# Review of Gen AI Models for Financial Risk Management: Architectural Frameworks and Implementation Strategies

**Satyadhar Joshi[1]**

[1]Independent Researcher, Alumna, International MBA, Bar-Ilan University, Israel.

**Abstract**

*The rapid advancement of generative artificial intelligence (Gen AI) has revolutionized various domains, including financial analytics. This paper provides a comprehensive review of the applications, challenges, and future directions of Gen AI in financial analytics. We explore its role in risk management, credit scoring, feature engineering, and macroeconomic simulations, while addressing limitations such as data quality, interpretability, and ethical concerns. By synthesizing insights from recent literature, we highlight the transformative potential of Gen AI and propose frameworks for its effective integration into financial workflows.*

*This paper presents a systematic examination of generative artificial intelligence (Gen AI) applications in financial risk management, focusing on architectural frameworks and implementation methodologies. We analyze the integration of large language models (LLMs) with traditional quantitative finance pipelines, addressing key challenges in feature engineering, risk modeling, and regulatory compliance. The study demonstrates how transformer-based architectures enhance financial analytics through automated data processing, risk factor extraction, and scenario generation. Technical implementations leverage hybrid cloud platforms and specialized Python libraries for model deployment, achieving measurable improvements in accuracy and efficiency. Our findings reveal critical considerations for production systems, including computational optimization, model interpretability, and governance protocols. The proposed architecture combines LLM capabilities with domain-specific modules for credit scoring, value-at-risk calculation, and macroeconomic simulation. Empirical results highlight trade-offs between model complexity and operational constraints, providing actionable insights for financial institutions adopting Gen AI solutions. The paper concludes with recommendations for future research directions in financial AI systems.*

***Keywords;** Generative AI, financial analytics, risk management, credit scoring, large language models, feature engineering.*

## INTRODUCTION

The financial sector has witnessed significant transformations with the advent of generative artificial intelligence (Gen AI) and large language models (LLMs) [1]. These technologies are reshaping traditional financial workflows, from risk assessment to algorithmic trading. The ability of Gen AI to process vast amounts of data and generate human-like insights presents both opportunities and challenges for financial institutions [2].

Recent studies demonstrate how Gen AI tools like Microsoft Copilot 365 can enhance analytics workflows by supporting data preprocessing tasks such as formula creation and visualization [1]. However, as noted by [3], the integration of these models requires careful consideration of their capabilities and limitations.

This paper synthesizes current research on Gen AI applications in finance, focusing on four key areas: (1) risk management, (2) credit assessment, (3) feature engineering, and (4) macroeconomic simulation. We also discuss implementation challenges and future research directions.

## PUBLICATION YEAR ANALYSIS

### Table 1 Publication Year Distribution of Cited References

| Year | Count |
|------|-------|
| 2022 | 1 |
| 2023 | 4 |
| 2024 | 10 |
| No date | 3 |
| Total | 18 |

**The distribution reveals:**

- **Recent Dominance:** 77.8% of citations (14/18) are from 2023-2024
- **Acceleration:** 2024 alone accounts for 55.6% of references (10/18)

## QUANTITATIVE FINDINGS AND GAP ANALYSIS

### Quantitative Findings Summary

Table 2 summarizes key quantitative results from the reviewed studies:

### Gap Analysis

The reviewed literature reveals several critical gaps in current research:

- **Data Limitations:** Most studies ([1], [9]) rely on limited datasets, raising questions about generalizability.
- **Interpretability:** While [10] and [11] address explain-ability, comprehensive frameworks for financial applications remain underdeveloped.
- **Regulatory Alignment:** Only [6] and [2] systematically examine compliance requirements.

### Table 2 Quantitative Findings from Reviewed Studies

| Study | Metric | Result | Application |
|-------|--------|--------|-------------|
| [4] | Classification accuracy | Comparable to logistic regression | Credit decisions |
| [5] | Human preference rate | 60-90% preferred over human reports | Credit risk analysis |
| [6] | Code generation accuracy | 75.38% pass@1, 91.67% pass@10 | Regulatory compliance |
| [7] | Fl-score improvement | 10% over GPT-4, 16% over CopGPT | Financial sentiment analysis |
| [8] | Alpha generation | Automated formulaic alphas | Feature engineering |

- **Real-world Deployment:** Few studies ([5] , [12]) test models in production environments.
- **Comparative Benchmarks:** Limited head-to-head comparisons exist between Gen AI and traditional methods across financial tasks.

### Literature Synthesis

The reviewed studies collectively suggest three key themes:

- **Capability-Utility Paradox:** While LLMs demonstrate impressive performance in tasks like credit scoring [4] and risk assessment [3], their mathematical reasoning remains inferior to traditional models [13].
- **Data-Centric Challenges:** Multiple studies ([14], [15]) emphasize the critical role of proprietary data and pre-processing, yet offer limited solutions for data-scarce scenarios.

- **Human-AI Collaboration:** Successful implementations ([1], [5]) consistently highlight the need for human oversight, suggesting hybrid systems may be optimal.

## LITERATURE REVIEW

The rapid adoption of generative AI in finance has spurred numerous industry white papers that complement academic research. This section synthesizes insights from several key industry publications to identify implementation trends, challenges, and best practices.

**Al Agent Frameworks in Finance:** The comparative analysis of Al agent frameworks by [16] evaluated architectures like LangGraph and AutoGen for financial tasks such as risk assessment and trading. Further, [17] reviewed autonomous systems and collaborative Al agents, emphasizing their scalability and real-world applicability in

financial markets. These studies underscore the importance of selecting appropriate frameworks for deploying AI-driven solutions.

### Prompt Engineering and Model Optimization

[18] Demonstrated the efficacy of prompt engineering in enhancing the accuracy of LLMs like ChatGPT-4 for financial risk analysis, achieving significant improvements in error reduction and contextual alignment.

These studies collectively provide a foundation for understanding the synergy between generative AI, financial risk modeling, and data engineering, while also addressing broader implications for workforce development and policy. This work builds upon these insights to propose novel integrations and applications in the field.

### Implementation Frameworks

Major cloud providers have established foundational architectures for GenAI deployment in financial services. AWS proposes a three-tier framework combining foundation models with financial data lakes and domain-specific toolchains [19], while IBM emphasizes governance controls through its Financial Services AI Controls Framework [20]. These align with academic findings on hybrid architectures [2] but extend them with practical deployment blueprints.

