Journal of
Information Technology,
Cybersecurity, and Artificial Intelligence

(RESEARCH)

Use of Generative Artificial Intelligence for Managerial Verification in Multinational Contract Management.

An Integrative Review and Inductive Study

Msc. Fernando Ivan Jaimes Rada
MR & Lord

Msc. Julia Enith Herrera Mendoza

Journal of Information Technology, Cybersecurity, and Artificial Intelligence, 2026, 3(3), 31-42

Article DOI: https://doi.org/10.70715/jitcai.2026.v3.i3.065

Abstract

The expansion of organizational activities across national boundaries has intensified regulatory compliance challenges due to heterogeneous legal, tax, and health, safety, and environmental (HSE) frameworks across jurisdictions. While large multinational corporations typically rely on permanent and specialized compliance structures, small and medium-sized companies often face resource constraints that limit their internal verification capacity.

This study examines how generative artificial intelligence (AI), specifically Microsoft Copilot, can be integrated into managerial verification processes to strengthen risk management and regulatory compliance in multinational projects. An integrative literature review, combined with an inductive qualitative case study, was conducted using data from 11 projects across Europe, Asia, South America, and Central America. The empirical setting highlights the challenges posed by heterogeneous regulatory environments and underscores the need for efficient oversight of contracts and compliance.

The findings show that generative AI, when properly contextualized and iteratively refined, enhances the precision and consistency of managerial verification, thereby reducing incoherence and regulatory misalignment. The study provides evidence that generative AI integrated into office productivity platforms and governed by human-in-the-loop frameworks serves as a cost-efficient tool for reinforcing internal compliance controls in small and medium-sized multinational companies, offering managers a structured approach to navigating complex, multi-country regulatory contexts.

Keywords: Artificial intelligence; Risk management; Managerial decision-making; Regulatory compliance; Governance; Multinational companies.

1. Introduction.

AI has progressively become integrated into organizational decision-making processes, supporting activities at operational, tactical, and strategic levels. Current applications frequently involve predictive analytics, classification, and knowledge management, thereby extending AI's role in scenario analysis and compliance-related verification tasks.

Within multinational organizations, regulatory compliance poses a persistent managerial challenge due to the coexistence of diverse, sometimes conflicting legal frameworks. Because comprehensive expertise across all relevant jurisdictions is rarely feasible at the individual level, organizations typically combine localized external advisory services with internal verification mechanisms. In this context, contract managers are required to adapt traditional review practices by incorporating technological instruments that strengthen oversight, consistency, and feedback.

The growing availability of generative AI tools integrated into office productivity environments, such as Microsoft 365, has expanded access to advanced analytical capabilities. However, their use in compliance-sensitive contexts requires carefully designed adoption models that preserve human judgment as the ultimate locus of responsibility.

---

2. Methodology.

This study adopts a qualitative, exploratory, and inductive research design. The methodological framework integrates two complementary components:

- An integrative literature review, conducted with systematic rigor, was conducted to establish the conceptual and regulatory foundations of the study.
- An inductive empirical case study, applied to multinational projects, to generate descriptive evidence on the role of generative AI in managerial verification.

Together, these components provide a complementary contribution: the integrative literature review establishes the conceptual and regulatory grounding of the study. It identifies existing gaps in the literature, while the empirical case study addresses these gaps by examining how generative AI is operationalized in managerial verification practices within multinational projects.

2.1. Integrative review.

The review comprised two complementary components: a narrative synthesis and a systematic review. The narrative synthesis was used to contextualize the use of artificial intelligence in risk management within a multinational company, incorporating business and contractual documents, as well as the management system's ethical and procedural guidelines. The systematic review was conducted in accordance with the PRISMA 2020 framework. It included explicit inclusion and exclusion criteria, a structured search of accredited sources using predefined keywords and time limits (2021–2025), and the refinement of an initial set of records to a final sample of eight articles. Although a meta-analysis was not performed, the systematic review provided rigor in the selection and organization of the evidence. At the same time, the narrative synthesis enabled the integration of findings and practices within a broader research context, ensuring that both components remained methodologically distinct yet analytically complementary.

2.1.1. Narrative review and synthesis.

Company documents were incorporated into the analysis; however, for confidentiality reasons, their links or references are not disclosed. These documents mainly comprised internal governance policies, management system guidelines, and corporate compliance frameworks relevant to risk and regulatory oversight.

