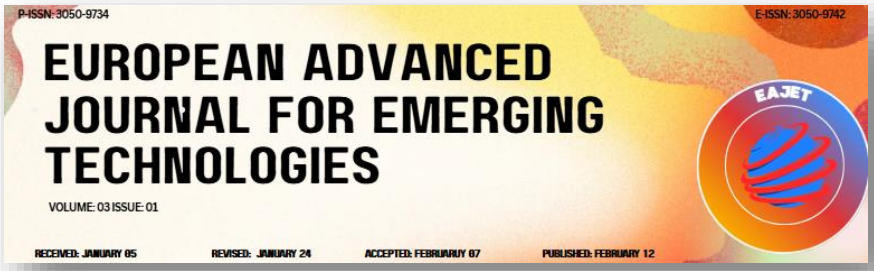


A Trust Architecture for Enterprise Decision Automation

Vikram Boga
Independent Researcher
bogavickram@gmail.com

Abstract : In fast-changing environments, enterprises must adapt quickly to survive. One increasingly common strategy is to automate parts of an organization by delegating business functions to intelligent agents—software systems capable of performing work autonomously. This paper describes an architecture in which such agents operate around a central execution orchestrator and decision workflows. However, a key limitation of current agentic systems is that decision-making speed—critical for throughput and reliability—is often insufficient, resulting in operational delays.

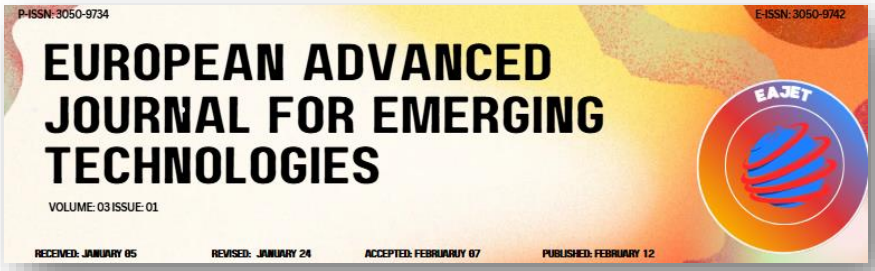
This limitation stems primarily from a gap in the underlying architecture. In enterprise settings, it is essential that the internal reasoning processes of agent-based decision-makers be coordinated through an orchestrator, and that the decision pipeline supporting this reasoning operate with adequate speed. This orchestration layer forms the core of secure, knowledge-centric decision-making: the generation of decisions, recommendations, and predictions that meet the quality and assurance standards required by the task domain. Secure knowledge-centric reasoning thus serves as a necessary counterpart to the agentic foundation being automated—enabling the speed, reliability, and trust required for decisions that directly affect business function execution.

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-



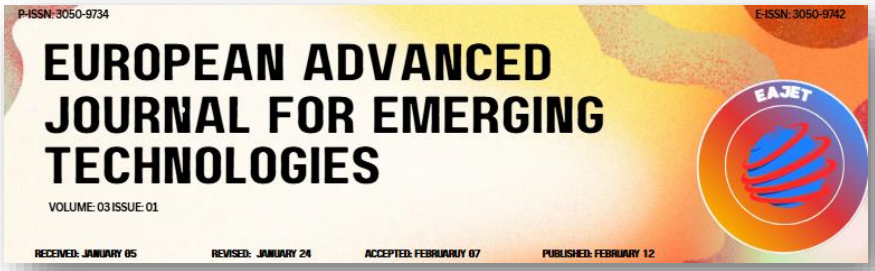
managed, compliant, and assured. The proposed governance architecture offers a comprehensive policy, risk, compliance, and assurance framework that addresses enterprise needs. It establishes cohesive regulatory, risk, and compliance design principles covering the life cycle of such enterprise services, 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



Module	Function	Input	Output
Monitoring	Tracks throughput, latency, faults	Execution telemetry	Performance metrics
Improvement (Runtime Learning)	Adjusts policy/plan based on outcomes	Metrics, escalation history	Updated plan heuristics

