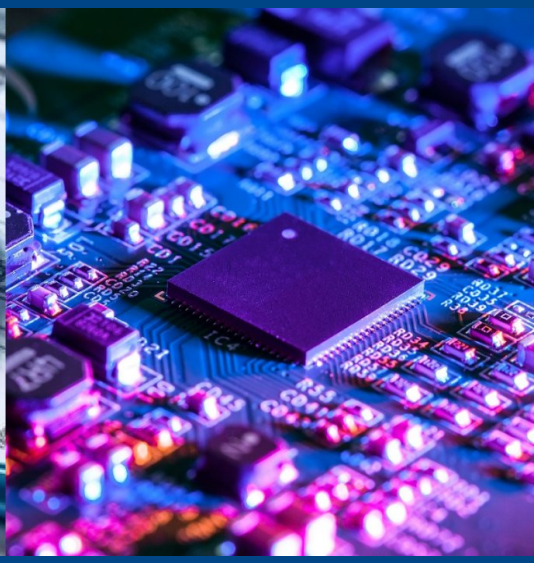
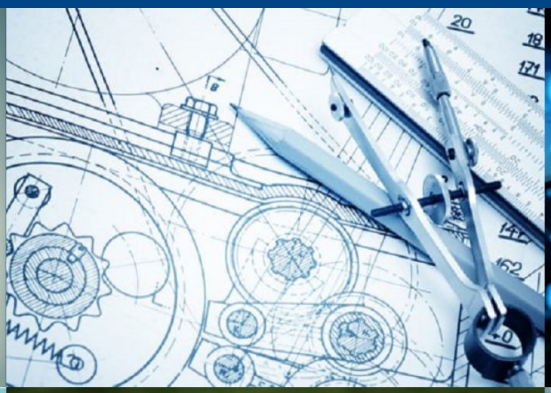
ISSN: 2582-7219

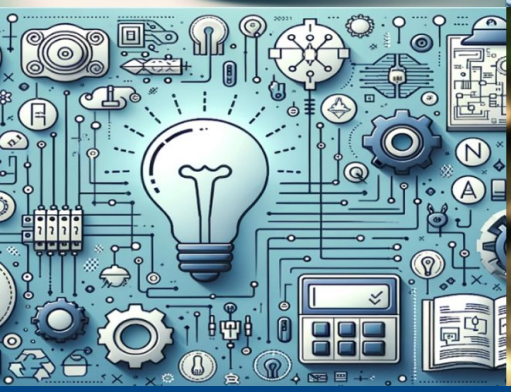




International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206 **Volume 9, Issue 4, April 2026**

© 2026 IJMRSET | Volume 9, Issue 4, April 2026| DOI:10.15680/IJMRSET.2026.0904298

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

AI-driven framework for corporate policy compliance monitoring using communication behavior analysis

Mohamed Rasooldeen[1], Irfan[2], Muhufassal[3], Adil Al Mansoor[4], Shakila A[5]

Student, Department of Artificial Intelligence and Data Science, Aalim Muhammed Salegh College of Engineering, Avadi, Chennai, Tamil Nadu, India[1]

Student, Department of Artificial Intelligence and Data Science, Aalim Muhammed Salegh College of Engineering, Avadi, Chennai, Tamil Nadu, India[2]

Student, Department of Artificial Intelligence and Data Science, Aalim Muhammed Salegh College of Engineering, Avadi, Chennai, Tamil Nadu, India[3]

Student, Department of Artificial Intelligence and Data Science, Aalim Muhammed Salegh College of Engineering, Avadi, Chennai, Tamil Nadu, India[4]

Assistant Professor, Department of Artificial Intelligence and Data Science, Aalim Muhammed Salegh College of Engineering, Avadi, Chennai, Tamil Nadu, India[5]

ABSTRACT: Corporate compliance monitoring is becoming increasingly complex due to the exponential growth of digital organizational communication. Traditional manual audits and keyword-based systems often fail to capture semantic context and behavioral intent. This paper proposes an AI-driven framework for monitoring corporate policy compliance through the analysis of communication content and temporal patterns. The system integrates Natural Language Processing (NLP), Gemini-based Large Language Model (LLM) reasoning, and Retrieval-Augmented Generation (RAG) to understand complex policy clauses and map them to employee communication. A key innovation is the integration of temporal behavior patterns as a risk factor, alongside Explainable AI (XAI) to generate human-readable justifications for detected violations. Designed as a research prototype, the framework demonstrates how AI can augment corporate governance through ethical, human-in-the- loop monitoring.