2.1.2. Systematic review.

The literature inclusion criteria focused on articles published in high-quality, accredited sources, such as peer-reviewed academic journals, and were defined according to the following elements:

2.1.3. Sources and scope.

The literature search was conducted in accredited academic databases and publisher platforms, including Emerald, Springer, Frontiers, IACIS, and SAGE, focusing on peer-reviewed journals and conference proceedings. The study began with a set of academic articles with DOIs addressing AI in managerial decision-making between 2023 and 2025, supplemented by a bibliometric analysis (2024) and a conceptual framework (2024–2025). Journals and conferences from recognized publishers (Emerald, Springer, Frontiers, IACIS, and SAGE) were included.

Search strategy.

Keyword combinations included "artificial intelligence" and "managerial decision-making", filtered by DOI. Searches were restricted to publications between 2021 and 2025, with preference for the period 2023–2025 to capture recent developments. Metadata (title, year, DOI, and publisher) were verified for accuracy.

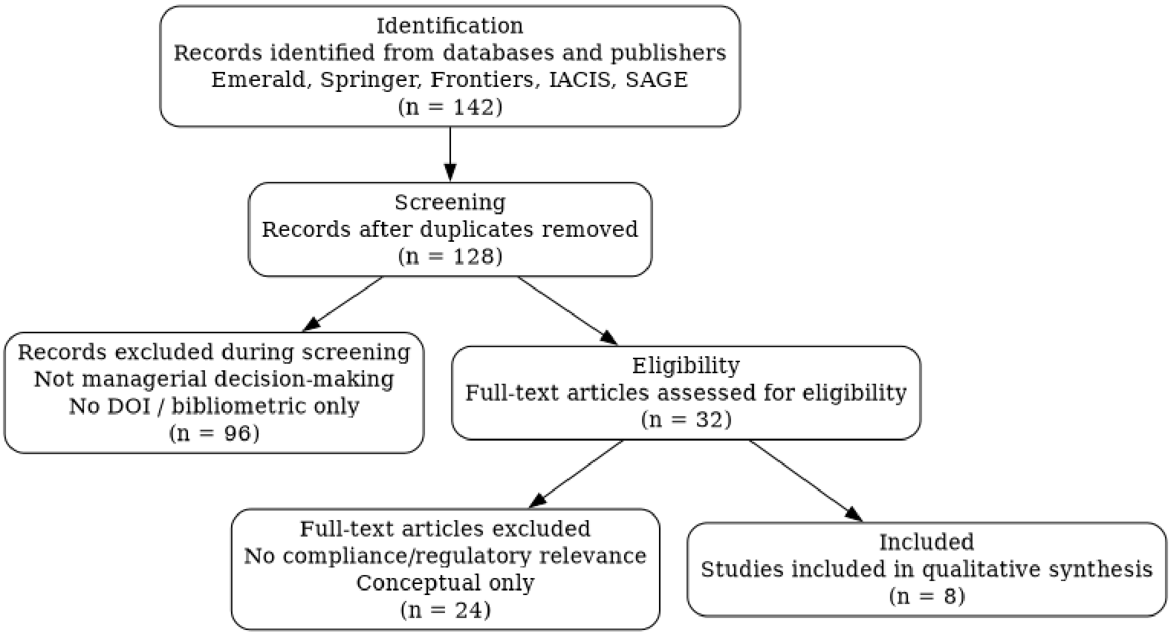
Selection Criteria.

Inclusion criteria comprised articles with a DOI, an explicit link to managerial decision-making (operational, strategic, or functional domains such as HR, finance, or trading), and a discussion of ethical or regulatory implications.

Exclusion criteria included studies without a DOI, works unrelated to managerial decision-making, duplicates, and purely bibliometric analyses without legal or compliance relevance.

The selected articles were further classified by (a) type of contribution (review, empirical, conceptual, or case-based studies) and (b) the presence of legal implications (direct, indirect, or contextual). Priority was given to studies with explicit or inferable legal and regulatory repercussions.

Figure 1. PRISMA 2020 flowchart of the literature review.



2.2. Inductive study.

2.2.1. Level constructs and anchors.

Information quality was assessed by manually verifying the sources and citations provided by the AI in its responses.

