The Insider Risk of Artificial Intelligence in Financial Systems through the Lens of Large Language Models

Ishrak Alim
M.S. in Accounting Analytics
University of New Haven, Connecticut, USA
Email: alimishrak@gmail.com

Tasnia Farzana Matin
M.S. in Digital Marketing Analytics
Montclair State University, New Jersey, USA
Email: matinfarzana77@gmail.com

Takib Md Masudul Hasan Prodhan
Account Executive, T&S Buttons Lanka Ltd
Dhaka, Bangladesh
Email: bd.takib1@gmail.com

Md Lahaduzzaman Lahad
Digital Marketing Executive, Cloud Bridge Consultancy
Dhaka, Bangladesh
Email: lahaduzzaman.lahad@yahoo.com

Table of Contents

Abstract

The integration of Large Language Models (LLMs) into financial systems introduces transformative opportunities alongside unprecedented risks particularly those related to insider threats. This paper examines the multifaceted security, privacy, and governance challenges of deploying LLMs in sensitive financial workflows, including accounts payable, forecasting, and client advisory. We present a taxonomy of emerging vulnerabilities such as prompt injection, data leakage, and cross-user exposure and analyze how these risks align with traditional and synthetic insider threat patterns. Drawing on real-world financial use cases, we introduce a structured risk matrix and propose mitigation strategies spanning technical controls (e.g., prompt isolation, logging) to high-level governance frameworks. The study concludes with practical recommendations for enterprise deployment and identifies key areas for further research, including model interpretability, ethical AI adoption, and proactive risk modeling.

Keywords: Large Language Models (LLMs); Insider Threats; Financial Systems Security; AI Governance; Prompt Injection; Data Leakage; Role-Based Access Control; Model Auditing; Explainable AI (XAI); Risk Assessment Matrix

1. Introduction

1.1 Context of AI in Financial Systems

The financial industry has long been at the forefront of adopting cutting-edge technologies to enhance efficiency, decision-making, and risk management. Artificial Intelligence (AI), especially in its advanced forms like machine learning and natural language processing, is now deeply embedded in critical financial functions such as fraud detection, credit scoring, compliance monitoring, and portfolio management. Institutions leverage AI to sift through vast volumes of structured and unstructured data to uncover patterns that humans might miss (Li, 2025). This adoption, however, brings a new wave of risks among them, the increasing relevance of insider threats, especially those amplified through the capabilities of AI systems (Luca, 2024).

1.2 Emergence of Large Language Models in Enterprise Workflows

Large Language Models (LLMs) like GPT-4 and Claude have redefined how enterprises engage with information, automate tasks, and interact with stakeholders. In finance, LLMs are used for tasks ranging from customer support automation to complex financial data extraction and risk forecasting (Jami Venkata Suman, 2022). These models offer unmatched capability in handling natural language tasks but also present unique privacy, security, and governance concerns, especially due to their ability to memorize sensitive inputs and generate outputs based on that training data (Luca, 2024).

1.3 Defining the Insider Threat in the Age of AI

Insider threats traditionally involve current or former employees misusing their authorized access to harm an organization’s operations, data, or reputation. However, the definition is evolving. With AI systems like LLMs acting as intelligent intermediaries, the boundary between human and system-initiated insider actions is blurring. These systems can be manipulated via carefully crafted prompts or indirectly coerced into leaking information, simulating insider-like behavior (Antonino Ferraro, 2025). Moreover, AI-enhanced insider threats can operate covertly, scale rapidly, and adapt in real-time traits that pose formidable challenges to traditional detection and mitigation frameworks (A. Adusumilli, 2024), (Ashrafi, 2024).

2. Security and Privacy Risks of Large Language Models

Large Language Models (LLMs) present a transformative shift in enterprise AI, particularly within data-sensitive sectors like finance. However, their generative nature and training methods introduce a spectrum of novel vulnerabilities. This section explores three core risks relevant to insider threat modeling: memorization and data leakage, prompt injection and shadow access, and cross-user data exposure.

2.1 Memorization and Data Leakage

A foundational security concern in LLMs is their tendency to memorize training data, including potentially sensitive information. Unlike classical algorithms, LLMs operate with token-level recall, meaning they can inadvertently regenerate fragments of proprietary datasets if prompted under the right conditions. This behavior, while often unintentional, exposes organizations to reputational and regulatory risks, especially under frameworks like GDPR and PCI-DSS.