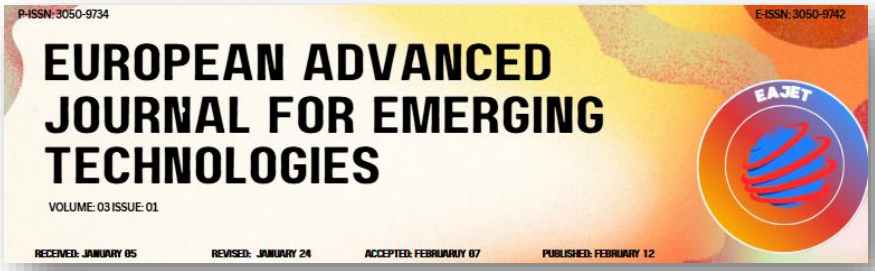
Table 3. KAI Trust Pillars (Section 5)

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.

P-ISSN: 3050-9734
E-ISSN: 3050-9742
EUROPEAN ADVANCED JOURNAL FOR EMERGING TECHNOLOGIES
VOLUME: 03 ISSUE: 01
RECEIVED: JANUARY 05 REVISED: JANUARY 24 ACCEPTED: FEBRUARY 07 PUBLISHED: FEBRUARY 12

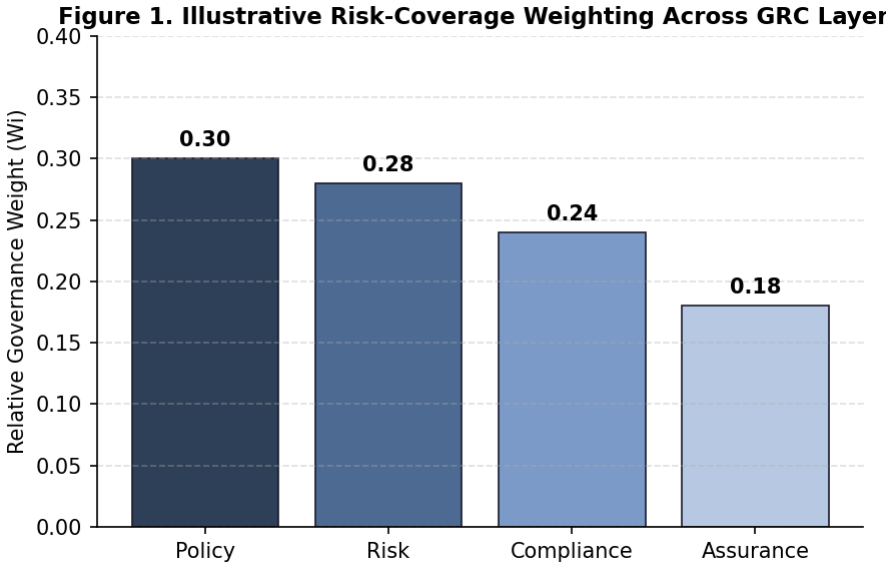


Figure 1. Illustrative risk-coverage weighting across the four GRC layers (Table 1), summing to unity.

