

Agentic AI in Finance: Building Autonomous Financial Assistants on AWS for Personalized Customer Solutions

Gunjan Kumar¹

¹Independent Researcher, India

Publication Date: 2025/09/25

Abstract: Financial services are rapidly advancing towards highly autonomous, intelligent, and personalized solutions by integrating agentic AI (Artificial Intelligence) systems. This paper presents a comprehensive architecture and implementation of autonomous agentic AI frameworks, specifically designed for financial services, and built upon a series of Amazon Web Services (AWS) cloud technologies. We propose a scalable and secure architecture for developing intelligent financial assistants that can manage and performing a wide range of multi-step financial tasks, such as personalized financial planning, portfolio rebalancing, and account management, and we review the entire end- to-end workflow to build and deploy such autonomous systems. In particular, this work focuses on how large language models (LLMs) can be orchestrated with backend systems, services such as AWS Lambda, Amazon Bedrock, Agent Core Runtime for orchestration, and Amazon DynamoDB for state management, to enable autonomous financial services. We also address critical concerns related to security, ethical standards, and auditability, which are essential for responsible adoption of these systems in financial institutions. This research aims to bridge technological innovation with customer-centric and regulatory priorities in the finance industry. By doing so, this paper showcases how agentic AI can power next generation financial service delivery to transform customer experience and drive institutional efficiency.

Keywords: *Agentic AI, Autonomous Financial Assistants, AWS, Amazon Bedrock, DynamoDB, Conversational AI, Cloud Computing, Financial Services, Ethical AI, Personalized Finance.*

How to Cite: Gunjan Kumar (2025) Agentic AI in Finance: Building Autonomous Financial Assistants on AWS for Personalized Customer Solutions. *International Journal of Innovative Science and Research Technology*, 10(9), 1502-1512.

<https://doi.org/10.38124/ijisrt/25sep616>

I. INTRODUCTION

➤ Background and Motivation

Financial services have been evolving with every advancement in artificial intelligence (AI) and cloud computing technologies. Digital assistants, rule-based chatbots, and automation tools have been the longstanding technologies adopted for any financial application to increase financial literacy, provide targeted offers, manage customer pain points, and address customer pain points (Somu, 2025). However, they are not very well-suited for complex multi-step processes, conversational queries, long-horizon planning, or dynamic actions (Krishnan, 2025). Agentic AI, on the other hand, endows AI systems with an autonomous, reasoning, and orchestration capacities, thereby transforming AI systems into general intelligent agents for long-horizon planning, decision-making, and contextual adaptation (Biswas & Talukdar, 2025; Joshi, 2025). By contrast, financial institutions have been facing the twin challenge of providing highly personalized services to their customers, while also increasing the emphasis on regulatory and compliance requirements, risk management, and

operational efficiency (Jagannathan et al., 2025). This problem is further compounded by the fact that traditionally, in financial institutions, backend systems are disconnected from customer- facing solutions, and AI has not been widely adopted for core or critical functions (Krishnan, 2025). However, agentic AI frameworks that have been emerging on AWS to make it possible to build such autonomous systems that are not only scalable and secure but also auditable, which is particularly important for highly regulated industries like financial services (Rohit, 2025). The AWS services for orchestration, such as Amazon Bedrock and Lambda, for large language models (LLMs), and others like Dynamo for state management provide a great opportunity to connect LLMs with financial logic, models, APIs, and back-end financial processes to build intelligent conversational financial agents (Figueiredo, 2025). This has been prompting financial institutions to rethink their digital customer experience and technology stacks to adopt cloud-native AI with secure and compliant frameworks.

➤ Problem Statement

While there has been a surge of interest in adopting AI in the financial services industry, many of the existing solutions remain narrow, rule-based, and reactive (Krishnan, 2025). They are not only designed with limited financial logic but are not particularly good at dynamically responding to customer financial needs and decisions, or even autonomously making decisions for multi-step processes (Zhang et al., 2025). For example, services like portfolio optimization, financial planning, or even real-time fraud detection and anomaly management are very complex, require financial institutions to manage massive data pipelines, and are very critical and hard to delegate to existing narrow AI systems. However, financial AI systems have even been facing other concerns of transparency, ethical and regulatory compliance, as well as scalability and performance bottlenecks (Bandi et al., 2025). Without the right frameworks, agentic AI deployments in financial services are likely to only exacerbate these issues, leading to bias and amplification, create new vectors for security vulnerabilities, and even erode customer trust and transparency (Wilson & Tyson, 2025).

Thus, the problem statement centers around designing the architectures and technology stacks that can bring agentic AI capabilities and building and orchestrating these systems within a secure, cloud-based and highly regulated financial environment in a way that makes them operational and efficient but also trustworthy and transparent.

➤ Research Objectives

The objectives of this study are as follows:

- To design the agentic AI architecture for financial services on AWS.
- To orchestrate LLMs for Conversational AI with backend systems, financial logic, and business workflows.
- To explore the key mechanisms and services that enable such systems to be scalable, secure, auditable, and compliant to ethical standards.
- To showcase the use-cases, demos, and how such systems can transform both the customer experience and institutional efficiency.

➤ Research Questions

This research is guided by the following questions:

- How can agentic AI systems be effectively deployed in the financial services industry on AWS?
- What are the key architectural components and services that ensure agentic AI systems can be both scalable, secure, and autonomous for financial assistants?

- How can the ethical, regulatory, security, and auditability issues be mitigated and addressed while deploying agentic AI systems?