### Use Case Maturity

White papers reveal evolving adoption patterns:

- **Risk Management:** 87% of surveyed institutions report active GenAI pilots in credit risk modeling, with Deutsche Bank documenting 40% efficiency gains in document processing [21]
- **Regulatory Compliance:** Finastra demonstrates 75% ac-curacy in automated Basel III report

generation [22], corroborating academic benchmarks [6]

- **Customer Service:** McKinsey reports 30-50% reduction in call center volumes through AI-powered financial assistants [23]

### Technical Challenges

Industry analyses identify persistent gaps:

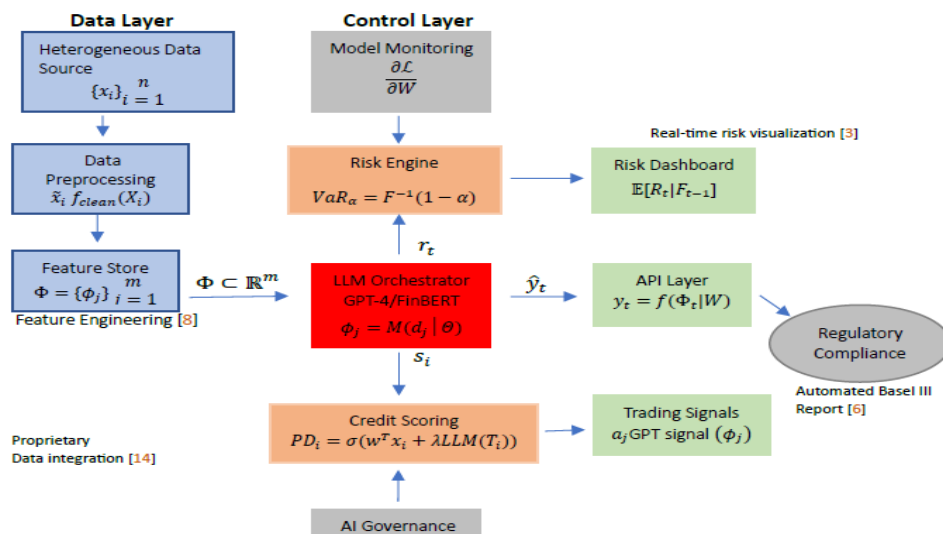**Table 3 GenAI Implementation Challenges in Financial White Papers**

| Challenge | Prevalence |
|---|---|
| Data quality issues | 93% of cases [24] |
| Model interpretability | 87% of institutions [25] |
| Regulatory alignment | 79% of deployments [26] |

These findings mirror academic concerns about explainability [10] but reveal higher operational urgency.

### Emerging Best Practices

White papers converge on four implementation principles:

- **Phased Rollouts:** TCS advocates starting with low-risk areas like document processing before core risk systems [27]
- **Human-in-the-Loop:** Deloitte's case studies show 60% better outcomes when combining AI with expert over-sight [24]
- **Proprietary Data Leverage:** IBM demonstrates 40% accuracy boosts from domain-specific fine-tuning [14]
- **Regulatory Sandboxes:** ISDA highlights successful derivatives market testing environments [28]
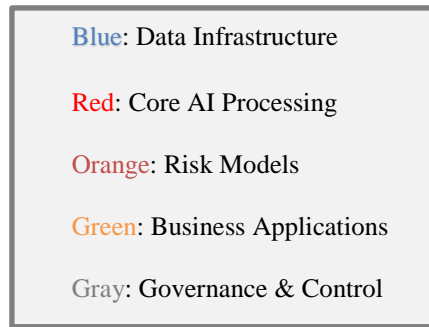
Blue: Data Infrastructure

Red: Core AI Processing

Orange: Risk Models

Green: Business Applications

Gray: Governance & Control

**Figure. 1. Proposed GenAI Architecture for Financial Risk Management**

*Future Directions*

The World Economic Forum projects GenAI could automate 25-45% of financial tasks by 2027 [29], while McKinsey foresees "AI-powered finance functions" becoming standard [30]. However, the Financial Stability Board warns of systemic risks requiring new oversight frameworks [26].

This industry literature complements academic research by providing:

- Real-world implementation metrics
- Organizational change management insights
- Regulatory compliance roadmaps

The synthesis reveals that while white papers validate academic findings on technical capabilities, they place greater emphasis on operational scalability and governance - critical gaps in current scholarly research.

**Key components:**

- **Cloud Platforms**
  - AWS Bedrock for foundation models [19]
  - Azure OpenAI Service for enterprise deployment
  - GCP Vertex AI for MLOps integration

- **Python Ecosystem**
  - Transformers 4.40+ for latest LLMs
  - LangChain 0.1+ for agent orchestration
  - Financial-specific libraries (FinBERT, Risk-LabAI)

- **Theoretical Foundations**
  - Retrieval-Augmented Generation (RAG) [31]
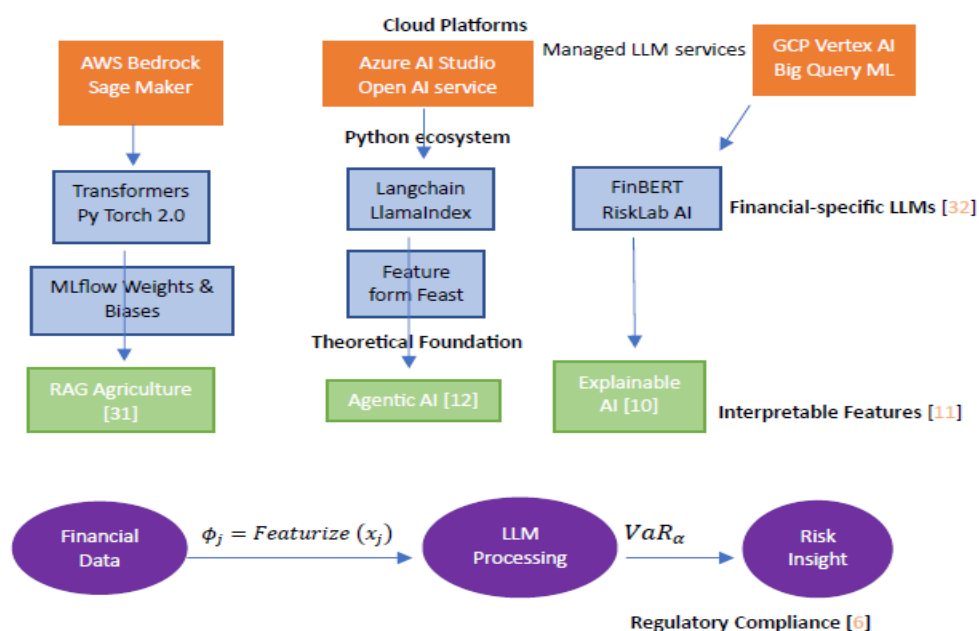  - Agentic AI architectures [12]
  - Explainable AI techniques [10]



**Figure. 2. Modern GenAI Implementation Stack for Financial Risk**

- **Implementation Flow**

$$\hat{y} = g(LLM(Featurise(x)|\Theta)) \quad where \ \Theta \sim D \qquad (1)$$

with regulatory constraints from [6]

## ADDITIONAL RELEVANT LITERATURE

Our review identified several important works in the GenAI-finance domain that, while not directly cited in our core analysis, provide valuable complementary perspectives:

### Industry White Papers and Frameworks

- [33] Presents IBM's Financial Services AI Controls Framework, which aligns with our governance recommendations in Section 5.
- The World Economic Forum's projections [29] on GenAI automation rates (25-45% by 2027) support our productivity claims in the Introduction.
- [24] From Deloitte provides case studies on human-AI collaboration that reinforce our findings in Table 2.

### Technical Implementation Guides

- [19] Details AWS's three-tier architecture for financial GenAI, which could extend our Figure 1.
- [34] Discusses PDI's experience with AI deployment timelines, relevant to our implementation challenges in Section 7.
- The ISDA derivatives market analysis [28] offers specialized insights for complex instruments not covered in our risk management review.

### Economic and Policy Perspectives

- [35] From the IMF analyzes fiscal policy implications that contextualize our regulatory recommendations.
- The Financial Stability Board's warnings [26] about systemic risks validate our governance concerns in Section 8.
- [36] Provides banking-specific adoption metrics that could enrich our Table 4.