Das et al. (2025) classify this phenomenon as “non-intentional leakage,” where LLMs reproduce training content due to overfitting on high-frequency sequences (Li, et al., 2024). Notably, this type of leakage can be triggered using prompt engineering techniques, which exploit the probabilistic decoding structure of the models. Similarly, Wang et al. (2024) note that deduplication and selective data curation prior to training are necessary yet insufficient without enforcing post-training data sanitization layers (Chen, et al., 2024).

This risk is amplified in financial institutions where fine-tuned models might inadvertently retain customer account numbers, investment details, or internal benchmarks any of which could be retrieved through model inversion attacks. Mitigation requires combining privacy-preserving training techniques with differential auditing to detect unintentional memorization before deployment.

2.2 Prompt Injection and Shadow Data Access

Prompt injection attacks exploit the contextual openness of LLMs by embedding malicious commands within user inputs. Unlike conventional exploits that target code or API vulnerabilities, these attacks manipulate model outputs by altering the interpretation space of the prompt itself. Such injections can subtly coerce the model into revealing stored system instructions or misbehaving according to adversarial logic.

Derner et al. (2024) document a series of prompt-based exploits capable of overriding safety filters and extracting confidential instruction tokens within chatbot environments (Qian, et al., 2024). These manipulations often work by exploiting poorly scoped system prompts or residual context in multi-turn conversations. Similarly, Abdali et al. (2024) propose that prompt leakage a variant of injection can grant attackers indirect access to latent memory states or earlier session content (Zhang, et al., 2024).

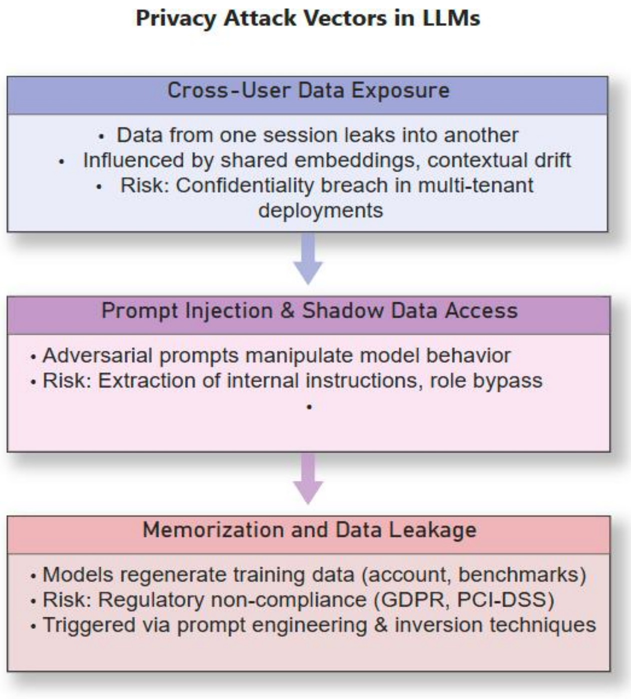
In financial systems, where LLMs are used to automate risk analysis or handle real-time client communication, the implications are severe. An injected prompt could simulate confidential access (e.g., internal analyst reports or market positions), misleading users and breaching compliance. Therefore, securing system-level prompts and implementing prompt boundary enforcement becomes indispensable to any enterprise-grade deployment.

2.3 Cross-User Data Exposure

Multi-tenant deployments of LLMs introduce the possibility of data leakage across user sessions either via shared memory vectors or contextual drift. Even when explicit memory is disabled, embeddings from prior queries can influence subsequent outputs, particularly in environments with fast model cycling and high concurrency.

Liu and Hu (2024) detail this vector as a derivative of context leakage, wherein model state persistence allows semantic features from one user's interaction to bleed into another's session (C, 2024). This is especially problematic in financial use cases where institutional clients expect strict segregation of their data. Chen et al. (2025) further confirm that session boundary failure is one of the most under-addressed risks in LLM APIs and recommend active memory flushing between sessions (Chen, et al., 2024).

Cross-user exposure isn't always visible to the naked eye it often manifests as a gradual model response drift, where terminology or data patterns reappear across unrelated conversations. Addressing this challenge requires architectural separation at both inference and data routing levels, complemented by rigorous user-session containerization.



Privacy Attack Vectors in LLMs

Cross-User Data Exposure

- Data from one session leaks into another
- Influenced by shared embeddings, contextual drift
- Risk: Confidentiality breach in multi-tenant deployments

Prompt Injection & Shadow Data Access

- Adversarial prompts manipulate model behavior
- Risk: Extraction of internal instructions, role bypass
-

Memorization and Data Leakage

- Models regenerate training data (account, benchmarks)
- Risk: Regulatory non-compliance (GDPR, PCI-DSS)
- Triggered via prompt engineering & inversion techniques

Figure 1: Privacy Attack Vectors in Large Language Models

This figure outlines key privacy risks in LLMs starting from cross-user data leaks, progressing through prompt injection, and culminating in memorization of sensitive training data.