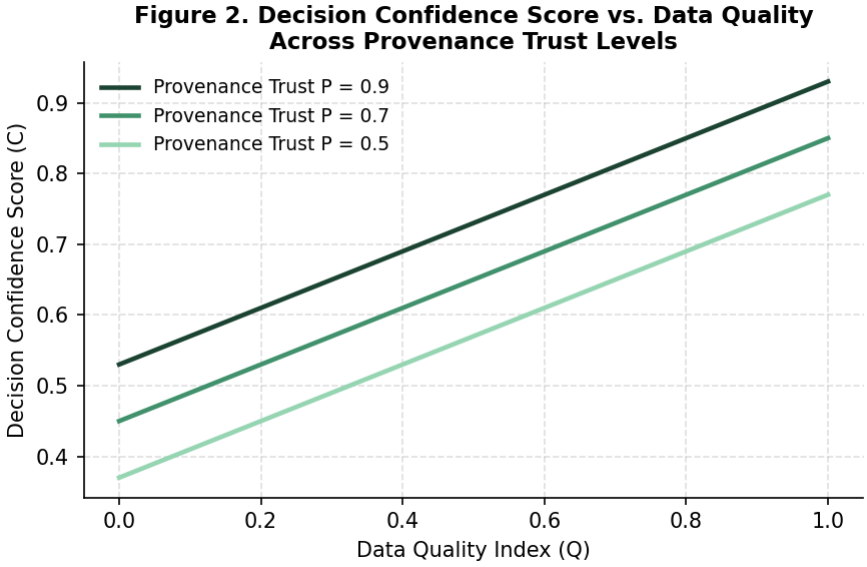


Figure 2. Decision confidence score (Eq. 1) as a function of data-quality index, plotted across three provenance-trust levels.

P-ISSN: 3050-9734 E-ISSN: 3050-9742
EUROPEAN ADVANCED JOURNAL FOR EMERGING TECHNOLOGIES
VOLUME: 03 ISSUE: 01
RECEIVED: JANUARY 05 REVISED: JANUARY 24 ACCEPTED: FEBRUARY 07 PUBLISHED: FEBRUARY 12

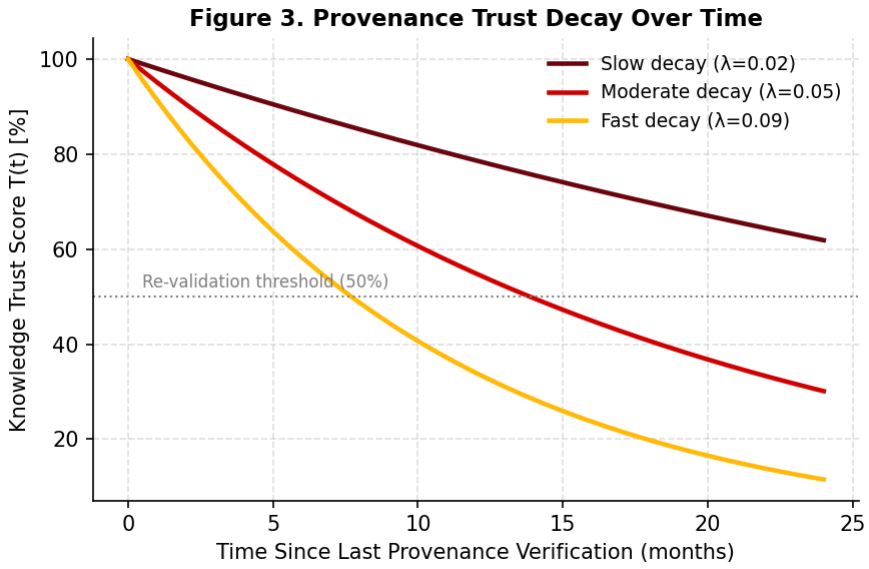


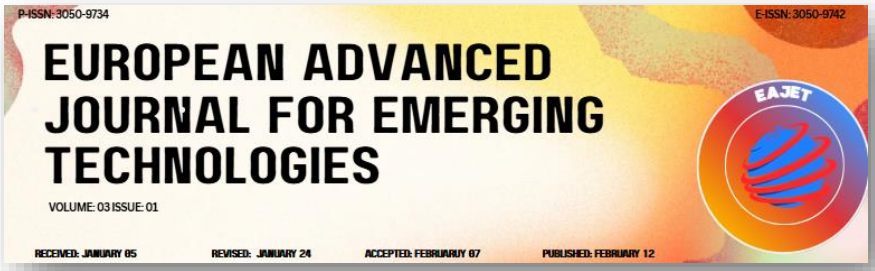
Figure 3. Exponential provenance trust decay (Eq. 3) under slow, moderate, and fast decay regimes, with re-validation threshold marked.