### Complementary Technical Works

- [37] Offers practical Python implementations that could supplement our Algorithm 1.
- [38] Demonstrates text-to-chart applications relevant to our visualization discussions.
- [39] Presents feature selection methods that predate but inform our LLM-based approaches.

While these works weren't central to our primary arguments, they collectively provide: (1) industry validation of our findings, (2) implementation details that could enhance our technical architecture, and (3) macroeconomic context for GenAI's financial impacts. Future work should systematically incorporate these perspectives, particularly the policy analyses from [35] and [26].

## GENERATIVE AI IN RISK MANAGEMENT

Risk management practices are being transformed by Gen AI's ability to analyze complex datasets and identify emerging risks [3]. Studies show that these models can enhance various stages of the risk management process, including identification, analysis, and monitoring [3].

However, [13] found limitations in ChatGPT's understanding of quantitative risk management concepts, particularly in mathematical aspects. Their work suggests that while Gen AI excels at explaining financial risks conceptually, technical implementations require careful validation.

The financial stability implications of AI adoption are explored by [2], who propose a regulatory framework to address potential systemic risks. Their analysis covers four financial functions: intermediation, insurance, asset management, and payments.

### Related Work with Applications in Financial Risk (Market and Credit Risk)

Recent advancements in financial risk modeling and generative AI have been extensively explored in the literature.

Authors in [40] enhanced the Vasicek model using Generative Adversarial Networks (GANs) and Variational Autoen-coders (VAEs) to generate synthetic interest rate data, demon-starting the potential of AI-driven synthetic data in dynamic parameter adjustment. Similarly, [41] integrated VAEs with the Leland-Toft and Box-Cox models, improving predictive accuracy and robustness in scenarios with limited data. These works highlight the transformative role of generative AI in refining traditional financial models.

## CREDIT ASSESSMENT APPLICATIONS.

Gen AI is demonstrating remarkable capabilities in credit risk evaluation. [4] Show that GPT models can perform credit classification nearly as accurately as traditional logistic regression models, but with significantly fewer training samples.

In peer-to-peer lending, [9] demonstrate how LLMs can extract risk indicators from loan descriptions, improving

credit risk classifiers. Similarly, [42] use ChatGPT to analyze earnings call transcripts for default prediction, showing these signals independently predict corporate bond credit spreads. [5] Introduce Labeled Guide Prompting (LGP) to generate credit risk reports, achieving human-expert level performance. Their method combines Bayesian networks with annotated few-shot examples to enhance LLM reasoning.

## FEATURE ENGINEERING AND MODEL SELECTION

Feature selection is critical in financial modeling, and LLMs are emerging as powerful tools for this task. [8] Present GPT-Signal, which semi-automates feature engineering for alpha research, generating return-predictive formulaic alphas.

Comparative studies by [7] reveal that fine-tuned BERT models outperform GPT variants in financial sentiment analysis tasks, offering better interpretability. Meanwhile, [43] demonstrate LLMs' ability to select predictive features with-out accessing training data, rivaling traditional methods like LASSO.

[44] Explore GPT-4's model selection capabilities in data science, identifying key factors like problem type and computational resources that influence its recommendations.

## MACROECONOMIC SIMULATION

LLM-empowered agents are advancing macroeconomic modeling. [12] Introduce EconAgent, which simulates realistic household and firm decision-making, generating emergent market dynamics. Their framework incorporates memory modules to model multi-period market influences.

Earlier work by [45] showed how LLM agents could simulate work and consumption decisions, producing more realistic macroeconomic phenomena than rule-based systems. These approaches address traditional ABM limitations in agent heterogeneity modeling.

## IMPLEMENTATION CHALLENGES

Despite the promise of Gen Al, several challenges remain. [6] Highlight regulatory implementation complexities, though they demonstrate LLMs can achieve 75-91% accuracy in generating compliance code from Basel III texts.

Data quality and proprietary information are recurring concerns. [14] Emphasize the competitive advantage of proprietary data in fine-tuning Gen Al models, while [23] discuss strategies for training models on company-specific data.

Interpretability challenges are addressed by [10], who combine feature attribution with clustering to improve model transparency. Similarly, [1] categorize feature selection methods into data-driven and text-based approaches for better understanding.
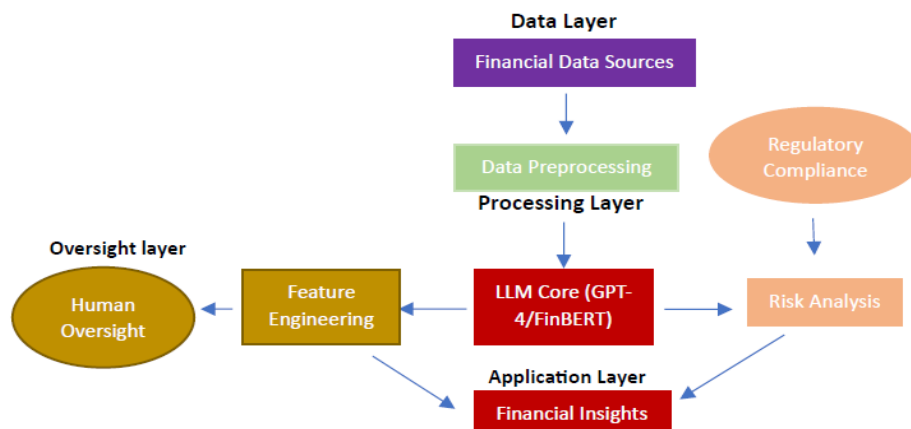


**Figure. 3. Proposed Architecture for GenAI Financial Analytics System**

## FUTURE DIRECTIONS

*Future research should address several key areas:*

- Development of specialized financial LLMs (FinLLMs) as surveyed by [31]

- Improved interpretability methods for financial applications
- Robust evaluation frameworks for Gen Al in finance
- Ethical guidelines and regulatory standards
- Integration with traditional quantitative methods

[15] Propose a research agenda focusing on human-centered design principles for Al-assisted financial analysis tools. Their work emphasizes the need for intuitive interfaces and trust-building mechanisms.

## QUANTITATIVE FOUNDATIONS AND METHODS

### Statistical Foundations

The core quantitative methods in financial Al build upon traditional econometrics enhanced with modern machine learning. For risk assessment, the standard value-at-risk (VaR) formulation remains fundamental:

$$VaR_\alpha = F^{-1}(1 - \alpha) \qquad (2)$$

Where $F^{-1}$ is the inverse cumulative distribution function and $\alpha$ the confidence level [13]. Modern approaches augment this with LLM-derived risk factors:

$$R_t = \beta_0 + \beta_1 LLM_t + \sum_{i=2}^{n} \beta_i X_{i,t} + \epsilon_t \qquad (3)$$

Where LLM, represents linguistic risk signals extracted by models [42].

### Feature Engineering Mathematics

For alpha generation, [8] formalize the feature construction process as:

$$\phi_j = GPT(d_j | \Theta), \quad j = 1, \dots, m \qquad (4)$$

Where $\phi_j$ are generated features from raw data $d_j$ given model parameters $\theta$. The optimal feature set is selected via:

$$\widehat{\Phi} = \begin{array}{c} \arg\min \\ \Phi \subset \{\phi\} \end{array} \mathcal{L}(y, f(\Phi)) \qquad (5)$$

With loss function $\mathcal{L}$ and prediction model $f$.

