-specific decay rate. Re-validation is triggered when T(t) crosses a governance-defined threshold (e.g., 50%).

Eq. (4) — GRC Compliance Index:

$$GCI = (A_p \cdot M_r \cdot C_c)^{(1/3)}$$

where GCI is the composite Governance-Risk-Compliance index (geometric mean, penalizing weak layers more than an arithmetic mean); A_p is policy-adherence coverage; M_r is risk-mitigation coverage; and C_c is regulatory compliance coverage, each normalized to (0,1].

Eq. (5) — Effective Agentic Throughput:

$$E_th = N \cdot (1/L) \cdot (1 - F)$$

where E_th is effective enterprise throughput, N is the number of concurrently active agents, L is mean end-to-end decision latency per task, and F is the observed fault/escalation rate requiring human intervention.

6. Conclusion

Research questions are addressed by presenting an integrated governance framework within enterprise service automation (ESA). The defined framework consists of four layers that facilitate the deployment of agentic AI while managing its inherent risks. Key components include policy formulation for agentic decision making, identification of stakeholder roles and responsibilities, mapping of operational controls, and establishment of audit trails for key system functions. The proposed solution remediates governance gaps that hinder the safe implementation of current and emerging technologies. In addition,



recommendations support compliance with widely adopted regulatory regimes and promote risk-aware usage of agentic AI.

The framework adapts principles of enterprise governance, risk management, and compliance (GRC) theory for the design and operation of agentic systems. Neither the risk management nor compliance layer is part of the agentic AI architecture but must therefore be specified separately. The latter considers the resources and systems required for assurance that agents, service orchestration, and runtime loops operate in accordance with the enterprise's policies, risk appetite, and business objectives. An integrated GRC solution provides the management of enterprise-specific policies, boundaries, and reporting functions while providing sufficient independence from the operational teams.

References

1. Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ... Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901.
2. Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115.
3. Thiebes, S., Lins, S., & Sunyaev, A. (2021). Trustworthy artificial intelligence. *Electronic Markets*, 31, 447–464.
4. Adusupalli, B., Elakkiya, M., Balakrishnan, N., Negi, A. S., Balaram, A., & Rathnakumari, E. A. (2025, May). Decentralized Blockchain Powered Citizen Participation Frameworks for Next Generation Transparent Elections and Trustworthy Governance. In *International Conference on Sustainability Innovation in Computing and Engineering (ICSICE 2024)* (pp. 367-380). Atlantis Press.
5. Gill, A. S., & Germann, S. (2022). Conceptual and normative approaches to AI governance for a global digital ecosystem supportive of the UN Sustainable Development Goals (SDGs). *AI and Ethics*, 2, 293–301.
6. Hickman, E., & Petrin, M. (2021). Trustworthy AI and corporate governance: The EU's Ethics Guidelines for Trustworthy Artificial Intelligence from a company law perspective. *European Business Organization Law Review*, 22, 593–625.
7. Mäntymäki, M., Minkinen, M., Birkstedt, T., & Viljanen, M. (2022). Defining organizational AI governance. *AI and Ethics*, 2, 603–609.
8. Wirtz, B. W., Weyerer, J. C., & Kehl, I. (2022). Governance of artificial intelligence: A risk and guideline-based integrative framework. *Government Information Quarterly*, 39(4), 101685.



9. Devi, M. U., Nandan, B. P., & Muchakhandi, S. (2025, October). Scene Text Recognition in Natural Images Using Convolutional Neural Network with Vision Transformer. In 2025 International Conference on Communication, Computer, and Information Technology (IC3IT) (pp. 1-6). IEEE.
10. Papagiannidis, E., Enhom, I. M., Dremel, C., Mikalef, P., & Krogstie, J. (2023). Toward AI governance: Identifying best practices and potential barriers and outcomes. *Information Systems Frontiers*, 25, 123–141.
11. Birkstedt, T., Minkinen, M., Tandon, A., & Mäntymäki, M. (2023). AI governance: Themes, knowledge gaps and future agendas. *Internet Research*, 33(7), 133–167.
12. Almeida, P. G. R. de, & Santos Júnior, C. D. dos. (2024). Artificial intelligence governance: Understanding how public organizations implement it. *Government Information Quarterly*, 41, 102003.
13. Kumar, B. H., Nuka, S. T., Recharla, M., Chakilam, C., Suura, S. R., & Pandugula, C. (2025, July). Addressing Ethical Challenges in AI-Driven Health Predictions. In 2025 2nd International Conference on Computing and Data Science (ICCDs) (pp. 1-6). IEEE.
14. Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé III, H., & Crawford, K. (2021). Datasheets for datasets. *Communications of the ACM*, 64(12), 86–92.
15. Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q. V., & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, 35.
16. Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C. L., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Simens, M., Askell, A., Welinder, P., Christiano, P., Leike, J., & Lowe, R. (2022). Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 35.
17. Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K. R., & Cao, Y. (2023). ReAct: Synergizing reasoning and acting in language models. *International Conference on Learning Representations*.
18. Mashetty, S. (2025). LEVERAGING DEEP LEARNING, NEURAL NETWORKS, AND DATA ENGINEERING FOR INTELLIGENT MORTGAGE LOAN VALIDATION. *INTERNATIONAL JOURNAL OF SOCIAL SCIENCE & INTERDISCIPLINARY RESEARCH* ISSN: 2277-3630 Impact factor: 8.036, 14(04), 51-65.
19. Schick, T., Dwivedi-Yu, J., Dessì, R., Raileanu, R., Lomeli, M., Zettlemoyer, L., Cancedda, N., & Scialom, T. (2023). Toolformer: Language models can teach themselves to use tools. *Advances in Neural Information Processing Systems*, 36.
20. Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., Chen, Z., Tang, J., Chen, X., Lin, Y., Zhao, W. X., Wei, Z., & Wen, J. (2024). A survey on large language model based autonomous agents. *Frontiers of Computer Science*, 18, 186345.
21. Liu, X., Yu, H., Zhang, H., Xu, Y., Lei, X., Lai, H., Gu, Y., Ding, H., Men, K., Yang, S., Zhang, S., Deng, X., Zeng, A., Du, Z., Zhang, C., Shen, S., Zhang,