Quality of Argumentation was evaluated based on logical coherence, the use of evidence, and the presence of counter-argumentation.

Traceability was examined by reviewing recorded decision steps, iterations, and sources, to ensure a transparent workflow that links AI interactions to final outputs.

The validation rubric was applied across the three analytical dimensions of information quality, argument quality, and traceability, in accordance with the study's design.

Table 1: Analytical evaluation rubric - Evaluation criteria for Quality with three levels (3-Good, 2-Acceptable, 1-Low).

Level	Descriptor	Observable criteria
3	Good	Adequate alignment with contextual constraints; accurate references that passed manual review.
2	Acceptable	Poor alignment with contextual constraints; references are partially accurate, and some passed manual review.
1	Low	Unjustified alignment with contextual constraints; inaccurate references that failed manual review

Table 2: Analytical evaluation rubric - Evaluation criteria for Argumentation with three levels (3-Good, 2-Acceptable, 1-Low).

Level	Descriptor	Observable criteria

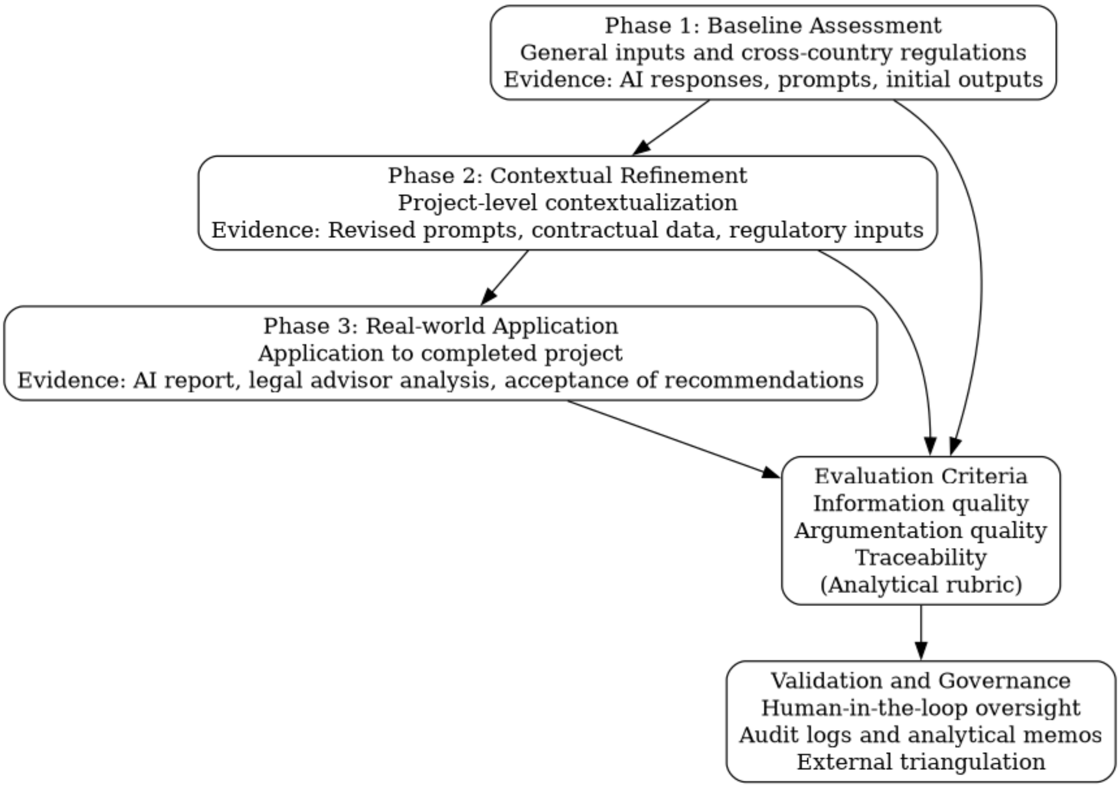
3	Good	Logically consistent reasoning; sufficient, real, and relevant references; good Argumentation.
2	Acceptable	Logically consistent reasoning; sufficient partial references and some irrelevant ones; good Argumentation.
1	Low	Poor logical structure; insufficient and irrelevant references; poor Argumentation.

Table 3: Analytical evaluation rubric - Evaluation criteria for traceability with three levels (3-Good, 2-Acceptable, 1-Low).