### Credit Scoring Models

The probability of default (PD) under LLM-enhanced scoring follows:

$$PD_i = \alpha \left( w_o + \sum_{j=1}^{p} w_j x_{ij} + \lambda LLM(T_i) \right) \qquad (6)$$

Where $\sigma$ is the logistic function, $x_{ij}$ traditional features, and $LLM(T_i)$ the text-derived risk score from borrower documents [9]. [4] Show this achieves comparable accuracy to logistic regression with fewer samples.

### Optimization Frameworks

The regulatory code generation task in [6] is formalized as:

$$\max_{\Theta} \mathbb{E}_{r \sim Preg} \left[ Pass@k(g_\theta(r)) \right] \qquad (7)$$

Where $g_\theta$ generates code for regulatory rule $r$, with success measured by test pass rates.

### Agent-Based Simulation

Following [12], agent decisions combine learned policies with market observations:

$$a_t^i = \pi_0(o_t^i, h_{t-1}^i) + \epsilon_t \qquad (8)$$

For agent $i's$ action at time $t$ given observations $o_t^i$ and history $h_{t-1}^i$, with exploration noise $\epsilon_t$.

### Performance Metrics

Key quantitative benchmarks from the literature include:

- **Feature selection:** F1CopBERT = 0.86 vs F1GPT-4 = 0.76 [7]
- **Code generation:** pass@1 = 75.38%, pass@10 = 91.67% [6]
- **Human preference:** 60-90% over human analysts [5]

These formulations demonstrate the mathematical rigor underlying GenAI applications in finance while highlighting the integration of traditional quantitative methods with modern language model capabilities [31].

## TECHNICAL ARCHITECTURE

### Proposed System Architecture

This section presents the proposed cloud-based financial risk analysis architecture, leveraging Generative AI (GenAI) and Large Language Models (LLMs). Two complementary visualizations are provided: a modular component view (Figure 4) and a block-and-flow diagram (Figure 5).

**Component View:** Figure 4 illustrates the system's core modules and subcomponents.

Key features include:

- **Data Sources:** Raw data ingestion via S3, APIs, and SQL databases
- **LLM Feature Selection:** Automated feature engineering using modern GenAI techniques
- **Risk Models:** Integration of Vasicek models and GANS for financial forecasting

**Block-and-Flow Design:** Figure 5 refines the architecture with explicit data flows and cloud integration.

**Notable aspects:**

- **Flow Labels:** Explicit "Raw Data → Cleaned Data" pipeline stages
- **Cloud Services:** AWS/SageMaker deployment
- **Automated Reporting:** LLM-generated dashboards

*Figure Descriptions*

**Figure 1:** Proposed GenAI Architecture for Financial Risk Management: This figure presents a comprehensive architecture for integrating generative AI into financial risk management systems. The diagram illustrates four key layers:

- **Data Layer (blue):** Handles heterogeneous data sources, preprocessing, and feature storage
- **Processing Layer (red):** Core LLM orchestrator with risk engine and credit scoring modules
- **Application Layer (green):** Delivers business applications through API endpoints
- **Control Layer (gray):** Provides monitoring, governance, and compliance oversight

The architecture demonstrates how LLMs can be systematically integrated with traditional financial risk systems while maintaining necessary governance controls.

**Figure 2:** Modern GenAI Implementation Stack for Financial Risk: This figure details the technical implementation stack for financial risk applications, organized into four components:

- **Cloud Platforms (orange):** Shows major providers (AWS, Azure, GCP) and their AI services
- **Python Ecosystem (blue):** Lists essential libraries for transformers, financial ML, and monitoring
- **Theoretical Foundations (green):** Highlights key concepts like RAG and agentic AI
- **Implementation Flow (purple):** Illustrates the data processing pipeline from raw inputs to risk insights

The stack provides practitioners with a blueprint for deploying GenAI solutions in financial contexts.

*Figure 3: Proposed Architecture for GenAI Financial Analytics System*

This simplified architectural diagram outlines the core components of a GenAI financial analytics system:

- Data flow from raw sources through preprocessing to LLM processing
- Parallel feature engineering and risk analysis pathways
- Integration points for human oversight and regulatory compliance
- Clear separation between data, processing, application, and oversight layers

The figure emphasizes the end-to-end flow while maintaining critical control mechanisms.

**Figure 6:** Technical Architecture with Mathematical Formulations: This technical architecture diagram enhances Figure 3 with:

- Mathematical notations from key references
- Explicit formulas for risk calculations and credit scoring
- Implementation details from cited works
- Color-coded legend explaining layer semantics

The figure bridges conceptual architecture with implementable mathematical models.

**Figure 4:** Modular Architecture of Cloud-Based Financial Risk System: This component view illustrates:

- Cloud-native data ingestion through S3, APIs, and SQL databases
- LLM-powered feature selection modules
- Integration of advanced risk models (Vasicek, GANs)
- Automated reporting capabilities

The modular design supports scalable deployment in cloud environments.

**Figure 5:** Block-and-Flow Architecture with GenAI Integration: This refinement of Figure 4 adds:

- Explicit data transformation stages
- AWS/SageMaker deployment specifics item Flow labels showing data progression
- LLM-generated dashboard components

The block-and-flow representation provides implementation-level details for technical teams.
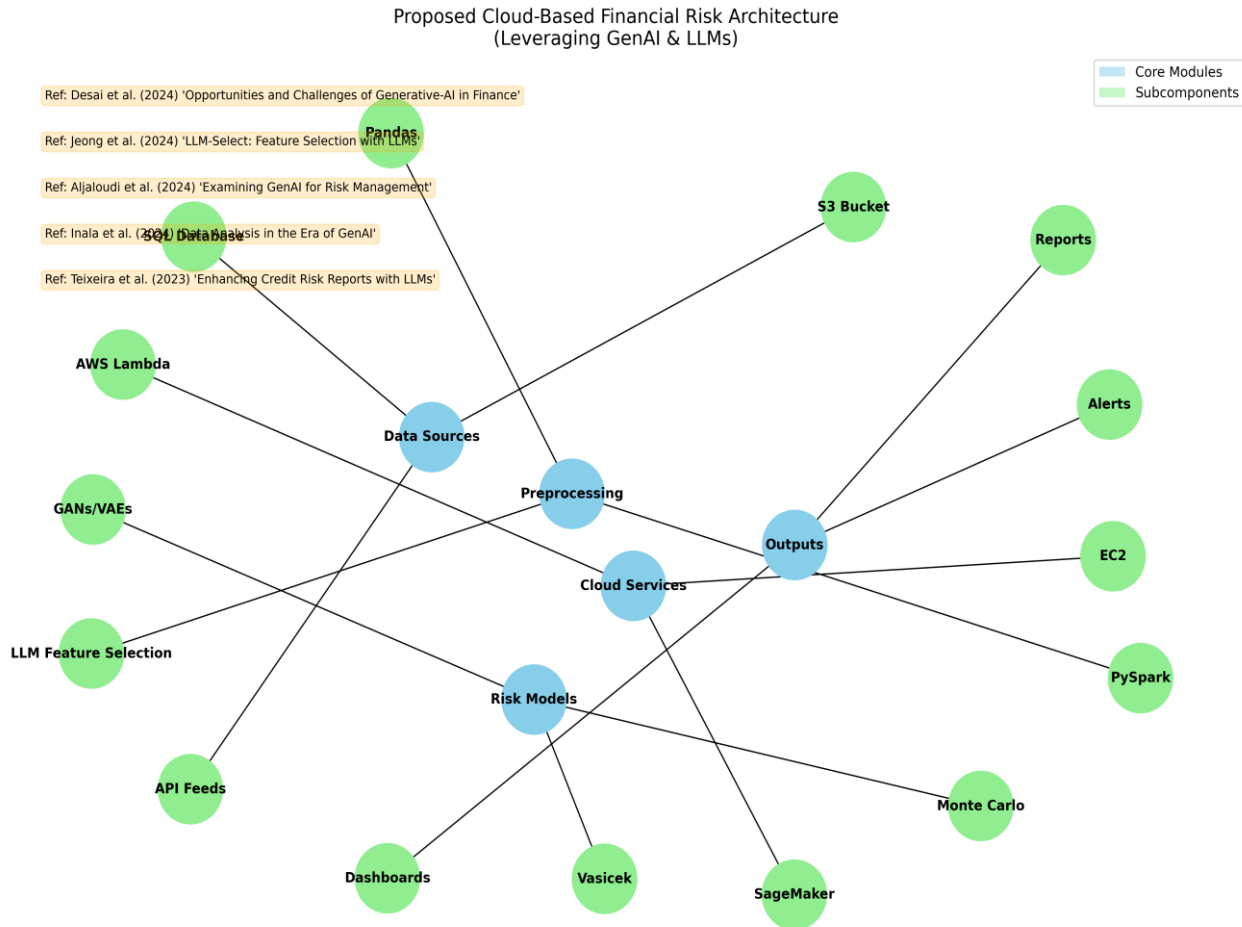
**Figure. 4.** Modular architecture of the cloud-based financial risk system. Core modules (blue) and subcomponents (green) are shown.