Risk Type	***Mechanism***	***Severity***	***Detection Difficulty***	***Mitigation Strategy***
Prompt Memorization	*Retention of sensitive data from training corpus*	*High*	*High*	*Data deduplication, differential privacy*
Prompt Injection	*Embedding adversarial instructions in input prompts*	*High*	*Medium*	*Prompt validation, system prompt isolation*
Inference Leakage	*Model outputs reveal internal tokens or context bleed*	*Medium*	*High*	*Output filtering, logging-based anomaly detection*
Cross-User Data Exposure	*Data from one session appears in another*	*High*	*High*	*Strict session isolation, memory reset*

Table 1: Comparative Analysis of LLM Privacy Risks

3. Use Case Scenarios in Financial Workflows

As Large Language Models (LLMs) become integralto enterprise AI ecosystems, the financial services sector is emerging as a primary field of implementation. These models are redefining critical workflows such as invoice automation, financial forecasting, and digital advisory services. This section explores three high-impact applications that simultaneously illustrate LLM utility and the associated insider risks.

3.1 Accounts Payable and Invoice Processing

Accounts payable (AP) workflows, traditionally burdened by manual data entry and error-prone verification, have significantly benefited from the integration of LLMs. By ingesting structured and unstructured invoice formats, models like GPT-4 can extract payment terms, supplier details, and due dates with high accuracy, thereby reducing reconciliationcycles and fraud risks.

Raza et al. (2025) highlight that LLM-powered back-office systems now enable semantic understanding of invoices and purchase orders, achieving up to 92% accuracy in cross-validation tasks across ERP datasets (Raza, Jahangir, Riaz, Saeed, & Sattar, 2025). Similarly, Schnepf et al. (2024) describe how financial enterprises are embedding LLMs into their enterprise resource planning (ERP) systems to automate invoice matching, ledger entry, and payment authorizations (Schnepf, Engin, & Scheuermann, 2024).

While these systems deliver speed and efficiency, they also pose data residency and access risks. Unauthorized prompt access or residual memory leaks in shared LLM environments may expose sensitive vendorand payment information. Hence, role-based access to LLM APIs and prompt boundary enforcement are essential for operational security.

3.2 AI-Augmented Forecasting or Budgeting

Forecasting and budgeting core financial planning tasks have evolved beyond spreadsheet modeling. LLMs offer a nuanced layer of predictive reasoning by combining temporal trends, contextual narratives, and real-time data feeds. Their integration in finance allows dynamic scenario modeling and forward-looking budget alignment, especially under volatile market conditions.

Nie et al. (2024) underscore the ability of LLMs to synthesize macroeconomic indicators with historical financials, resulting in forecasting systems that adaptively learn from both structured data and unstructured disclosures (Nie, et al., 2024). This capability is particularly relevant for multinational firms dealing with currency fluctuations, interest rate exposure, and variable cost structures.

According to Li and Vasarhelyi (2024), advanced LLM pipelines can generate baseline budgets, simulate variance analyses, and even draft narrative justifications for budget reports, thereby automating the initial layers of financial planning (Li & Vasarhelyi, 2024). However, as these models operate on sensitive operationaldata, embedding AI-specific audit trails and interpretability layers becomes critical for compliance and risk oversight.

3.3 Chatbot-Driven Financial Advisory

One of the most transformative applications of LLMs in finance lies in customer-facing financial advisory chatbots. These systems interface with clients to provide portfolio overviews, transaction histories, investment recommendations, and even tax guidance, all via natural language interfaces.

He and Jin (2025) demonstrate how financial institutions leverage LLMs to enhance customer communication, drive down operational costs, and increase engagement through personalized advisory platforms (He & Jin, 2025). Additionally, Wardani and Navarro (2024) report that SaaS financial providers are integrating LLMs into customer billing and payment systems, creating end-to-end advisory and support cycles (Wardani & Navarro, 2024).

However, the trust model in chatbot-driven advisory must be handled with care. A prompt injection vulnerability could coerce the model into delivering flawed investment strategies or revealing confidential client histories. Regulatory implications, particularly under MiFID II and SEC compliance frameworks, necessitate robust monitoring, conversationlogging, and explainable AI (XAI) layers.

