E-ISSN: 1115 - 8530 P-ISSN: 3026-8958
MPRI-JFRBD | INTERNATIONAL JOURNAL OF FINANCIAL RES. AND BUSINESS DEVELOPMENT
VOL. 13 NO. 7 – JULY, 2026

TRUSTWORTHY GENERATIVE AI FOR EXECUTIVE DECISION-MAKING (TGAI-EDM) FRAMEWORK

AMAGWULA DANIEL EBUKA
Department of International Business Management, Buckinghamshire New University.
danielamagwula@gmail.com
DOI: https://doi.org/10.70382/mejfrbd.v13i7.123

Abstract

Generative Artificial Intelligence (GenAI) is rapidly transforming executive decision-making in enterprises, introducing both unprecedented opportunities and critical risks. This paper proposes a Trustworthy Generative AI for Executive Decision-Making (TGAI-EDM) framework designed to enable responsible, transparent, and accountable AI adoption in intelligent enterprises. The framework integrates governance-by-design, explainability, human-AI collaboration, risk-adaptive orchestration, and continuous assurance mechanisms. It aligns with emerging global standards including the EU AI Act (2024), NIST AI RMF 1.0 (2023), and ISO/IEC 42001:2023. The study contributes a layered architectural model that embeds trust directly into executive decision workflows rather than treating governance as a post-deployment control mechanism.

Keywords:
Generative AI, Decision Making, Trustworthy, Risk Management, Governance Mechanisms

Introduction

The adoption of Generative Artificial Intelligence marks a fundamental shift in how enterprises create, process, and act on information. Unlike earlier forms of automation that followed explicit rules, López-Solís et al. (2025), maintained that GenAI systems such as large language models and multimodal foundation models can synthesize new content, reason across ambiguous data, and propose strategic alternatives with minimal human prompting. In boardrooms and C-suites, this capability is moving GenAI beyond support roles in marketing copy or customer service and into the

EDITOR-IN-CHIEF
Prof. R. Rena [South Africa]

core of executive decision-making. Executives according to Saari (2026) are now using AI to generate market entry scenarios, model the financial impact of geopolitical events, draft M&A theses, and simulate competitor responses in near real time. This transition creates what can be called a "cognitive partnership" between humans and machines (Saup et al., 2026). In this partnership, GenAI acts as a pattern-recognition engine and scenario generator, while human leaders provide judgment, context, and accountability (Bick et al., 2024). The promise is clear: faster insight generation, broader exploration of strategic options, and the ability to process data volumes that exceed human capacity. For multinational enterprises operating in volatile markets, this speed can be a decisive competitive advantage.

However, the same properties that make GenAI powerful also make it risky for executive use. First, GenAI outputs are probabilistic. Two prompts with the same intent can yield different answers, and the system cannot guarantee factual accuracy without external grounding. Second, these models are prone to hallucination, bias inherited from training data, and subtle errors that are hard to detect without domain expertise. Third, the lack of deterministic logic makes traditional software validation methods inadequate (Sajadieh et al. 2026). You cannot "unit test" a strategic recommendation the way you test an accounting rule. These technical challenges intersect with governance and legal challenges. The EU AI Act (2024) classifies AI used in employment, credit, and critical infrastructure as high-risk and imposes strict requirements for documentation, human oversight, and conformity assessment. Bonfigliol et al. (2024) calls for organizations to map risks, measure impacts, and manage them across the AI lifecycle. ISO/IEC 42001:2023 further requires an auditable AI management system. In this regulatory environment, an executive who relies on an AI-generated recommendation must be able to answer three questions: Where did this insight come from? How confident are we in it? And who is responsible if it is wrong?