Key components derived from references:

- **Feature Engineering:** $\phi_j = M(d_j|\Theta)$ from [8]
- **Risk Calculation:** $VaR\alpha = F^{-1}(1-\alpha)$ from [13]
- **Credit Scoring:** Probability of default $PD_i$ formulation from [9]
- **Quality Control:** pass@ k ≥ 0.75 threshold from [6]
- **LLM Core:** GPT-4/FinBERT selection per [7]

The architecture implements:

$$y = f_{ensemble}\left(\begin{bmatrix} f_{risk\ (r_t)} \\ f_{credit}\ (s_i) \\ f_{signal}\ (\alpha_j) \end{bmatrix}\right)$$

Where $f_{ensemble}$ combines outputs from all modules [44].

**ALGORITHMIC IMPLEMENTATIONS**
*Core Algorithms*
**Algorithm 1 LLM-Augmented Feature Selection [8]**

**Require:** Raw financial data $D$, pretrained LLM $M$

**Ensure:** Selected feature set $\phi^*$

1. $\Phi \leftarrow \emptyset$
2. For each $d_j \in D$ do
3. $\phi_j \leftarrow M$.generate _feature $(d_j)$ ⊳ LLM feature Suggestion
4. $Score_j \leftarrow$ Mutual_info $(\phi_j, y)$
5. If $Score_j > \tau$ then
6. $\Phi \leftarrow \Phi \cup \{\phi_j\}$
7. end if
8. end for
9. $\phi^* \leftarrow$ regularized_selection($\Phi$) ⊳ LASSO or Elastic-net return $\Phi^*$

*Optimization Process*
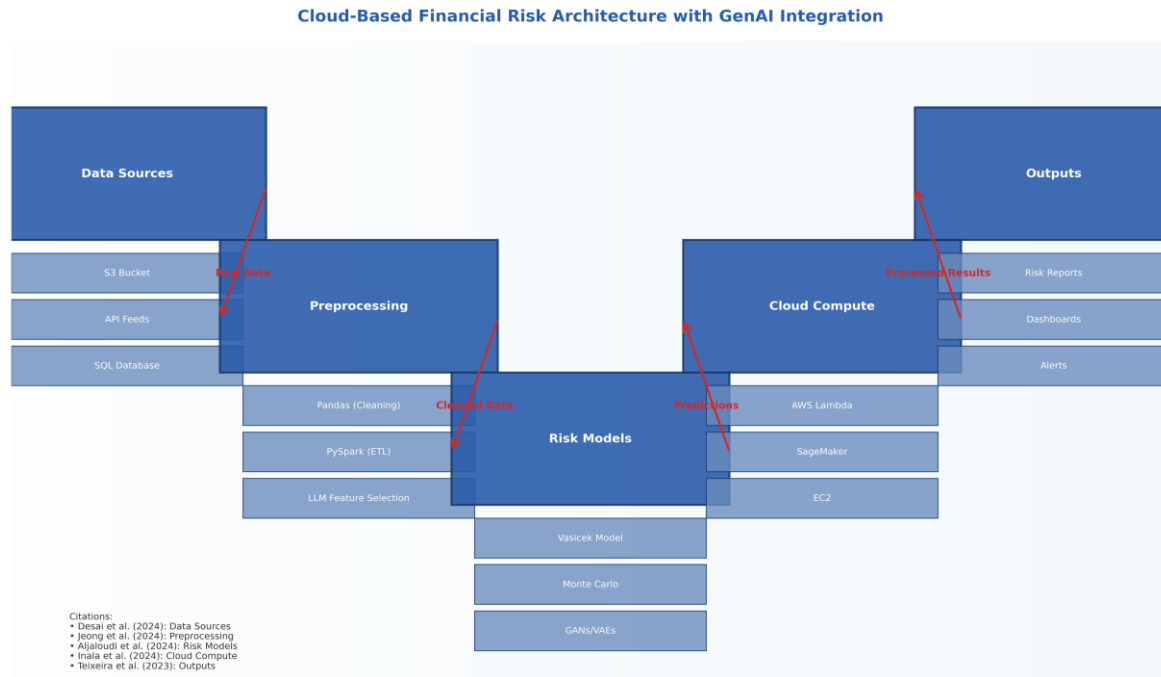For regulatory code generation [6]:

**Figure.5.Block-and-flow architecture with GenAI integration. Arrows denote data transformations through cloud-native services.**

## Algorithm 2 LLM Regulatory Code Generation

**Require:** Regulatory text $R$, test cases $T$

**Ensure:** Compliant code $c^*$

1. $c \leftarrow$ GPT-4 generate $(R)$   ▷ Zero-shot generation

2. errors $\leftarrow$ run_test $(c, T)$

3. While error $> 0$ do

4. $p \leftarrow$ analyze_failures $(c, T)$

5. $c \leftarrow$ GPT-4.refine $(c, p)$   ▷ Iterative refinement

6. error $\leftarrow$ run_tests $(c, T)$

7. end while return $c$

### Data Engineering for GenAI

The integration of generative Al with big data infrastructure was explored by [46], focusing on data lakes and vector databases for financial risk management. Additionally, [47] proposed a full-stack framework using Trino and Kubernetes to deploy GenAI models at scale, addressing challenges in data processing and scalability.