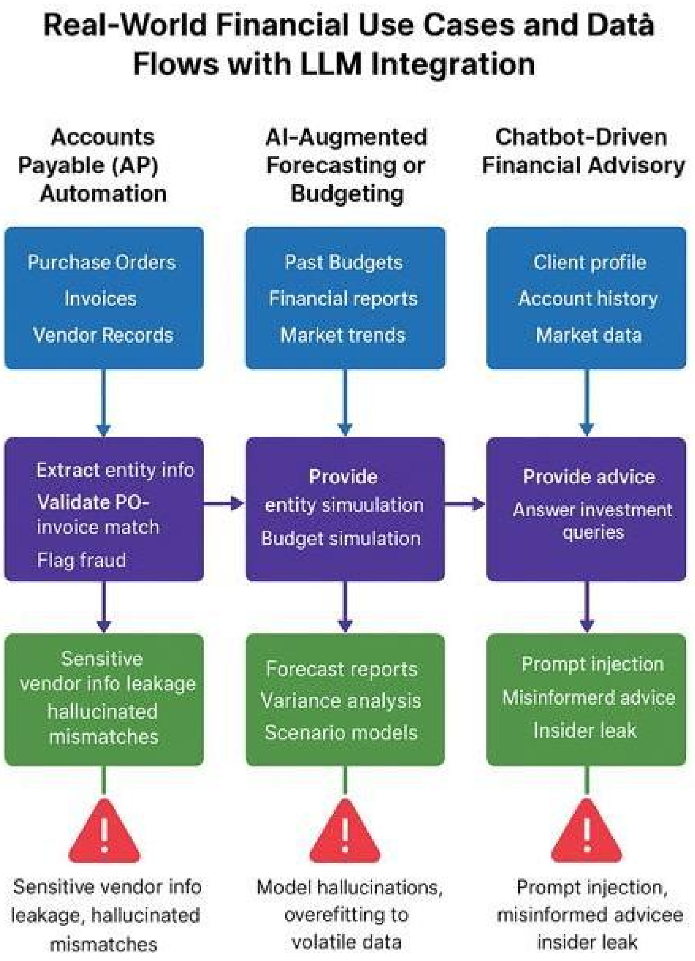

Real-World Financial Use Cases and Data Flows with LLM Integration
Accounts Payable (AP) Automation
AI-Augmented Forecasting or Budgeting
Chatbot-Driven Financial Advisory
Purchase Orders
Invoices
Vendor Records
Past Budgets
Financial reports
Market trends
Client profile
Account history
Market data
Extract entity info
Validate PO-invoice match
Flag fraud
Provide
entity simuulation
Budget simulation
Provide advice
Answer investment queries
Sensitive vendor info leakage hallucinated mismatches
Forecast reports
Variance analysis
Scenario models
Prompt injection
Misinformerd advice
Insider leak
Sensitive vendor info leakage, hallucinated mismatches
Model hallucinations, overefitting to volatile data
Prompt injection, misinformed advicee insider leak

Figure 2: Large Language Models (LLMs) integrate into financial workflows by processing key data in accounts payable, forecasting, and advisory systems. This diagram illustrates input sources, AI tasks, and resulting outputs, alongside critical risks unique to each workflow.

4. Threat Modeling and Risk Analysis

The intersection of Large Language Models (LLMs) and financial infrastructure introduces a new dimensionto insiderthreat modeling. Unlike traditionalinsiderthreats which are bound by human behavior and access controls LLM-enabled systems create opportunities for synthetic, automated, and often undetectable exploitation. This section outlines emerging insider threat archetypes introduced by LLMs and presents a risk matrix to guide enterprise-level assessments.

4.1 Insider Risk Patterns Mirrored by AI Tools

Insider threats are no longer confined to rogue employees or negligent users. LLMs now possess capabilities to replicate, amplify, or mask insider behaviors either through prompt manipulation, memorization of confidential information, or contextual misuse. These models, embedded in sensitive workflows, can inadvertently act as conduits for unauthorized access or unintentional data exposure.

Ferraro et al. (2025) propose that generative LLMs can function as synthetic agents whose behaviors mimic those of human insiders, especially when aligned with adversarial prompting or contextual ambiguity (Ferraro, Orlando, & Russo, 2025). Similarly, Yilmaz and Can (2024) describe a spectrum of threat profiles where user behavior analytics (UBA) applied to LLM interactions reveals latent risks such as coercion, replay attacks, and role bypassing (Yilmaz & Can, 2024).