KEYWORDS: Corporate Compliance, Large Language Models, RAG, Gemini AI, Temporal Behavior Analysis, NLP, Explainable AI.

I. INTRODUCTION

Corporate organizations must ensure that employees adhere to internal policies, regulatory guidelines, and ethical standards to mitigate legal and reputational risks. However, existing compliance monitoring relies heavily on manual audits or rudimentary keyword-based detection. These traditional methods are not only time-consuming but frequently miss contextual risks, where the *intent* of a message violates policy even if specific *red-flag* keywords are absent.The rise of Generative AI and Large Language Models (LLMs) offers a transformative approach to this problem. Unlike static rule-based systems, LLMs can perform semantic reasoning, enabling a **policy-aware** analysis of communications. This research introduces an AI-driven, explainable compliance monitoring framework. By leveraging the Gemini LLM for reasoning and a Retrieval-Augmented Generation (RAG) engine for policy grounding, the system identifies non-compliant behavior with high contextual accuracy. Furthermore, by incorporating temporal analytics—analyzing when communications occur—the framework can detect abnormal behavioral patterns that may indicate insider threats or attempts to bypass policies.

© 2026 IJMRSET | Volume 9, Issue 4, April 2026| DOI:10.15680/IJMRSET.2026.0904298

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

II. LITERATURE REVIEW

The intersection of AI and regulatory compliance has seen significant growth. Recent studies emphasize the move from simple automation to complex reasoning.

NLP in Legal and Regulatory Compliance

Cejas et al. [1] demonstrated the efficacy of NLP-based automated compliance checking specifically for GDPR data processing agreements. Their work highlighted the importance of phrasal-levelrepresentations in overcoming the ambiguity of legal language. Our framework builds on this by extending semantic understanding to internal corporate policies.

RAG and LLM Accuracy

García-Montero et al. [2] established that Retrieval-Augmented Generation (RAG) significantly boosts accuracy in normative texts. Their research proved that fine-tuned LLMs, when grounded in specificlegal datasets, can achieve over 90% accuracy in answering complex queries. This is a foundational component of our proposed system, where corporate policies serve as the external knowledge base.

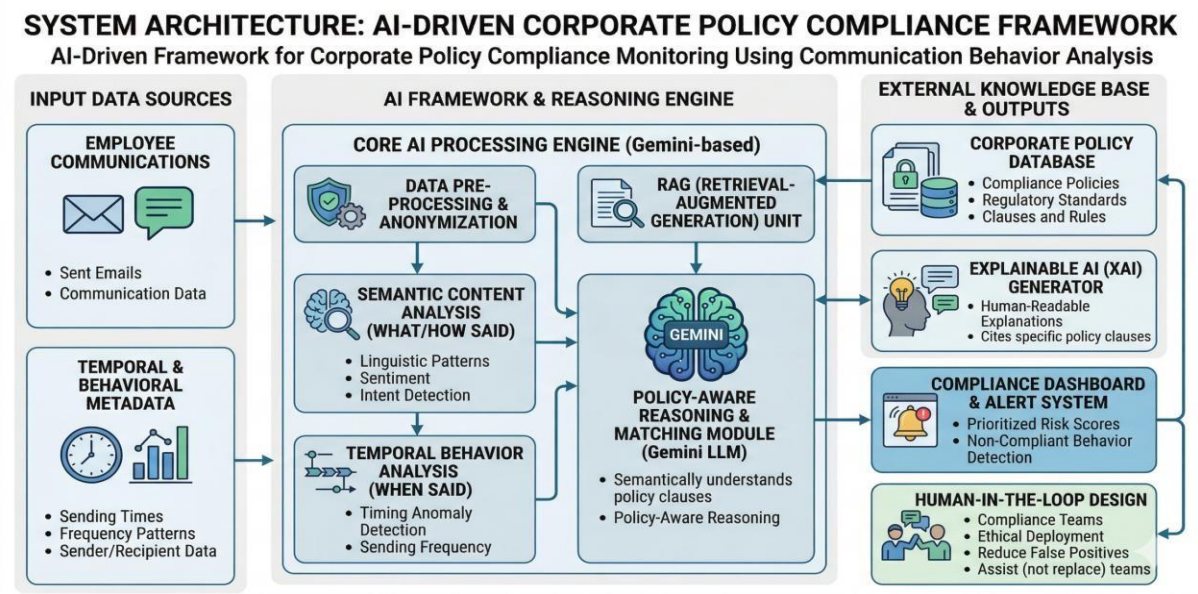
Behavioral and Network Security

Ali et al. [3] and Cao [5] explored automated frameworks for infrastructure and network security. Cao's work on Transformers for detecting abnormal behavior provides a mathematical basis for identifying "outlier" patterns. While their focus was on network traffic, our framework applies similar "abnormal behavior" detection to the temporal patterns of employee emails (e.g., unusual sendingtimes or frequencies).