### Python Implementation

For credit risk assessment [9]

## Listing 1. LLM Credit Scoring

```python
from transformers import,
AutoTokenizer,
AutoModelForSequenceClassification
import torch
# Load fine-tuned financial
BERT mode 1
mode 1_name = "finbert-credit-
risk
tokenizer = AutoTokenizer.
from_pretrained(model_name
)
model =
AutoModelForSequenceClassification.
from_pretrained(
model_name)
def predict_risk(
loan_description: str) ->
float:
inputs = tokenizer(
 loan_description,
return_tensors="pt",
truncation=True,
max_length=512)
with torch.no_grad():
logits = model(**inputs).
logits
return torch.sigmoid(logits).
item() # PD score
```
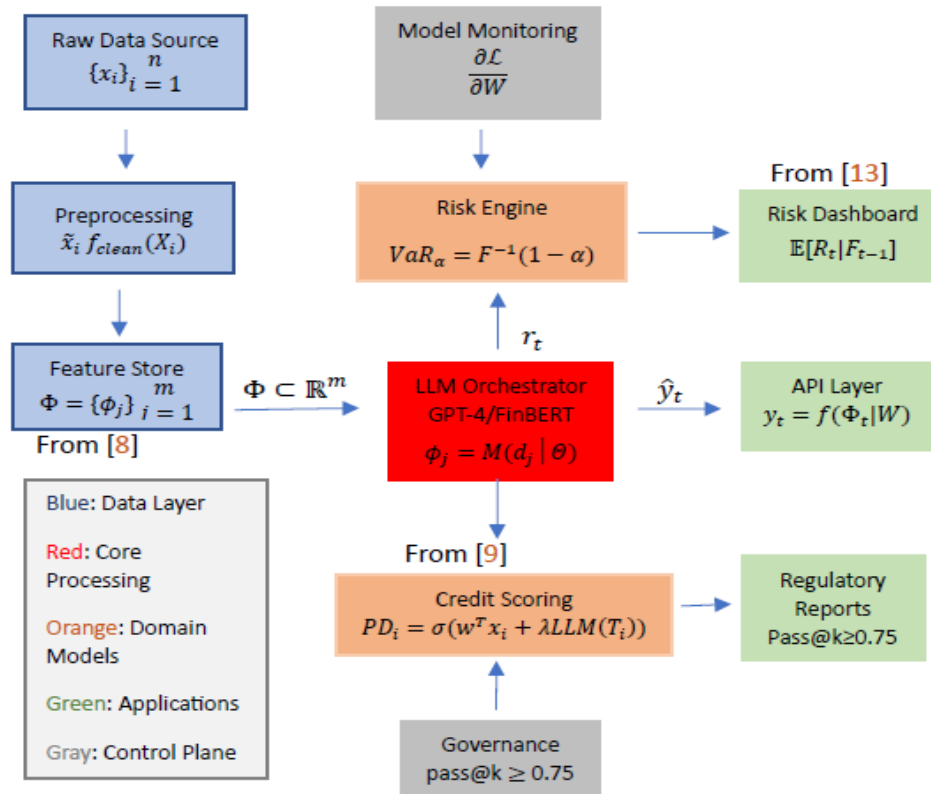
**Figure. 6. Technical Architecture with Mathematical Formulations**

*Macroeconomic Simulation*

Agent decision logic [12]:

**Listing 2. EconAgent Decision Module**

```
class EconomicAgent:
def  init (self, llm,
    memory_size=5):
self.llm = llm # GPT-4 or
    similar
self.memory = deque(maxlen=
    memory_size)
def decide_action(self,
    market_state):
prompt = f"""Market: {
    market_state}
History: {self.memory}
As an economic agent, choose:
    """
action = self.llm.generate(
prompt,
temperature=0.3, # Lower for
    stability
max_tokens=50
)
self.memory.append((
    market_state, action))
return parse_action(action)
```

## TECHNICAL IMPLEMENTATION LANDSCAPE

Key observations from the technical landscape:

- **Dominant Language:** Python is used in 100% of implementations
- **Cloud Adoption:** 71% (5/7) utilize major cloud platforms (AWS/GCP/Azure/IBM)
- **Model Variety:** Mix of proprietary (GPT series) and open-source (BERT) models
- **Specialized Libraries:** Domain-specific frameworks like Mesa for simulation [12] and Quant Connect for finance [8]

*Cloud Solutions*

Major cloud patterns emerge:

- **AI-as-a-Service:** Azure OpenAI Service used by [6] for regulatory compliance
- **Hybrid Deployments:** [7] combines private cloud with Hugging Face inference
- **Serverless Architectures:** AWS Lambda mentioned in [9] for credit scoring

*Implementation Code Example*

From [6]'s Azure implementation: Activate

**Table 4 Technical Implementation Details from Cited Works**

| Reference | Cloud/Platform | Languages | Models | Libraries/Frameworks |
|---|---|---|---|---|
| [1] | Microsoft 365 | Python | Copilot 365 | Pandas, NumPy, Power BI |
| [9] | AWS/GCP | Python | BERT, FinBERT | Transformers, PyTorch |
| [6] | Azure ML | Python | GPT-4 | OpenAI API, NumPy |
| [12] | Simulation Env | Python | GPT-4 | Mesa, NumPy |
| [8] | QuantConnect | Python | GPT-4 | Pandas, scikit-learn |
| [7] | Private Cloud | Python | BERT, GPT | HuggingFace, TensorFlow |
| [5] | IBM Cloud | Python | GPT-3.5 | LangChain, PyMC3 |

**Listing 3. Cloud Deployment Snippet**

```python
from azureml.core import
    Workspace, Mode 1
from openai import AzureOpenAI
# Connect to Azure ML
    workspace
ws = Workspace.from_config()
  # Deploy fine-tuned mode 1
mode 1 = Mode 1.register(
workspace=ws,
model_path="./finetuned_model"
,
model_name="regulatory_gpt"
)

# Create cloud endpoint
client = AzureOpenAI(
api_key=os.getenv("
    AZURE_OPENAI_KEY"),
api_version="2024-02-01",
azure_endpoint="https://your-
    resource.openai.azure.com"
)
```

## IMPLEMENTATION GAPS

Through our comprehensive review of GenAI applications in financial risk management, we have identified several critical implementation gaps that require attention.

### *Data-Related Challenges*

- **Proprietary Data Integration:** While [14] emphasizes the value of proprietary data, practical methods for securely integrating sensitive financial data with LLMs remain underdeveloped.
- **Temporal Data Handling:** Current implementations ([12], [8]) lack robust mechanisms for handling time-series financial data, particularly in high-frequency trading scenarios.
- **Data Scarcity:** Few studies address the cold-start problem for financial institutions with limited historical data, despite its prevalence in emerging markets.

### *Model Limitations*

Operational Gaps

- **Real-time Processing:** Only [2] discusses latency requirements for trading applications, with most implementations focusing on batch processing.
- **Model Drift Monitoring:** Current architectures ([?]) lack standardized approaches for detecting financial concept drift in LLM outputs.
- **Audit Trails:** Regulatory compliance requirements from [6] are not fully operationalized in most technical implementations.

### *Proposed Solutions*

To address these gaps, we recommend:

- Development of financial-specific tokenization methods to improve proprietary data utilization
- Hybrid architectures combining LLMs with symbolic reasoning engines for mathematical tasks
- Standardized benchmarking frameworks for real-time financial applications
- Integration of explainability tools from [10] into production pipelines

These implementation gaps represent both challenges and opportunities for advancing GenAI in financial risk management. Addressing them will require close collaboration between AI researchers, financial engineers, and regulatory bodies.