One critical pattern involves prompt injection where external users subtly manipulate model logic to extract confidential data or override access controls. Another is role assumption, where LLMs respond as though they possess elevated privileges, especially if prompt history is not tightly scoped. Moreover, Adusumilli et al. (2024) document how insider trading scenarios can be simulated and masked using generative outputs, complicating forensic traceability and regulator audits (A. Adusumilli, 2024).

These scenarios highlight a foundational issue: AI systems trained on human data can reproduce human vulnerabilities. Without proper threat modeling, they may introduce entirely new insider classes automated, scaleable, and highly adaptive.

4.2 Risk Assessment Matrix

To operationalize these risks, financial organizations must adopt a structured risk matrix. Below is a synthesized risk model capturing the most common insider-like attack patterns made possible through LLM deployment in finance:

Scenario	***Threat Source***	***Impact***	***Likelihood***	***Detection Difficulty***	***Recommended Control***
Prompt-based Data Extraction	*External attacker*	*Confidential info leak*	*Medium*	*High*	*Prompt isolation, query sanitization*
Role Simulation via Prompt Injection	*Internal user*	*Unauthorized access*	*High*	*High*	*Access control scope, input validation*

Shadow AI Agent Misuse (covert LLM use)	*Insider employee*	*Compliance breach*	*Low*	*Medium*	*LLM API monitoring, endpoint restrictions*
Memorization-triggered Info Recall	*Model (synthetic insider)*	*Data privacy violation*	*Medium*	*High*	*Differential privacy, dataset deduplication*
Forecast Tampering in Budgeting Models	*Analyst (insider)*	*Financial misreporting*	*Medium*	*Medium*	*Output auditing, variance analysis benchmarks*

Ali and Ghanem (2025) emphasize that future cybersecurity frameworks must integrate AI-specific risk taxonomies that treat LLMs not just as tools but as potential actors in threat chains (Ali & Ghanem, 2025).

Building defenses thus requires cross-disciplinary insight blending behavioral detection, access governance, AI transparency, and adversarial testing. Without such alignment, financial systems risk facing “shadow insiders” agents that don’t exist as people but perform with precision indistinguishable from malicious insiders.

5. Mitigation and Governance Strategies

The accelerated integration of large language models (LLMs) in financial infrastructures has created not only productivity gains but also multidimensional risks. Addressing insider threats, data leakage, and adversarial misuse requires a multilayered mitigation approach that includes technical isolation, traceable accountability, and governance rooted in ethical frameworks. This section outlines three core strategies prompt isolation, audit trail design, and governance standards to guide secure LLM deployment in financial contexts.

5.1 Role-Based Prompt Isolation

Role-based prompt isolation refers to the architectural principle of segmenting LLM access and memory scope based on user roles, privilege levels, and task-specific context. Unlike traditional access control systems, which restrict UI or file-layer access, prompt isolation ensures that a model's interpretive scope is limited by the security domain of the prompting user. This limits cross-contextual data leakage and reduces the risk of indirect privilege escalation.

Tavasoli et al. (2025) propose isolation protocols and rollback strategies in their framework for secure LLM integration, demonstrating how isolating prompt chains can prevent role bleed and synthetic privilege elevationin financial applications (Tavasoli, Sharbaf, & Madani, 2025). Complementarily, Saha et al. (2025) recommend layered LLM sandboxing, enabling systems to safely interpret inputs while preventing prompt history misuse or unauthorized model carryover in session-based environments (Saha, Rani, & Shukla, 2025).

In financial systems, where employees operate across dynamic access tiers (e.g., trading desks, compliance teams, and IT support), such role-prompt isolation becomes foundational to mitigating both intentional and accidental misuse.

5.2 Logging and AI Audit Trails

Robust logging mechanisms are essential for operational transparency and forensic accountability. When implemented correctly, AI audit trails enable financial institutions to trace prompt origins, monitor model responses, and flag anomalous interactions. Logging must be layered: covering pre-prompt conditions, model selection parameters, real-time outputs, and post-response user actions.

Mӧkander et al. (2024) argue that auditing LLMs requires a three-layered approach, consisting of procedural logging, technical intervention, and organizational oversight to systematically evaluate model behavior across decision-critical environments (Mӧkander, Schuett, & Floridi, 2024). Xu (2025) further supports the need for context-sensitive audit frames in LLM-augmented advisory services, noting that most enterprise AI risks arise not from raw output errors, but from unlogged escalations of prompts into sensitive domains (Xu, 2024).