Despite this, many organizations are deploying GenAI in an ad hoc manner. Governance is often added after deployment, as a compliance checklist. This approach creates a trust deficit. Executives hesitate to rely on AI, legal teams block its use, and the organization fails to capture value. The central problem this paper addresses is therefore: how can enterprises embed trustworthiness into GenAI-assisted executive decision-making from the start, without sacrificing agility? We argue that trust cannot be bolted on. It must be designed into the workflow through layered controls, clear human-AI roles, and continuous monitoring (Brynjolfsson & Raymond, 2025). To address this, there is need to propose the Trustworthy Generative AI for Executive Decision-Making (TGAI-EDM) framework. TGAI-EDM treats governance as an operating system for decision-making, not as a separate audit function. It provides a practical architecture that aligns technical safeguards with organizational processes and regulatory requirements. The goal is to

enable executives to use GenAI with justified confidence, knowing that risks are managed, decisions are traceable, and accountability is preserved.

Literature Review and Theoretical Foundations

The academic and policy literature on AI governance has grown rapidly since 2020, driven by the deployment of large language models in enterprise settings. Three streams of research are most relevant to executive decision-making: AI risk management, generative AI safety, and hybrid intelligence.

The first stream is anchored by formal standards. Eloundou et al. (2023) provides a voluntary but widely adopted structure organized around four functions: Govern, Map, Measure, and Manage. It emphasizes that AI risk is socio-technical and must be addressed at the organizational level, not just the model level. ISO/IEC 42001:2023 complements this by specifying requirements for an AI management system, including policy, resources, and continual improvement. These standards share a core principle: accountability requires documentation and oversight throughout the lifecycle.

The second stream addresses the unique risks of generative AI. Studies from 2023 to 2026 have documented three persistent problems. Lehman et al. (2023), hallucination remains common even in state-of-the-art models, especially when asked to synthesize information outside their training cutoff. Bias amplification occurs when models reproduce and exaggerate patterns in corporate or societal data, leading to discriminatory recommendations in hiring or lending. Data leakage and prompt injection present new security vectors, where sensitive enterprise data can be exposed through model outputs or adversarial prompts. Regulatory responses have followed quickly. The EU AI Act (2024) imposes the strictest obligations on high-risk AI systems, requiring risk management systems, data governance, technical documentation, record-keeping, transparency, human oversight, and accuracy.

The third stream is hybrid intelligence. This concept rejects both full automation and full human control in favor of collaboration. Research in organizational behavior shows that executives perform best when AI handles information synthesis and scenario generation, while humans handle value judgments and final approval (Levine et al. 2026, Patel & Dey, 2026). Key design principles from this literature include maintaining human agency, providing explainable outputs, and designing interfaces that make uncertainty visible (Page & Kallapur, 2025). Without these, automation bias sets in: decision-makers over-rely on AI because it is fast and confident, even when it is wrong (Li & Tian, 2026). A gap remains. Existing frameworks are strong on principles but weak on operationalization for the executive context (Mishra et al. 2023). They do not specify how to embed governance into the moment of decision. They also rarely address the political reality of the C-suite, where speed and competitive advantage are prioritized. TGAI-EDM builds on these

foundations but adds a workflow-level architecture, risk-adaptive controls, and metrics that executives can actually use.

Proposed Framework: TGAI-EDM

The Trustworthy Generative AI for Executive Decision-Making framework is a seven-layer architecture. Each layer embeds specific controls into the decision workflow.

Layer 1: Governance and Policy Layer. This layer defines who can use GenAI for what decisions, and under what conditions. It includes a board-approved AI use policy, a catalog of pre-approved use cases, and a decision authority matrix. High-impact decisions such as capital allocation or workforce reduction are automatically classified as high-risk and require additional controls. This aligns with ISO 42001 leadership requirements and EU AI Act risk classification.

Layer 2: Data Integrity and Provenance Layer. Executive decisions are only as good as the data behind them. This layer ensures all data sources used by GenAI are catalogued, with lineage and freshness tracked. Retrieval-Augmented Generation is used to ground model outputs in enterprise documents, and all sources are cited. PII is redacted, and access controls prevent unauthorized data from entering prompts.

Layer 3: Model Trust and Validation Layer. Before any output reaches an executive, it is scored on multiple dimensions: factual consistency, bias, and hallucination risk. Confidence scores are calibrated, and counterfactual tests are run to check robustness. A model card is attached to each recommendation, documenting limitations.