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:

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

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:

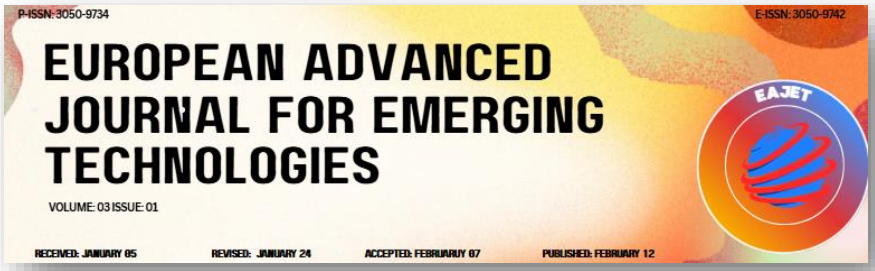
$$T = ¼(E + Tr + Co + Au)$$

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.

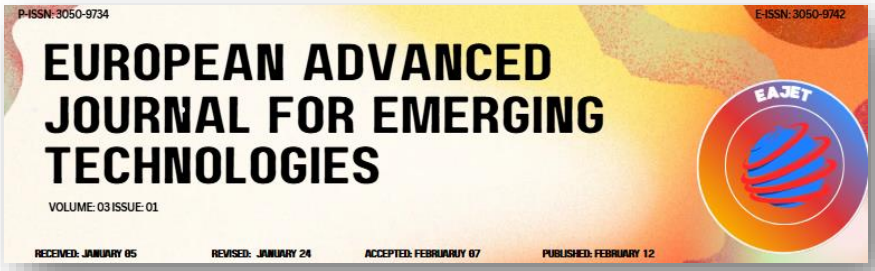
6. Conclusion

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.

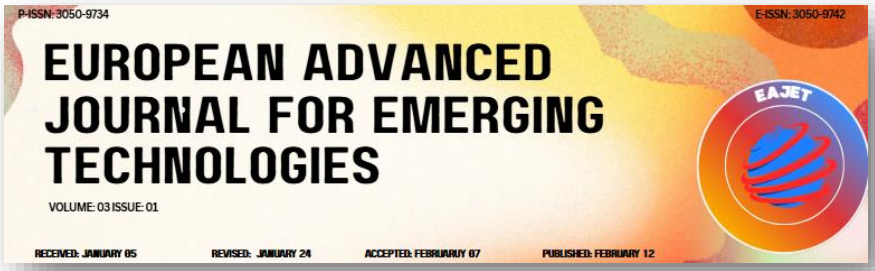
The framework adapts principles of enterprise governance, risk management, and compliance (GRC) theory for the design and operation of agentic systems. Neither the risk management nor compliance layer is part of the agentic AI architecture but must therefore be specified separately. The latter considers 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

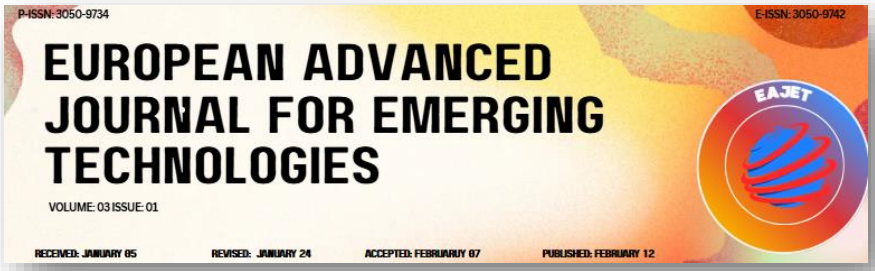
1. Bao, Y., Cheng, X., de Vreede, T., & de Vreede, G. J. (2021). Investigating the relationship between AI and trust in human-AI collaboration. Proceedings of the Annual Hawaii International Conference on System Sciences, 324–333.