Level	Descriptor	Observable criteria
3	Good	Adequate correlation between the situation presented and the AI's response, and between the AI's response and the real response from another company used as an example for the practical exercise; appropriate use of references and information provided within the prompt and the contextual inputs supplied.
2	Acceptable	Adequate correlation between the situation presented and the response given by the AI and the real response given by another company taken as an example for the practical exercise; partially adequate use of the references and information provided within the prompt and the contextual inputs supplied.
1	Low	Inadequate correlation between the situation presented and the response given by the AI and the real response given by another company taken as an example for the practical exercise; inadequate use of the references and information provided within the prompt and the contextual inputs supplied.

Reflexive Thematic Analysis (RTA) emphasized transparency through the use of analytical memos and audit logs rather than reliability metrics. Content validity and traceability were supported through triangulation using documented processes from other organizations, real-world cases treated as hypothetical exercises, and publicly available reports and information published on official web pages.

Figure 2. Evaluation flow by phases.



2.2.2. Ethical considerations.

No confidential or personally identifiable information was collected. All documents consisted of work products used solely for research purposes. The data were securely stored and analyzed in aggregate form, and both the company and the projects were anonymized.

2.2.3. Data adequacy and arrest criteria.

The study was conducted over twelve months, from January to December 2025. During the first eleven months, the AI system was iteratively contextualized and progressively refined through repeated, project-specific interactions, with sampling continuing until the outputs exhibited a decline in novelty and informational value. In the final month (December 2025), the system was applied to a real-world case study, and the evaluation was concluded upon completion of the project. The AI system was not trained or tested in a technical sense; rather, it was progressively contextualized through iterative, scenario-based interactions under human supervision.

2.2.4. Limits of evidence and non-inferential stance.

The findings are observational and descriptive; therefore, no causal claims are made. Statements are presented as patterns associated with instrumental use rather than as proven effects. The analysis is grounded in project implementation contexts and specific empirical observations, including reports, evaluations, and regulatory documentation, which may limit the generalizability of the findings to other settings.

Empirical evidence was drawn from a purposive sample of eleven projects implemented across five countries under the responsibility of a single contract manager. Data sources included contractual records, client-facing reports, internal progress documentation, tax and financial materials, and jurisdiction-specific regulatory requirements. Case selection emphasized geographic diversity and variation in regulatory exposure. No inferential statistical techniques were applied; instead, the analysis focused on identifying descriptive patterns that support analytical rather than statistical generalization.

Table 4: Description of variables.

Variable	Description / Value
Country	Colombia
Public Client	1
Private Client	1
Number of projects	2
Country	Malaysia
Public Client	0
Private Client	2
Number of projects	2
Country	Costa Rica
Public Client	1
Private Client	0
Number of projects	1
Country	United Arab Emirates
Public Client	1
Private Client	0
Number of projects	1
Country	Belgium
Public Client	0
Private Client	5
Number of projects	5
Total number of projects	11

The application of AI followed three sequential stages: an initial baseline phase, a project-level contextualization phase, and a final real-world application. System performance was assessed at each stage using an analytical rubric addressing information quality, argumentative coherence, and traceability. This phased approach enabled observation of changes in AI performance over time while consistently maintaining human supervision as the guiding mechanism for decision-making.

3. Results.

3.1. Results of integrative review.

From the initial set of identified records, eight articles were retained after applying the inclusion and exclusion criteria. This limited number reflects the specific focus on managerial decision-making with explicit compliance or regulatory implications, rather than on general applications of artificial intelligence.

During the identification phase, 142 records were retrieved from academic databases and publisher platforms, including Emerald, Springer, Frontiers, IACIS, and SAGE. After duplicate removal, 128 records remained and were screened by title and abstract. Of these, 96 records were excluded because they did not address managerial decision-making, lacked a DOI, or consisted solely of bibliometric analyses.

In the eligibility phase, 32 full-text articles were assessed for relevance. Following full-text review, 24 articles were excluded due to the absence of explicit compliance or regulatory relevance or because they were purely conceptual.

Ultimately, eight studies were included in the qualitative synthesis of the integrative review.