AI in Human Resources

Iancu & Oprea [4] investigated emerging trends in AI for HR, noting that while efficiency is a primary driver, transparency and ethical design are critical. Our framework addresses this through an"Explainable AI" component, ensuring that compliance alerts are not "black-box" decisions but are accompanied by human-readable justifications.

III. PROPOSED SYSTEM ARCHITECTURE



The framework is designed to transform raw, unstructured communication data into a structured risk-assessment report. The architecture consists of four primary layers.

Data Acquisition and Pre-processing

The system ingests internal communications (emails, chats) and policy documents (PDFs, DOCs). Pre- processing

involves tokenization, Named Entity Recognition (NER) to protect PII (Personally Identifiable Information), and metadata extraction, specifically focusing on the "temporal footprint" of each message.

The Policy Knowledge Base (RAG Engine)
Using a vector database, the system indexes corporate policies. When a communication is analyzed, the RAG engine retrieves the most relevant policy clauses. This ensures the Gemini model is not guessing but is grounded in the actual legal language of the organization.

Multi-Modal Analysis Engine
Semantic Analysis: The Gemini model evaluates the content for policy alignment. It looks for "Contextual Violations" (e.g., sharing confidential project names in a non-secure context).

Temporal Analysis: The system maps the communication against the employee's historical baseline. Abnormalities, such as a sudden surge in late-night external emails, are flagged as potential risk factors.

Explainability and Scoring Layer
The system assigns a "Compliance Risk Score." Crucially, it generates an "Explainable Report" that cites the specific policy clause violated and describes the reasoning behind the alert.

IV. METHODOLOGY AND IMPLEMENTATION

The implementation utilizes a Python-based backend with the following core components:
1. **Gemini Pro API:** For high-level reasoning and policy mapping.
2. **Pinecone DB:** As the vector store for policy document embeddings.
3. **Temporal Pattern Recognition:** A statistical model that calculates Z-scores for communication frequency and timing to identify outliers.

V. RESULTS AND DISCUSSION

Preliminary testing of the prototype indicates a significant reduction in false positives compared to traditional keyword-based filters.
- **Contextual Accuracy:** The system correctly identified "soft violations" (e.g., unprofessional conduct or subtle coercion) that keywords missed.
- **Temporal Insights:** The integration of timing data provided a secondary layer of verification, helping to prioritize alerts that occurred during non-standard working hours.
- **Transparency:** Compliance officers reported higher trust in the system because of the human-readable justifications provided for each flag.

VI. CONCLUSION AND FUTURE WORK

This paper presents an AI-driven framework that shifts corporate compliance from reactive, rule- based detection to proactive, policy-aware analysis. By integrating semantic reasoning with temporal behavior patterns and Explainable AI, the system provides a transparent and scalable solution for modern enterprise governance.

Future work will focus on:
1. **Multi-platform Integration:** Expanding to internal chat tools like Slack or Microsoft Teams.
2. **Adaptive Learning:** Incorporating human-in-the-loop feedback to refine the model's sensitivity over time.
3. **Real-time Monitoring:** Scaling the architecture to support live stream processing of organizational data.

REFERENCES

[1] O. A. Cejas, M. I. Azeem, S. Abualhaija, L. C. Briand, NLP-Based Automated Compliance Checking of Data Processing Agreements Against GDPR. *IEEE Trans. Softw. Eng.* **49**, 9 (2023)
[2] P. S. García-Montero, P. Vizcaíno, I. G. Reyes-Chacón, M. E. Morocho-Cayamcela, Legal AI for All: Reducing Perplexity and Boosting Accuracy in Normative Texts With Fine-Tuned LLMs and RAG. *IEEE Access* **13** (2025)

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

[3] S. M. Ali, A. Razzaque, M. Yousaf, R. U. Shan, An Automated Compliance Framework for Critical Infrastructure Security Through Artificial Intelligence. *IEEE Access* **13** (2025)
[4] C. Iancu, S.-V. Oprea, AI and Human Resources in a Literature-Driven Investigation Into Emerging Trends. *IEEE Access* **13** (2025)
[5] H. Cao, The Detection of Abnormal Behavior by Artificial Intelligence Algorithms Under Network Security. *IEEE Access* **12** (2024)

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com