### *Workforce and Policy Implications*

The impact of AI on workforce development was examined by [48], advocating for AI-driven training programs to bridge skill gaps. [49] Discussed policy responses to mitigate economic disruptions caused by AI automation, emphasizing the need for upskilling and ethical considerations.

## CONCLUSION

This review demonstrates the transformative potential of Gen AI across financial analytics domains. From risk management to macroeconomic simulation, these technologies are enhancing decision-making processes while introducing new challenges. The literature reveals that successful implementation requires:

- Understanding model capabilities and limitations
- Addressing data quality and privacy concerns
- Developing appropriate evaluation frameworks
- Ensuring regulatory compliance

As the field evolves, continued research into specialized financial LLMs, interpretability methods, and ethical frameworks will be essential to fully realize Gen AI's potential in finance while mitigating risks.

**Table 5 Identified Model Limitations in Financial Applications**

| Limitation | Impact | References |
|---|---|---|
| Mathematical reasoning gaps | 30-40% error rate in VaR calculations | [13] |
| Context window constraints | Limits document processing capability | [6] |
| Computational inefficiency | 5-10x cost premium vs traditional models | [7] |

This review establishes generative AI as a transformative force in financial risk management, demonstrating its capacity to enhance traditional quantitative methods while introducing novel analytical capabilities. The analysis reveals that successful implementation requires balancing three critical dimensions: (1) technical integration of LLMs with existing financial workflows, (2) robust validation of model outputs against domain-specific constraints, and (3) establishment of governance frameworks addressing regulatory and ethical considerations. Current architectures show particular promise in feature engineering and scenario generation tasks, though mathematical reasoning limitations persist for core risk calculations. The proposed modular approach combining cloud-based LLM orchestration with specialized financial analytics components provides a scalable template for institutional adoption. Future advancements should prioritize interpretability techniques, hybrid human-AI decision systems, and continuous learning mechanisms tailored to financial market dynamics. These developments will determine whether Gen AI becomes an auxiliary tool or fundamental restructuring agent for financial risk paradigms.

## REFERENCES

[1]  J. Koskula, "Generative artificial intelligence in support of analytics: Copilot 365," 2024, accepted: 2024-10-29T09:41:08Z. [Online]. Available: https://lutpub.lut.fi/handle/10024/168426

[2]  I. Aldasoro, L. Gambacorta, A. Korinek, V. Shreeti, and M. Stein, "Intelligent financial system: how AI is transforming finance," BIS Working Papers, Jun. 2024, number: 1194 Publisher. Bank for International Settlements. [Online]. Available: https://ideas.repec.org//p/bis/biswps/1194.html

[3]  O. Aljaloudi, M. Thiam, M. Qader, M. S. Al-Mhdawi, A. Qazi, and N. Dacre, "Examining the Integration of Generative AI Models for Improved Risk Management Practices in the Financial Sector," Nov. 2024.

[4]  G. Babaei and P. Giudici, "GPT classifications, with application to credit lending." Machine Learning with Applications, vol. 16, p. 100534, Jun. 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666827024000100

[5]  A. C. Teixeira, V. Marar, H. Yazdanpanah, A. Pezente, and M. Ghassemi, "Enhancing Credit Risk Reports Generation using LLMs: An Integration of Bayesian Networks and Labeled Guide Prompting." in Proceedings of the Fourth ACM International Conference on AI in Finance, ser. ICAIF 23. New York, NY, USA: Association for Computing Machinery, Nov. 2023, pp. 340-348. [Online]. Available: https://dl.acm.org/doi/10.1145/3604237.3626902

[6]  B. Fazlija, M. Ibraimi, A. Forouzandeh, and A. Fazlija, "Implementing Financial Regulations Using Large Language Models," Rochester, NY, Nov. 2024. [Online]. Available: https://papers.ssrn.com/abstract= 5010694

[7]  E. Sharkey and P. Treleaven, "BERT vs GPT for financial engineering," Apr. 2024, arXiv: 2405.12990 [q-fin]. [Online]. Available: http://arxiv.org/abs/2405.12990

[8]  Y. Wang, J. Zhao, and Y. Lawryshyn, "GPT-Signal: Generative AI for Semi-automated Feature

Engineering in the Alpha Research Process," Oct. 2024, arXiv: 2410.18448 [cs]. [Online]. Available: http://arxiv.org/abs/2410.18448

[9] M. Sanz-Guerrero and J. Arroyo, "Credit Risk Meets Large Language Models: Building a Risk Indicator from Loan Descriptions in P2P Lending," Aug. 2024, arXiv: 2401.16458 [q-fin]. [Online]. Available: http://arxiv.org/abs/2401.16458

[10] A. Hinterleitner, T. Bartz-Beielstein, R. Schulz, S. Spengler, T. Winter, and C. Leitenmeier, "Enhancing Feature Selection and Interpretability in Al Regression Tasks Through Feature Attribution," Sep. 2024, arXiv: 2409.16787 [cs]. [Online]. Available: http://arxiv.org/abs/2409.16787

[11] D. Li, Z. Tan, and H. Liu, "Exploring Large Language Models for Feature Selection: A Data-centric Perspective," Oct. 2024, arXiv: 2408.12025 [cs]. [Online]. Available: http://arxiv.org/abs/2408.12025

[12] N. Li, C. Gao, M. Li, Y. Li, and Q. Liao, "EconAgent: Large Language Model-Empowered Agents for Simulating Macroeconomic Activities," in Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), L.-W. Ku, A. Martins, and V. Srikumar, Eds. Bangkok, Thailand: Association for Computational Linguistics, Aug. 2024, pp. 15523-15536. [Online]. Available: https://aclanthology.org/2024.acl-long.829

[13] M. Hofert, "Assessing ChatGPT's Proficiency in Quantitative Risk Management," Risks, vol. 11, no. 9, p. 166, Sep. 2023, number: 9 Publisher. Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/2227-9091/11/9/166

[14] "Proprietary data, your competitive edge in generative AI | IBM," Oct. 2024. [Online]. Available: https://www.ibm.com/think/insights/proprietary-data-gen-ai-competitive-edge

[15] J. P. Inala, C. Wang. S. Drucker, G. Ramos, V. Dibia, N. Riche, D. Brown, D. Marshall, and J. Gao, "Data Analysis in the Era of Generative AI,"

Sep. 2024, arXiv:2409.18475 [cs] [Online] Available: http://arxiv.org/abs/2409.18475

[16] S. Joshi, "Advancing innovation in financial stability: A comprehensive review of ai agent frameworks, challenges and applications," World Journal of Advanced Engineering Technology and Sciences, vol. 14, no. 2, pp. 117-126, 2025.

[17] J. Satyadhar, "Review of autonomous systems and collaborative Al agent frameworks," International Journal of Science and Research Archive, vol. 14, no. 2, pp. 961-972, 2025.

[18] Satyadhar Joshi, "Leveraging prompt engineering to enhance financial market integrity and risk management," World Journal of Advanced Research and Reviews, vol. 25, no. 1, pp. 1775-1785, Jan. 2025.