Table 5: Articles included:

Reference (DOI)	Type	Decisional domain
https://doi.org/10.1108/MD-08-2023-1331	Structured Review (SLR)	Managerial (forecasting, classification, knowledge)
https://doi.org/10.7494/manage.2025.26.1.77	Quick review	General management (AI adoption)
https://doi.org/10.48009/4_iis_2024_116	Conceptual with cases	HR, credit, trading
https://doi.org/10.1007/s00187-025-00396-7	Experimental	Operational and strategic
https://doi.org/10.3389/forgp.2025.1419403	Empirical (surveys)	HR (recruitment, performance)
https://doi.org/10.32038/mbrq.2025.33.03	Conceptual review	Strategy (planning, scenarios)
https://doi.org/10.51505/IJEBMR.2024.8404	Bibliometrics	Meta -level
https://doi.org/10.48009/2_iis_2025_138	Cases	Finance and accounting

Taken together, the integrative review indicates that artificial intelligence is predominantly employed as a decision-support mechanism rather than as an autonomous decision-maker. Across the reviewed publications, AI applications are primarily focused on predictive analytics, classification, and knowledge management, particularly in domains such as human resources, finance, accounting, and strategic planning.

Study [2] highlights that legal and regulatory risks associated with AI use are not homogeneous but tend to emerge more strongly in sensitive or high-impact processes, especially under conditions of high automation, limited model transparency, or the presence of algorithmic biases.

Research reported in [3] proposes a categorization of decision types and associated risks, emphasizing that transparency and traceability enhance organizational acceptance of AI and are closely linked to accountability and compliance mechanisms by facilitating audits, justifications, and internal supervision.

Findings from [4] demonstrate that excessive delegation of decisions to opaque systems underscores the need to define explicit human responsibilities. Similarly, the analysis in [5] identifies risks related to discrimination and bias in human resource processes, while [6] shows that AI is increasingly reshaping strategic decision-making toward scenario-based planning practices.

The work described in [7] provides a bibliometric overview that contextualizes the field's evolution. In contrast, the study in [8] highlights tangible benefits of AI adoption in finance and accounting while also warning of privacy-related risks and potential biases.

Overall, the integrative review supports the interpretation of AI as a decision-support tool whose organizational value depends less on the degree of automation and more on its integration into governance structures, sustained human oversight, and control mechanisms. While AI offers substantial operational and analytical benefits, these advantages coexist with legal and ethical risks that require explicit regulatory and organizational responses.

DOI: https://doi.org/10.70715/jitcai.2026.v3.i3.065

These conclusions should be interpreted in light of certain limitations. The synthesis is narrative in nature due to the heterogeneity of methodological designs and evaluative metrics across the reviewed studies, and several publications do not explicitly address legal dimensions. In addition, potential availability biases related to language or access to sources may be present, and no meta-analysis was conducted.

In terms of reproducibility, the review process can be replicated using the documented DOIs and editorial metadata of the analyzed studies, provided that extraction dates and the defined inclusion and exclusion criteria are maintained. Collectively, the reviewed literature supports a set of compliance-oriented recommendations, including the promotion of AI transparency and explainability, the implementation of bias audits and algorithmic impact assessments, the definition of explicit human responsibilities through human-in-the-loop schemes, the strengthening of data governance, and the establishment of limits on AI autonomy in high-impact strategic decisions.

The integrative review thus provides the conceptual and regulatory foundation for the empirical component of the study. In particular, it highlights the limited empirical work examining how generative AI is operationalized in managerial compliance verification processes, which informed and motivated the design of the inductive case study of 11 multinational projects presented in the subsequent section.

3.2. Inductive study.

The empirical results are presented as a progressive evaluation of generative AI performance in managerial verification tasks. The study was structured into sequential stages to examine how contextual adjustments enhanced system reliability and alignment with regulatory and compliance requirements over time. Each stage was evaluated using an analytical rubric focused on information quality, argumentative coherence, and traceability.

3.2.1. Results in Phase 1.

In the initial phase, the AI system was applied using general inputs and cross-country regulatory data. The system exhibited significant inconsistencies, including the conflation of legislation from different jurisdictions and the generation of fabricated information to "fill gaps", which are clear indicators of hallucination arising from ambiguous or incomplete references.