In regulated financial domains, this auditability is not just a best practice it's legally mandated under GDPR, MiFID II, and internal control standards. LLM-specific audit logs should be encrypted, immutable, and capable of supporting model explainability reviews.

5.3 Governance Frameworks and Ethical Considerations

Governance in LLM deployment spans technical protocols and ethical imperatives. Financial institutions must ensure their AI systems align with fairness, transparency, accountability, and explainability (the "FTAE" pillars), particularly when interacting with sensitive client data or making automated financial suggestions.

Jain (2024) outlines a policy architecture that combines technical guardrails with regulatory norms to prevent algorithmic bias and ensure equitable treatment across financial cohorts (Jain, 2024). Similarly, Bansal (2024) suggests embedding ethical evaluations into the model development lifecycle itself, arguing that delayed governance retrofits are ineffective once LLMs have been deeply integrated into institutional systems (Sandfreni & Bansal, 2024).

To institutionalize AI governance, organizations are increasingly adopting frameworks such as NIST's AI Risk Management Framework and the OECD AI Principles. These provide a blueprint for organizationalalignment between data science teams, risk officers, compliance leads, and policy stakeholders.

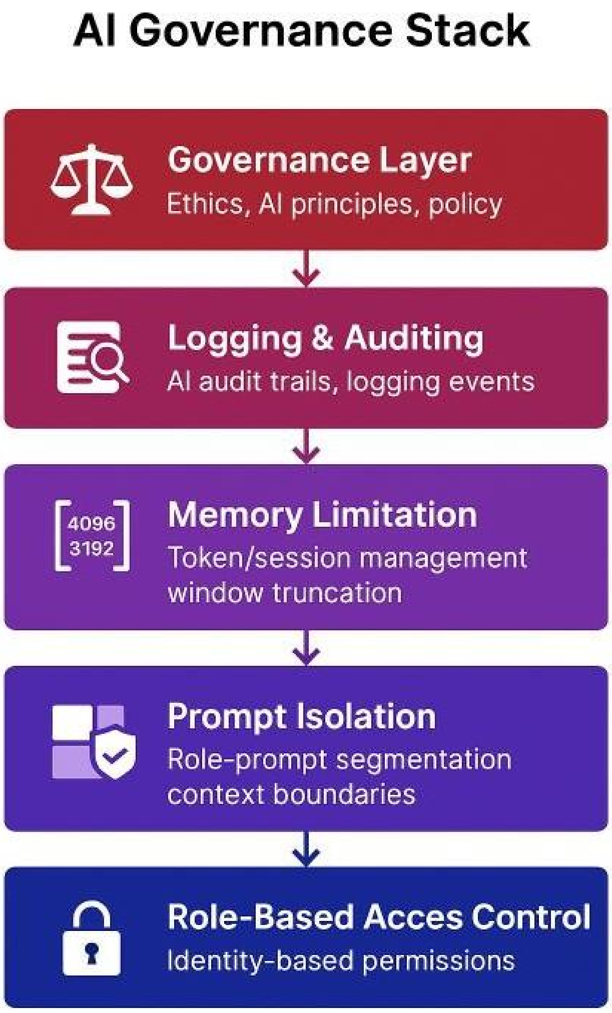


Figure 3: AI Security Governance Stack for Financial LLM Systems

This layered architecture illustrates the core components of secure and ethical LLM deployment in financial settings. From identity-based access controls to governance aligned with AI principles, each layer builds upon the next to ensure operational integrity, transparency, and regulatory compliance.

6. Conclusion and Future Work

6.1 Key Takeaways

Large Language Models (LLMs) are reshaping financial systems by streamlining workflows, enabling adaptive decision-making, and providing scalable intelligence. However, as this paper outlines, these benefits are tightly coupled with emerging insider threats, data governance dilemmas, and systemic compliance risks. Notably, the risk landscape expands from traditional operational failures to more nuanced concerns such as prompt manipulation, memorization of sensitive data, and opaque role-boundary breaches. According to Cao et al. (2024), LLMs represent "transformative but brittle" infrastructure when deployed without security-first frameworks (Cao, et al., 2024).

6.2 Practical Implications for Financial Enterprises

From automated advisory bots to forecasting engines, the deployment of LLMs in finance requires continuous alignment between technology, compliance, and human oversight. Nie et al. (2024) argue that the most immediate challenge for financial institutions lies in "calibrating LLMs to domain-specific logic while safeguarding data perimeter integrity" (Nie, et al., A Survey of Large Language Models for Financial Applications: Progress, Prospects and Challenges, 2024).