- T., Su, Y., Sun, H., ... Tang, J. (2024). AgentBench: Evaluating LLMs as agents. International Conference on Learning Representations.
22. Wu, Q., Bansal, G., Zhang, J., Wu, Y., Li, B., Zhu, E., Jiang, L., Zhang, X., Zhang, S., Awadallah, A., White, R. W., Burger, D., & Wang, C. (2024). AutoGen: Enabling next-gen LLM applications via multi-agent conversation. Proceedings of the Conference on Language Modeling.
23. Jimenez, C. E., Yang, J., Wettig, A., Yao, S., Pei, K., Press, O., & Narasimhan, K. R. (2024). SWE-bench: Can language models resolve real-world GitHub issues? International Conference on Learning Representations.
24. Bommasani, R., Liang, P., & Lee, T. (2023). Holistic evaluation of language models. *Annals of the New York Academy of Sciences*, 1525(1), 140–146.
25. South, T., Marro, S., Hardjono, T., Mahari, R., Deslandes Whitney, C., Chan, A., & Pentland, A. (2025). Position: AI agents need authenticated delegation. Proceedings of the 42nd International Conference on Machine Learning, 267, 82211–82231.
26. Subramanian Sendil Kumar, Sneha Singireddy, Botlagunta Preethish Nandan, Mahesh Recharla, Anil Lokesh Gadi, & Srinivasarao Paleti. (2025). Optimizing Edge Computing for Big Data Processing in Smart Cities . *Metallurgical and Materials Engineering*, 31(3), 31–39. <https://doi.org/10.63278/1317>
27. Chan, A., Wei, K., Huang, S., Rajkumar, N., Perrier, E., Lazar, S., Hadfield, G. K., & Anderljung, M. (2025). Infrastructure for AI agents. Governance of AI Research Institute.
28. Jha, S., Arora, R. R., Watanabe, Y., Yanagawa, T., Chen, Y., Clark, J., Bhavya, B., Verma, M., Kumar, H., Kitahara, H., Zheutlin, N., Takano, S., Pathak, D., George, F., Wu, X., Turkkan, B. O., Vanloo, G., Nidd, M., Dai, T., ... Puri, R. (2025). ITBench: Evaluating AI agents across diverse real-world IT automation tasks. Proceedings of the 42nd International Conference on Machine Learning, 267, 27134–27197.
29. Luo, J., Zhang, W., Yuan, Y., Zhao, Y., Yang, J., Gu, Y., Wu, B., Chen, B., Qiao, Z., Long, Q., Tu, R., Luo, X., Ju, W., Xiao, Z., Wang, Y., Xiao, M., Liu, C., Yuan, J., Zhang, S., ... Zhang, M. (2025). Large language model agent: A survey on methodology, applications and challenges. arXiv.
30. Kolt, N. (2025). Governing AI agents. SSRN Electronic Journal.
31. Bommarito, J., Katz, D. M., & Bommarito, M. J. (2025). Governing AI agents: Risk, compliance, and accountability in law and finance. SSRN Electronic Journal.
32. Mäntymäki, M., Minkinen, M., Birkstedt, T., & Viljanen, M. (2022). Defining organizational AI governance. *AI and Ethics*, 2, 603–609.
33. Papagiannidis, E., Enholm, I. M., Dremel, C., Mikalef, P., & Krogstie, J. (2023). Toward AI governance: Identifying best practices and potential barriers and outcomes. *Information Systems Frontiers*, 25, 123–141.
34. Wirtz, B. W., Weyerer, J. C., & Kehl, I. (2022). Governance of artificial intelligence: A risk and guideline-based integrative framework. *Government Information Quarterly*, 39(4), 101685.



35. South, T., Marro, S., Hardjono, T., Mahari, R., Deslandes Whitney, C., Chan, A., & Pentland, A. (2025). Position: AI agents need authenticated delegation. Proceedings of the 42nd International Conference on Machine Learning, 267, 82211–82231.