Table 6: Phase 1 Results:

Country	Main findings	Quality	Argumentation	Traceability
Colombia	Mixture of local and international regulations; inconsistencies in tax references.	Low	Low	Low
Malaysia	Confusion surrounding tax regulations; AI-generated fabricated information to fill gaps.	Low	Low	Low
Costa Rica	Hallucinations in environmental regulations; incomplete references.	Low	Low	Low
United Arab Emirates	Mixture of commercial legislation; lack of precision in legal requirements.	Low	Low	Low
Belgium	Inconsistencies in European tax regulations; confusing references.	Low	Low	Low

All countries scored Low across the three evaluation criteria. These results indicate that the AI system lacked sufficient contextual grounding and required a more structured, project-specific approach to prevent regulatory confusion.

3.2.2. Results in Phase 2.

The second phase was introduced as a corrective response to the hallucinations and jurisdictional mixing observed in Phase 1. The analysis was reorganized at the project level, with separate sessions conducted for each project and the incorporation of refined contextual inputs based on contractual and regulatory information specific to each context.

Regulatory and legal inputs were reintroduced to reflect that different projects—even within the same country—may entail distinct contractual documentation and technical-legal requirements. These differences were particularly pronounced between private and public clients, as public projects are subject to stricter oversight by governmental authorities and regulatory bodies in their respective jurisdictions.

This adjustment enabled more precise alignment with applicable contractual and regulatory requirements. As a result, inconsistencies were substantially reduced, and the traceability of managerial decisions improved markedly.

Table 7: Phase 2 Results:

Country	Main findings	Quality	Argumentation	Traceability
Colombia	Greater precision in tax and contractual regulations; clear differences between public and private clients.	Good	Acceptable	Good
Malaysia	Improved consistency in tax requirements; reduction of inconsistencies.	Good	Good	Good
Costa Rica	Responses are more aligned with environmental regulations and strengthened traceability.	Good	Acceptable	Good
United Arab Emirates	Greater precision in trade regulations; better alignment with contextual constraints	Good	Good	Good
Belgium	Consistency in European tax regulations; significant reduction of errors.	Good	Good	Good

The system demonstrated a marked improvement in precision, contextual alignment, and traceability. Scores increased to Good or Acceptable across all evaluation criteria, with consistent improvements observed in tax, environmental, and trade-related regulatory assessments.

Performance improvements were assessed by comparing rubric scores and qualitative indicators, including reduced hallucinations, more accurate citation of applicable legislation, and improved alignment with contractual requirements. Contextual segmentation combined with iterative, project-specific inputs significantly enhanced system reliability and reduced informational incoherence.

3.2.3. Results in Phase 3.

In the third phase, corresponding to a real-world project case, the AI-generated report (Microsoft Copilot) fully aligned with the analysis and conclusions of the document prepared by the local legal advisor in the country where the project was implemented. An initial assessment was conducted without providing the AI system with the advisor's document, and no substantive inconsistencies or contradictions were identified when the outputs were compared with the expert analysis.

All evaluation criteria achieved Good ratings, indicating high levels of precision, argumentative coherence, and traceability. At this stage, the system demonstrated operational maturity, functioning as a reliable internal verification tool under continuous human supervision.

DOI: https://doi.org/10.70715/jitcai.2026.v3.i3.065

When the AI system was subsequently asked to analyze the legal advisor's document, its observations and recommendations remained consistent with the external legal assessment. In addition, the system identified formal (non-legal) inconsistencies and provided constructive feedback on the overall strategy. The acceptance and incorporation of these recommendations by the local legal advisor served as qualitative external validation of the Good ratings assigned in Phase 3.

Taken together, the progression across the three phases indicates that generative AI can evolve from producing inconsistent and hallucinatory outputs to supporting reliable managerial verification when guided by structured contextual inputs and human oversight. Improvements were assessed using comparative rubric scoring, qualitative analysis of coherence, and external qualitative validation from a professional legal advisor.

3.2.4. Authorship, governance, and limitations.

The final legal document remains the sole authorship and responsibility of the legal advisor. The AI system was used exclusively as an internal control mechanism and for secondary verification purposes. The use of AI did not directly modify the original document; instead, it generated recommendations that the legal advisor could accept or reject based on professional judgment, consistent with a human-in-the-loop approach.