2. Chatila, R., Dignum, V., Fisher, M., Giannotti, F., Morik, K., Russell, S., & Yeung, K. (2021). Trustworthy AI. In R. Braunschweig & M. Ghallab (Eds.), Reflections on artificial intelligence for humanity (pp. 13–39). Springer.
3. Floridi, L., Cowls, J., King, T. C., & Taddeo, M. (2021). How to design AI that benefits people and society. Harvard Data Science Review, 3(1).
4. Sondinti, L. R. K., & Pandugula, C. (2023). The Convergence of Artificial Intelligence and Machine Learning in Credit Card Fraud Detection: A Comprehensive Study on Emerging Trends and Advanced Algorithmic Techniques. International Journal of Finance (IJFIN), 36(6), 10-25.
5. 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.
6. Green, B. (2022). The flaws of policies requiring human oversight of government algorithms. Computer Law & Security Review, 45, 105681.
7. Hanif, A., Zhang, X., & Wood, S. (2021). A survey on explainable artificial intelligence techniques and challenges. Proceedings of the 2021 IEEE 25th International Enterprise Distributed Object Computing Conference Workshops, 81–89.
8. Habbal, A., Ali, M. K., & Abuzaraida, M. A. (2024). Artificial Intelligence Trust, Risk and Security Management (AI TRiSM): Frameworks, applications, challenges and future research directions. Expert Systems with Applications, 240, 122442.
9. Kaur, D., Usul, S., Rittichier, K. J., & Durresi, A. (2022). Trustworthy artificial intelligence: A review. ACM Computing Surveys, 55(2), Article 39.
10. Laux, J. (2024). Institutionalised distrust and human oversight of artificial intelligence: Towards a democratic design of AI governance under the European Union AI Act. AI & Society, 39, 2853–2866.
11. Challa, K., Burugulla, J. K. R., Pandiri, L., Pamisetty, V., & Paleti, S. (2022). Optimizing Digital Payment Ecosystems: Ai-Enabled Risk Management, Regulatory Compliance, And Innovation In Financial Services. Migration Letters, 19, 1748-1769.
12. Li, B., Qi, P., Liu, B., Di, S., Liu, J., Pei, J., Yi, J., & Zhou, B. (2021). Trustworthy AI: From principles to practices. arXiv.
13. Mäntymäki, M., Minkkinen, M., Birkstedt, T., & Viljanen, M. (2022). Defining organizational AI governance. AI and Ethics, 2, 603–609.
14. Mäntymäki, M., Birkstedt, T., Minkkinen, M., Tandon, A., & Mäntymäki, M. (2023). AI governance: Themes, knowledge gaps and future agendas. Internet Research.
15. Mökander, J., & Floridi, L. (2022). Operationalising AI governance through ethics-based auditing: An industry case study. AI and Ethics, 3, 451–468.
16. Srinivas Kalisetty, D. A. S. (2024). Leveraging Artificial Intelligence and Machine Learning for Predictive Bid Analysis in Supply Chain Management: A Data-Driven Approach to Optimize Procurement Strategies.
17. Mökander, J., Morley, J., Taddeo, M., & Floridi, L. (2021). Ethics-based auditing to develop trustworthy AI. Minds and Machines, 31, 323–327.
18. 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, 947–960.