Layer 4: Decision Orchestration and Human-AI Collaboration Layer. This is the interaction layer. The system presents 3 to 5 scenarios, not a single answer. Each scenario includes assumptions, risks, and a confidence score. A human-in-the-loop validation is required for high-risk decisions. All overrides are logged with rationale.

Layer 5: Risk-Adaptive Controls Layer. Not all decisions need the same level of scrutiny. This layer dynamically adjusts governance based on impact and uncertainty. For example, a routine budget forecast may need light review, while an acquisition recommendation triggers full audit and legal review.

Layer 6: Continuous Assurance and Monitoring Layer. After a decision is made, outcomes are tracked and compared to predictions. Drift detection monitors if model performance degrades. Audit logs are maintained for regulatory review.

Layer 7: Organizational Enablement and Culture Layer. Technology alone is not enough. This layer includes executive training on AI literacy, incentives that reward responsible use, and ethics workshops. The goal is to build a culture where questioning AI is expected. Together, these layers ensure that trust is built into every step, from data to decision to review.

EDITOR-IN-CHIEF
Prof. R. Rena [South Africa]

AI-Augmented Executive Decision-Making Model

Traditional executive decision-making follows a linear path: gather data, analyze options, consult experts, and decide. GenAI compresses and expands this path. It can gather and synthesize data in minutes and generate options that a human team might never consider. But it also introduces new failure points.

The TGAI-EDM decision loop has six stages.

Stage 1: Sense. The AI continuously monitors internal data such as sales, HR, and finance, plus external data such as news, regulations, and competitor moves. It flags anomalies and strategic signals.

Stage 2: Generate. When an executive poses a question, the AI produces multiple scenarios. For example, for market entry, it might generate an organic growth scenario, an acquisition scenario, and a partnership scenario. Each includes projected outcomes and key assumptions.

Stage 3: Govern. Automated filters check each scenario against policy and risk rules. Does it violate sanctions? Does it exceed budget authority? High-risk items are flagged.

Stage 4: Validate. The executive reviews the scenarios with an interface that shows data sources, confidence levels, and reasoning traces. The executive can ask follow-up questions and probe the logic.

Stage 5: Decide and Log. The final decision is recorded with full metadata: which AI version was used, what data was retrieved, what the confidence was, and what human judgment was applied.

Stage 6: Learn. After the outcome is known, the system compares prediction to result and updates its models and risk parameters.

This loop preserves human authority while leveraging AI speed. It also creates an audit trail that satisfies regulators and boards.

Risk Management and Governance Mechanisms

Five core risks dominate GenAI in executive contexts.

Hallucination is mitigated through RAG, source citation, and validation against trusted databases. Bias is addressed through regular dataset audits, fairness metrics, and diverse testing teams. Model drift is caught via continuous benchmarking against historical baselines. Over-reliance is prevented by confidence thresholds and mandatory human review for decisions above a certain financial or reputational threshold. Data leakage is controlled with enterprise guardrails that block sensitive information from prompts and outputs. Governance is mapped to regulation. For the EU AI Act, we implement conformity assessments and documentation. For NIST, we operationalize the Map-Measure-Manage cycle. For ISO 42001, we maintain an AI management system with clear roles and reviews.

EDITOR-IN-CHIEF
Prof. R. Rena [South Africa]

Conclusion

Generative AI will not replace executive judgment. It will reshape it. The organizations that win will be those that learn to partner with AI responsibly. The TGAI-EDM framework offers a path to do that by embedding trust into the workflow rather than adding it later. By combining governance, data integrity, model validation, human collaboration, adaptive risk controls, continuous assurance, and cultural enablement, enterprises can use GenAI for strategic decisions with confidence and accountability. The framework is aligned with current global standards and designed to evolve as regulation and technology change. Future research should test TGAI-EDM in live enterprise environments and develop industry-specific variants. The goal is not to slow innovation, but to make innovation sustainable and trustworthy.

REFERENCES

Bick, A., Blandin, A., & Deming, D. J. (2024). *The rapid adoption of generative AI* (NBER Working Paper No. 32966). National Bureau of Economic Research. https://doi.org/10.3386/w32966

Bonfiglioli, A., Crinò, R., Filomena, M., & Gancia, G. (2025). *Comparative advantage in AI-intensive industries: Evidence from US imports* (CEPR Discussion Paper No. 19883). CEPR Press.