3.2.5. Disclosure and Traceability.

At the conclusion of the exercise, disclosure of the use of artificial intelligence (Microsoft Copilot) in the official document submitted to authorities was not required, as the legal advisor remained the sole author and the document was formally signed by the company's legal representative. The AI system functioned exclusively as an internal verification tool for the contract manager, reinforcing compliance oversight without altering authorship or accountability.

The use of AI was recorded and documented in accordance with the company's management system's traceability and auditing practices, including the registration of prompts, document inputs, and generated recommendations. This approach ensures the availability of evidence for subsequent audits and aligns with established principles of transparency and explainability.

4. Discussion.

The findings support positioning AI as an assistive managerial tool rather than a substitute for professional judgment. Across the phased analysis, effectiveness was shown to depend on progressive enrichment of contextual inputs, explicit limits on automation, and sustained human-in-the-loop oversight, particularly when AI is embedded within office productivity platforms such as Microsoft 365.

The results further indicate that improvements in reliability, coherence, and traceability emerge when AI systems integrated into everyday managerial tools are deployed within structured governance and verification frameworks, rather than used as autonomous decision-making technologies. In compliance-sensitive contexts, excessive delegation to opaque systems increases risk exposure and reinforces the need for clearly defined human responsibility structures, supported by transparency, auditability, and external professional validation.

An additional implication concerns multilingual work environments typical of multinational contract management. The findings suggest that AI functionalities embedded in office productivity platforms, such as Microsoft 365, can help reduce language-related barriers in contractual and regulatory verification tasks. Prior contextualization through project-specific contractual, technical, and regulatory inputs supports more accurate interpretation and use of highly specialized legal and technical terminology across languages, limiting the risks associated with generic translation. In this way, AI-assisted verification can facilitate cross-jurisdictional work without language differences becoming a structural constraint, while preserving human oversight over legal meaning, accountability, and final judgment.

5. Conclusion and future work.

From a contract management perspective, these findings extend prior evidence by illustrating how AI-based verification tools can be operationalized as managerial support mechanisms within multinational compliance and risk-management processes, without displacing professional responsibility. For contract managers operating across multiple jurisdictions, AI functionalities embedded in office productivity platforms represent a cost-efficient means of strengthening internal controls, supporting contractual verification, and improving traceability when governed by transparent, human-in-the-loop frameworks.

Future research may extend this analysis across additional industries and contractual settings, incorporate quantitative performance indicators relevant to contract oversight, and examine the long-term governance implications of AI-assisted compliance verification within contract management functions.

6. References.

[1] M. Oppioli, MJ Sousa, M. Sousa, and E. de Nuccio, "Artificial intelligence in managerial decision-making: A structured literature review," *Management Decision,* 2023, doi: 10.1108/MD-08-2023-1331.

[2] R. Rydzewski, "Artificial intelligence and managerial decision-making: Adoption, opportunities, and accountability," *Managerial Economics,* vol. 26, no. 1, pp. 77–89, 2025, doi: 10.7494/manage. 2025.26.1.77.

[3] R. Boncella, "Ethical alignment of artificial intelligence in management decision making," *Information Systems Education Journal (IACIS IIS),* vol. 22, no. 4, 2024, doi: 10.48009/4_iis_2024_116.

[4] GHC Souza, CA Wanderley, and AB de Aguiar, "Transparency and acceptance of artificial intelligence in management decisions," *Journal of Management Control,* 2025, doi: 10.1007/s00187-025-00396-7.

[5] Y. Wen, J. Wang, and X. Chen, "Trust and decision weight of artificial intelligence in human resource management," *Frontiers in Psychology – Organizational Psychology,* 2025, doi: 10.3389/forgp. 2025.1419403.

[6] O. Rivero, "Artificial intelligence and strategic decision-making: A conceptual review," *European Journal of Studies in Management and Business,* Vol. 33, No. 3, 2025, doi: 10.32038/mbrq.2025.33.03.

[7] RC Jalagat Jr., "Artificial intelligence in management decision-making: A bibliometric analysis," *International Journal of Economics, Business and Management Research (IJEBMR)*, Vol. 8, no. 4, 2024, doi: 10.51505/IJEBMR.2024.8404.

[8] M. Numbi and G. Elongha, "Artificial intelligence for managerial optimization in finance and accounting," *Information Systems Education Journal (IACIS IIS)*, vol. 23, no. 2, 2025, doi: 10.48009/2_iis_2025_138.