19. Puranam, P. (2021). Human–AI collaborative decision-making as an organization design problem. Journal of Organization Design, 10, 75–80.
20. Schiff, D. S., Jackson Schiff, K., & Pierson, P. (2021). Assessing public value failure in government adoption of artificial intelligence. Public Administration, 99, 365–383.
21. Seppälä, A., Birkstedt, T., & Mäntymäki, M. (2021). From ethical AI principles to governed AI. Proceedings of the 42nd International Conference on Information Systems.
22. Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. International Journal of Human-Computer Studies, 146, 102551.
23. Vereschak, O., Bailly, G., & Caramiaux, B. (2021). How to evaluate trust in AI-assisted decision making? A survey of empirical methodologies. Proceedings of the ACM on Human-Computer Interaction, 5(CSCW2), Article 327, 1–39.
24. Wells, L., & Bednarz, T. (2021). Explainable AI and reinforcement learning—A systematic review of current approaches and trends. Frontiers in Artificial Intelligence, 4, 550030.
25. 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.
26. Pamisetty, A. (2024). Leveraging Big Data Engineering for Predictive Analytics in Wholesale Product Logistics. Available at SSRN 5231473.
27. Vyhmeister, E., Castane, G., Östberg, P.-O., & Thevenin, S. (2023). A responsible AI framework: Pipeline contextualisation. AI and Ethics, 3, 175–197.
28. Bork, D., Ali, S. J., & Dinev, G. M. (2023). AI-enhanced hybrid decision management. Business & Information Systems Engineering, 65, 179–199.
29. Westphal, M., Vössing, M., Satzger, G., Yom-Tov, G. B., & Rafaeli, A. (2023). Decision control and explanations in human-AI collaboration: Improving user perceptions and compliance. Computers in Human Behavior, 145, 107714.
30. Kaur, D., Uslu, S., Rittichier, K. J., & Durresi, A. (2022). Trustworthy artificial intelligence: A review. ACM Computing Surveys, 55(2), Article 39.
31. Gianni, R., Lehtinen, S., & Nieminen, M. (2022). Governance of responsible AI: From ethical guidelines to cooperative policies. Frontiers in Computer Science, 4, 873437.
32. Danda, R. R., Yasmeen, Z., & Maguluri, K. K. (2020). AI-driven healthcare transformation: Machine learning, deep learning, and neural networks in insurance and wellness programs. JEC PUBLICATION.
33. Mökander, J., & Floridi, L. (2023). Ethics-based auditing of automated decision-making systems: Intervention points and policy implications. AI & Society, 38, 153–171.
34. Zicari, R. V., Brodersen, J., Brusseau, J., Düdder, B., Eichhorn, T., Ivanov, T., Kararigas, G., Kringen, P., McCullough, M., Möslein, F., Mushtaq, N., Roig, G., Stürtz, N., Tolle, K., Tithi, J. J., van Halem, I., & Westerlund, M. (2021). Z-Inspection®: A process to assess trustworthy AI. IEEE Transactions on Technology and Society, 2(2), 83–97.
35. Zicari, R. V., Ahmed, S., Amann, J., Braun, S. A., Brodersen, J., Bruneault, F., Brusseau, J., Campano, E., Coffee, M., Dengel, A., Düdder, B., Gallucci, A., Gilbert,



T. K., Gottfrois, P., Goffi, E., Haase, C. B., Hagendorff, T., Hickman, E., Hildt, E., Holm, S., Kringen, P., Kühne, U., Lucieri, A., Madai, V. I., Moreno-Sánchez, P. A., Medlicott, O., Ozols, M., Schnebel, E., Spezzatti, A., Tithi, J. J., Umbrello, S., Vetter, D., Volland, H., Westerlund, M., & Wurth, R. (2021). Co-design of a trustworthy AI system in healthcare: Deep learning based skin lesion classifier. Frontiers in Human Dynamics, 3, 688152.

36. Zicari, R. V., Brusseau, J., Blomberg, S. N., Christensen, H. C., Coffee, M., Ganapini, M. B., Gerke, S., Gilbert, T. K., Hickman, E., Hildt, E., Holm, S., Kühne, U., Madai, V. I., Osika, W., Spezzatti, A., Schnebel, E., Tithi, J. J., Vetter, D., Westerlund, M., Wurth, R., Amann, J., Antun, V., Beretta, V., Bruneault, F., Campano, E., Düdder, B., Gallucci, A., Goffi, E., Haase, C. B., Hagendorff, T., Kringen, P., Möslein, F., Ottenheimer, D., Ozols, M., Palazzani, L., Petrin, M., Tafur, K., Tørresen, J., Volland, H., & Kararigas, G. (2021). On assessing trustworthy AI in healthcare: Machine learning as a supportive tool to recognize cardiac arrest in emergency calls. Frontiers in Human Dynamics, 3, 673104.
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, 947–960.
38. Giudici, P., Centurelli, M., & Turchetta, S. (2024). Artificial intelligence risk measurement. Expert Systems with Applications, 235, 121220.
39. Habbal, A., Ali, M. K., & Abuzaraida, M. A. (2024). Artificial Intelligence Trust, Risk and Security Management (AI TRiSM): Frameworks, applications, challenges and future research directions. Expert Systems with Applications, 240, 122442.
40. Thomas, C., Roberts, H., Mökander, J., Tsamados, A., Taddeo, M., & Floridi, L. (2024). The case for a broader approach to AI assurance: Addressing "hidden" harms in the development of artificial intelligence. AI & Society, 40, 1469–1484.