Brynjolfsson, E., Li, D., & Raymond, L. R. (2025). Generative AI at work. *The Quarterly Journal of Economics, 140*(2), 889–942. https://doi.org/10.1093/qje/qjae044

Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2023). *GPTs are GPTs: An early look at the labor market impact potential of large language models*. arXiv. https://arxiv.org/abs/2303.10130

European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) *Official Journal of the European Union*. https://eur-lex.europa.eu/eli/reg/2024/1689

Lehman, E., Hernandez, E., Mahajan, D., Wulff, J., Smith, M. J., Ziegler, Z., Nadler, D., Szolovits, P., Johnson, A., & Alsentzer, E. (2023). *Do we still need clinical language models?* arXiv. https://arxiv.org/abs/2302.08091

Levine, J., Goldberg, T., Teng Wade, J., Sukin, A., Nagda, B., & Scheller, J. (2026, April 16) . Bessemer predicts: Robotics and physical AI. Bessemer Venture Partners. https://www.bvp.com/atlas/bessemer-predicts-robotics-and-physical-ai

Li, H., & Tian, F. (2026). Advancing decision-making through AI-human collaboration: A systematic review and conceptual framework. *Group Decision and Negotiation*, 35(26). https://doi.org/10.1007/s10726-026-09980-1

López-Solís, O., Luzuriaga-Jaramillo, A., Bedoya-Jara, M., Naranjo-Santamaría, J., Bonilla-Jurado, D., & Acosta-Vargas, P. (2025). Effect of generative artificial intelligence on strategic decision-making in entrepreneurial business initiatives: A systematic literature review. *Administrative Sciences*, 15(2), 66. https://doi.org/10.3390/admsci15020066

Mishra, S., Koopman, R., De Prato, G., Rao, A., Osorio-Rodarte, I., Kim, J., Spatafora, N., Strier, K., & Zaccaria, A. (2023). AI specialization for pathways of economic diversification. *Scientific Reports, 13*, Article 19475. https://doi.org/10.1038/s41598-023-45723-x

Page, S. E., & Kallapur, A. (2025). Replace, augment, disrupt: AI & organizational decision-making. *Journal of Organization Design*. https://doi.org/10.1007/s41469-025-00194-4

Patel, P., & Dey, M. (2026). Generative AI and organisational collective intelligence: A dependency-structured framework. *Business Information Review*, 43(1). https://doi.org/10.1177/02663821261429686

Saari, L. J. M. (2026). Capital ecosystem of European AI: Patriotic billionaires, development banks, and the evolution of state-finance nexus. *Competition & Change*. *Advance online publication.* https://doi.org/10.1177/10245294261429545

EDITOR-IN-CHIEF
Prof. R. Rena [South Africa]

Sajadieh, S., Fattorini, L., Perrault, R., Gil, Y., Parli, V., Santarlasci, L., Pava, J., Maslej, N., Altman, R., Brynjolfsson, E., Brodley, C., Clark, J., Dignum, V., Kumar, V., Landay, J., Lyons, T., Manyika, J., Niebles, J. C., Shoham, Y., Tabassi, E., Wald, R., Walsh, T., & Weld, D. (2026). *The AI Index 2026 annual report*. AI Index Steering Committee, Institute for Human-Centered AI, Stanford University.

Saup, T.-O., Asghar, J., Kanbach, D. K., & Kraus, S. (2026). From pilots to decision systems: Embedding generative AI into strategic decision-making through a socio-technical and governance lens. *Journal of Decision Systems,* 35(1). https://doi.org/10.1080/12460125.2025.2597835

Sharma, S., Esposito, M., & Akbar, Y. H. (2025). Reframing strategic cognition: A model of AI augmented sensemaking and foresight. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.5240928

Übellacker, T. (2025). Making sense of AI limitations: How individual perceptions shape organizational readiness for AI adoption. *arXiv.* https://arxiv.org/abs/2502.15870

EDITOR-IN-CHIEF
Prof. R. Rena [South Africa]