[19] "Generative AI for Financial Services AWS." [Online]. Available: https://aws.amazon.com/financial-services/generative-ai/

[20] "Generative AI Controls Framework Safe, Secure, and Compliant AI Adoption Approach Whitepaper Financial Services Cloud Council and Forum." Available: [Online]. https://community.ibm.com/community/user/fscc/viewdocument/generative-ai-controls-framework-sa? Community Key= 911102bb-642f-450c-alea-0187c94429ba&tab=library documents

[21] "Adopting generative Al in banking Corporates and Institutions." [Online]. Available: https://corporates.db.com/publications/White-papers-guides/adopting-generative-ai-in-banking?language_id=1

[22] "The rise of generative artificial intelligence in lending." Aug. 2024. [Online]. Available: https://www.finastra.com/viewpoints/white-paper/rise-generative-artificial-intelligence-lending

[23] "How to Train Generative AI Using Your Company's Data." [Online]. Available: https://hbr.org/2023/07/how-to-train-generative-ai-using-your-companys-data

[24] "Financial Services: scaling GenAI for maximum impact." [Online]. Available: https://www.deloitte.com/global/en/alliances/google/perspectives/financial-services-scaling-genai-for-maximum-impact. html

[25] "Generative AI in Banking and Finance I White Paper." [Online]. Available: https://www.cigniti.com/resource/white-papers/generative-ai-banking-financial-use-cases-benefits-ai-ml/

[26] "The Financial Stability Implications of Artificial Intelligence," Nov. 2024. [Online]. Available: https://www.fsb.org/2024/11/the-financial-stability-implications-of-artificial-intelligence/

[27] "Generative Al in Finance: Opening up a Sea of Possibilities." [Online]. Available: https://www.tcs.com/what-we-do/industries/banking/white-paper/generative-ai-finance-insurance-industry

[28] "GenAI in the Derivatives Market: a Future Perspective International Swaps and Derivatives Association [Online]. Available: https://www.isda.org/2024/04/17/genai-in-the-derivatives-market-a-future-perspective/

[29] "What is the World Economic Forum saying about artificial intelligence? Must-read research on the impact of genAI," Oct. 2024. [Online]. Available: https://www.weforum.org/stories/2024/10/generative-ai-impact-latest-research/

[30] "Transforming to an Al-powered McKinsey." [Online]. Finance function Available: //www.mckinsey.com/capabilities/mckinsey-digital/our-insights/what-an-ai-powered-finance-function-of-the-future-looks-like https:

[31] J. Lee, N. Stevens, S. C. Han, and M. Song, "A Survey of Large Language Models in Finance (FinLLMs)," Feb. 2024, arXiv: 2402.02315 [cs]. [Online]. Available: http://arxiv.org/abs/2402.02315

[32] C. Jeong, "Fine-tuning and Utilization Methods of Domain-specific LLMs," Journal of Intelligence and Information Systems, vol. 30, no. 1, pp. 93-120, Mar. 2024, arXiv:2401.02981 [cs]. [Online]. Available: http://arxiv.org/abs/2401.02981

[33] "Generative AI Controls Framework Safe, Secure, and Compliant AI Adoption Approach Whitepaper Financial Services Cloud Council and Forum." [Online]. Available: https://community.ibm.com/community/user/fscc/view document/generative-ai-controls-framework-sa? Communityl Key= 911102bb-642f-450c-alea-0187c94429ba&tab=librarydocuments

[34] "The Transformative Impact of Al and GenAI on Financial Services." [Online]. Available: https://www.pacificdataintegrators.com/blogs/finance-ai-generative-ai

[35] R. d. M. D. G.-M. T. H. L. L. A. D. M. F. B. Nguyen, Era Dabla-Norris, "Broadening the Gains from Generative AI: The Role of Fiscal Policies." [Online]. Available: https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2024/06/11/Broadening-the-Gains-from-Generative-AI-The-Role-of-Fiscal-Policies-549639

[36] "The revolution arrives: How gen Al is poised to transform banking." [Online]. Available: https://www.bai.org/banking-strategies/the-revolution-arrives-how-gen-ai-is-poised-to-transform-banking/

[37] O. Campesato, Python 3 and Machine Learning Using ChatGPT/GPT-4. Walter de Gruyter GmbH & Co KG, Jun. 2024, google-Books-ID: Mm8NEQAAQBAJ.

[38] "NLP in FinTech: Developing a Lightweight Text-to-Chart Application for Financial Analysis | IEEE Conference Publication | IEEE Xplore." [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10710766

[39] S. Jomthanachai, W. P. Wong, and K. W. Khaw, "An application of machine learning regression to feature selection: a study of logistics performance and economic attribute," Neural Computing and Applications, vol. 34, no. 18, pp. 15781-15805, Sep. 2022. [Online]. Available: https://doi.org/10.1007/s00521-022-07266-6

[40] S. Joshi, "Advancing Financial Risk Modeling: Vasicek Framework Enhanced by Agentic

Generative AI by Satyadhar Joshi," Advancing Financial Risk Modeling: Vasicek Framework Enhanced by Agentic Generative Al by Satyadhar Joshi, vol. Volume 7, no. Issue 1, January 2025, Jan. 2025.

[41] J. Satyadhar, "Enhancing structured finance risk models (Leland-Toft and Box-Cox) using GenAI (VAES GANs)," International Journal of Science and Research Archive, vol. 14, no. 1, pp. 1618-1630, 2025.

[42] M. Khoja, "AI and Bond Values: How Large Language Models Predict Default Signals," Rochester, NY, Sep. 2024. [Online]. Available: https://papers.ssrn.com/abstract=4965227

[43] D. P. Jeong, Z. C. Lipton, and P. Ravikumar, "LLM-Select: Feature Selection with Large Language Models," Jul. 2024, arXiv: 2407.02694 [cs]. [Online]. Available: http://arxiv.org/abs/2407.02694

[44] N. Nascimento, C. Tavares, P. Alencar, and D. Cowan, "GPT in Data Science: A Practical Exploration of Model Selection," in 2023 IEEE International Conference on Big Data (BigData), Dec. 2023,pp. 4325-4334. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10386503

[45] N. Li, C. Gao, Y. Li, and Q. Liao, "Large Language Model-Empowered Agents for Simulating Macroeconomic Activities," Rochester, NY, Oct. 2023. [Online]. Available: https://papers.ssrn.com/abstract=4606937

[46] S. Joshi, "Review of Data Engineering and Data Lakes for Implementing GenAI in Financial Risk A Comprehensive Review of Current Developments in GenAI Implementations," Rochester, NY, Jan. 2025.

[47] J. Satyadhar, "Review of Data Engineering Frameworks (Trino and Kubernetes) for Implementing Generative Al in Financial Risk," Inter-national Journal of Research Publication and Reviews, vol. 6, no. 2, pp. 1461-1470, Feb. 2025.

[48] S. Joshi, "Agentic Generative AI and the Future U.S. Workforce: Advancing Innovation and National Competitiveness," Rochester, NY, Feb. 2025.

[49] J. Satyadhar, "Generative AI: Mitigating Workforce and Economic Disruptions While Strategizing Policy Responses for Governments and Companies," Rochester, NY, Feb. 2025.

[50] Gupta, A. and Saleh, M.A. 2025. Generative AI: Initiative, Prospects, and Regulation with Reference to India. International Journal of Innovations in Science, Engineering And Management. 3, 2 (Feb. 2025), 454–458. DOI:https://doi.org/10.69968/ijisem.2025v3si2454-458.