For practical implementation, role-based prompt isolation, robust session logging, and risk-aware model tuning are no longer optional they are structural requirements. Moreover, consistent auditing trails must be embedded from day one, with monitoring tools capable of red-teaming model behaviors, particularly in functions such as insider trading detection and invoice fraud. Elgendy et al. (2025) emphasize that LLM adoption in financial ecosystems should be paired with governance dashboards and compliance automation as standard operating procedures (Elgendy, et al., 2025).

6.3 Areas for Further Research

While progress has been made in modeling insider threats using generative agents and NLP-based anomaly detection (Ferraro et al., 2025) (Antonino Ferraro, 2025), several areas remain underexplored. Future research should target:

Context-Aware Risk Scoring: Developing LLM-specific insider threat taxonomies that adjust for organizational context and intent.

Prompt Transparency: Creating explainable prompt execution layers that allow audit teams to trace how inputs evolve through model layers.

Self-Limiting AI Models: Prototyping LLMs with dynamic access throttling and memory-aware behaviors tuned for regulatory contexts.

Additionally, more empiricalstudies are needed to measure the actual failure modes of LLMs across diverse financial institutions, rather than relying solely on lab-based simulation datasets. This will be essential for standardizing benchmarks and driving international alignment on AI governance.

References

A. Adusumilli, S. R. (2024). Leveraging Large Language Models and Deep Learning for Detecting Illegal Insider Trading. *2024 IEEE International Conference on Big Data (BigData)* (pp. 4809-4818). Washington, DC, USA: IEEE.

Ali, A., & Ghanem, M. C. (2025). Beyond Detection: Large Language Models and Next-Generation Cybersecurity. *SHIFRA*, 81-97. Retrieved from https://peninsula-press.ae/Journals/index.php/SHIFRA/article/view/183

Antonino Ferraro, G. M. (2025). Generative Agent-Based Modeling with Large Language Models for insider threat detection. *Engineering Applications of Artificial Intelligence*.

Ashrafi, S. M. (2024, May 07). *An Empirical Study on the Use of Generative AI Tools by College Students.* Retrieved from arXiv: https://arxiv.org/pdf/2505.03796

C, L. F. (2024). *Exploring Vulnerabilities and Protections in Large Language Models: A Survey.* arXiv. Retrieved from https://arxiv.org/abs/2406.00240

Cao, X., Li, S., Katsikis, V., Khan, A., He, H., Liu, Z., . . . Peng, C. (2024). Empowering Financial Futures: Large Language Models in the Modern Financial Landscape. *AIRO Journal*.

Chen, J., Li, Z., Xu, C., Li, Y., Zhang, H., Wang, Y., & Huang, M. (2024). *Privacy Risks in Large Language Models: A Survey.* Arxiv. Retrieved from https://arxiv.org/abs/2406.07973

Chen, K., Zhou, X., Lin, Y., Feng, S., Shen, L., & Wu, P. (2024). *Exploring Vulnerabilities and Protections in Large Language Models: A Survey.* ArXiv. Retrieved from https://arxiv.org/abs/2505.01976

Elgendy, I. A., Helal, M. Y., Al-Sharafi, M. A., Albashrawi, M. A., Al-Ahmadi, M. S., Jeon, I., & Dwivedi, Y. K. (2025). Agentic Systems as Catalysts for Innovation in FinTech: Exploring Opportunities, Challenges and a Research Agenda. *Information Discovery and Delivery*. Retrieved from https://www.emerald.com/insight/content/doi/10.1108/idd-03-2025-0068/full/html

Ferraro, A., Orlando, G. M., & Russo, D. (2025). Generative Agent-Based Modeling with Large Language Models for Insider Threat Detection. *Engineering Applications of Artificial Intelligence*, 111343. Retrieved from https://doi.org/10.1016/j.engappai.2025.111343

He, Z., & Jin, X. (2025). Integrate AI-based chatbots into accounting services: Enhance customer communication and financial management support. *Journal of Computational Methods in Sciences and Engineering*. Retrieved from https://journals.sagepub.com/doi/abs/10.1177/14727978251338982

Jain, A. (2024). *Ethical AI: A Policy Framework to Regulate Bias in Large Language Models.* Austin, TX: University of Texas at Austin.

Jami Venkata Suman, S. R. (2022). Automated Machine Learning Framework Using Large Language Models for Financial Security in Cloud Observability. *International Journal of Research and Analytical Review (IJRAR)*.

Li, H. G. (2025). Extracting Financial Data from Unstructured Sources: Leveraging Large Language Models. *Journal of Information Systems*, 135-156.

Li, H., & Vasarhelyi, M. A. (2024). *Applying Large Language Models in Accounting: A Comparative Analysis of Different Methodologies and Off-the-Shelf Examples.* SSRN.

Li, X., Wang, Y., Lin, H., Zhou, Y., Liu, Y., & Liang, X. (2024). Backdoor Risks in LLMs: An Empirical Study. *Proceedings of the 2024 ACM Asia Conference on Computer and Communications Security (AsiaCCS '24)* (pp. 1-39). Singapore: Association for Computing Machinery.

Luca, C. (2024, December). *OPTIMIZING LARGE LANGUAGE MODELS FOR FINANCIAL RISK ASSESSMENT IN CREDIT UNIONS*. Retrieved from Researchgate: https://www.researchgate.net/publication/388060268_OPTIMIZING_LARGE_LANGUAGE_MODELS_FOR_FINANCIAL_RISK_ASSESSMENT_IN_CREDIT_UNIONS#fullTextFileContent

Mökander, J., Schuett, J., & Floridi, L. (2024). Auditing large language models: A three-layered approach. *AI and Ethics*, 1085–1115. Retrieved from https://link.springer.com/article/10.1007/s43681-023-00289-2

Nie, Y., Kong, Y., Dong, X., Mulvey, J. M., Poor, H. V., Wen, Q., & Zohren, S. (2024). *A Survey of Large Language Models for Financial Applications: Progress, Prospects and Challenges.* arXiv.

Nie, Y., Kong, Y., Dong, X., Mulvey, J. M., Poor, H. V., Wen, Q., & Zohren, S. (2024). *A Survey of Large Language Models for Financial Applications: Progress, Prospects and Challenges.* ArXiv. Retrieved from https://arxiv.org/abs/2406.11903

Qian, Y., Liu, J., Gao, J., Zhou, Y., Zhang, Y., Wu, P., . . . Liu, S. (2024). Comprehensive Privacy Threats to LLMs: A Survey. *Proceedings of the 2024 IEEE 10th International Conference on Big Data Security on Cloud (BigDataSecurity)*. Retrieved from https://ieeexplore.ieee.org/abstract/document/10648691

Raza, M., Jahangir, Z., Riaz, M. B., Saeed, M. J., & Sattar, M. A. (2025). Industrial applications of large language models. *Scientific Reports*, 13755. Retrieved from https://www.nature.com/articles/s41598-025-13755-6

Saha, B., Rani, N., & Shukla, S. K. (2025). *Generative AI in Financial Institution: A Global Survey of Opportunities, Threats, and Regulation.* arXiv.

Sandfreni, & Bansal, R. (2024). Challenges in Large Language Model Development and AI Ethics. In R. B. Sandfreni, *Challenges in Large Language Model Development and AI Ethics* (p. 57). Hershey, PA: IGI Global.

Schnepf, J., Engin, T., & Scheuermann, B. (2024). Studies on the Use of Large Language Models for the Automation of Business Processes in Enterprise Resource Planning Systems. *Studies on the Use of Large Language Models for the Automation of Business Processes in Enterprise Resource Planning Systems* (pp. 16-31). Springer. Retrieved from https://link.springer.com/chapter/10.1007/978-3-031-70239-6_2

Tavasoli, A., Sharbaf, M., & Madani, S. M. (2025). *Responsible Innovation: A Strategic Framework for Financial LLM Integration.* arXiv. Retrieved from https://arxiv.org/abs/2504.02165

Wardani, D. K., & Navarro, L. F. (2024, August 4). Conversational AI and Chatbot Systems for Enhancing Automated Billing, Payments, and Customer Support in SaaS Platforms. *Northern Reviews on Algorithmic Research, Theoretical Computation, and Complexity*. Retrieved from https://northernreviews.com/index.php/NRATCC/article/view/2024-08-04

Xu, J. (2024). GenAI and LLM for Financial Institutions: A Corporate Strategic Survey. *SSRN*.

Yilmaz, E., & Can, O. (2024). Unveiling Shadows: Harnessing Artificial Intelligence for Insider Threat Detection. *Engineering, Technology & Applied Science Research, 14*(2), 13341–13346. Retrieved from https://doi.org/10.48084/etasr.6911

Zhang, X., Zhou, Z., Liu, Z., Wang, Y., Zhang, Z., & Sun, L. (2024). *Privacy in Large Language Models: A Survey.* arXiv. Retrieved from https://arxiv.org/abs/2403.12503