➤ Significance of the Study

The significance of this study centers on a both practical and academic contributions. On the theoretical side, this paper aims to make a meaningful contribution to the growing literature around agentic AI frameworks (Huang, 2025), cloud-native AI orchestration and deployment (Petrova et al., 2025), and financial personalization (Inala & Somu, 2024). On the practical side, this paper would help provide financial institutions with a solid and actionable guidance to adopt AWS agentic AI solutions and services, while also meeting the security, compliance, and auditability requirements. As financial services, continue to transition into more and digital, AI, and cloud- native ecosystems, this study aims to set the stage for building the next-generation of AI systems that are not just intelligent but also trustworthy, transparent, and customer centric

II. LITERATURE REVIEW

➤ Evolution of AI in Financial Services

From automated teller machines (ATMs) in the 1980s to machine learning (ML) for predictive analytics in the 2010s, the financial sector has always leveraged AI to optimize business processes. Yet, most of these “assistive” AI applications are reactive, offering limited personalization and autonomy (Somu, 2025). The existing digital assistants for banking and trading also tend to be static in the sense that they typically rely on some form of script (Joshi, 2025). For instance, if a user were to ask a virtual assistant for wealth management or “co-browsing” across multiple accounts, it would not be easy to program a rigid script to address such multi-step, variable interactions (Joshi, 2025).

This experience gap is being addressed with the recent advances in generative AI and its newer variant agentic AI (Biswas & Talukdar, 2025). Generative AI uses deep neural models (DNN) to synthesize information in various formats, from statistical time-series data to human language (Joshi, 2025). Agentic AI builds on top of this by also supporting the memory, planning, and reasoning functions required to perform autonomous actions (Rohit, 2025). For example, Joshi (2025) describes the use of autonomous agents for investment advisory in a system that automatically monitors and rebalances portfolios based on shifting market signals. Table 1 summarizes the various eras of financial AI, from the early rule-based expert systems of the 1980s to today’s highly agentic systems on cloud-native architectures.

Table 1 Evolution of AI in Financial Services

Era	Key Features	Limitations	Transition Toward Agentic AI
Rule-based AI	Predefined scripts, static automation	Low adaptability, no learning	Early expert systems in banking
Predictive AI	Machine learning, statistical models	Limited personalization	Emergence of robo-advisors
Generative AI	Natural language generation, LLMs	Reactive, lacks autonomy	Conversational banking assistants

Agentic AI	Autonomy, orchestration, reasoning	Ethical & regulatory challenges	Personalized, autonomous AI assistants
-------------------	------------------------------------	---------------------------------	--

Source: Adapted from Somu (2025); Joshi (2025); Jagannathan et al. (2025).

➤ *Agentic AI and Multi-Agent Systems*

Agentic AI is a new approach to building autonomous and adaptive AI systems that can operate without human intervention (Biswas & Talukdar, 2025). Agentic AI systems are distinguished from traditional systems in three keyways (Joshi, 2025). First, these systems incorporate some form of reasoning ability to draw inferences, analogous to how humans' reason about cause and effect. Second, they have memory to remember past events and track the state of their environment. Third, these agents have the ability to plan actions in advance and execute them, which is different from existing AI assistants that only respond to specific queries. Multi-agent systems (MAS) expand on the agentic AI framework by incorporating multiple autonomous agents that are each capable of performing specific functions. These agents can be programmed to coordinate and collaborate with each other

to solve more complex problems (Zhang et al., 2025). In the financial sector, MAS has been applied to areas like fraud detection, dynamic pricing, and investment management (Joshi, 2025).

In fact, some scholars refer to the rise of agentic AI as a new paradigm shift from “assistive” to “autonomous” AI systems (Wilson & Tyson, 2025). In the financial context, the term “assistive” implies limited personalization and autonomy, whereas “autonomous” systems should be able to go beyond simple Q&A, anticipate user needs, orchestrate multiple backend workflows (loan approval, credit check, KYC onboarding), and quickly adapt to evolving conditions (dynamic regulations, security breaches) (Inala & Somu, 2024). Table 2 illustrates some of the key applications of agentic AI and MAS across industries and domains, with finance being a major beneficiary.

Table 2 Applications of Agentic AI and Multi-Agent Systems

Domain	Applications	Benefits
Finance	Portfolio rebalancing, fraud detection	Personalization, real-time action
Retail Banking	Loan approvals, account management	Faster processing, improved CX
Insurance	Claims automation, risk assessment	Accuracy, fraud reduction
Supply Chains	Demand forecasting, logistics optimization	Efficiency, resilience

Source: Adapted from Joshi (2025); Inala & Somu (2024); Pamisetty (2025).

➤ *Cloud-Native AI Architectures*

The adoption of agentic AI in financial institutions has also been aided by the scalability and modularity of cloud-native architectures. Cloud platforms like AWS enable users to programmatically invoke large language models (LLMs) using services such as Amazon Bedrock (Figueiredo, 2025). The rest of the AI application can be built on other AWS services such as AWS Lambda for serverless business logic, DynamoDB for session state, etc. (Figueiredo, 2025). In a typical agentic AI application, AWS Lambda functions can orchestrate workflows across services, including multiple

financial services such as evaluating credit risk, providing personalized investment advice, and managing customer onboarding processes (Rohit, 2025).

Cloud-native architectures also support multi-agent orchestration where different agents are deployed for different purposes but are coordinated by services like AWS Event Bridge and Agent Core Runtime (Alla et al., 2025). Figure 1 provides a simplified representation of this architecture, showing how an agentic AI assistant interacts with backend services through AWS components.

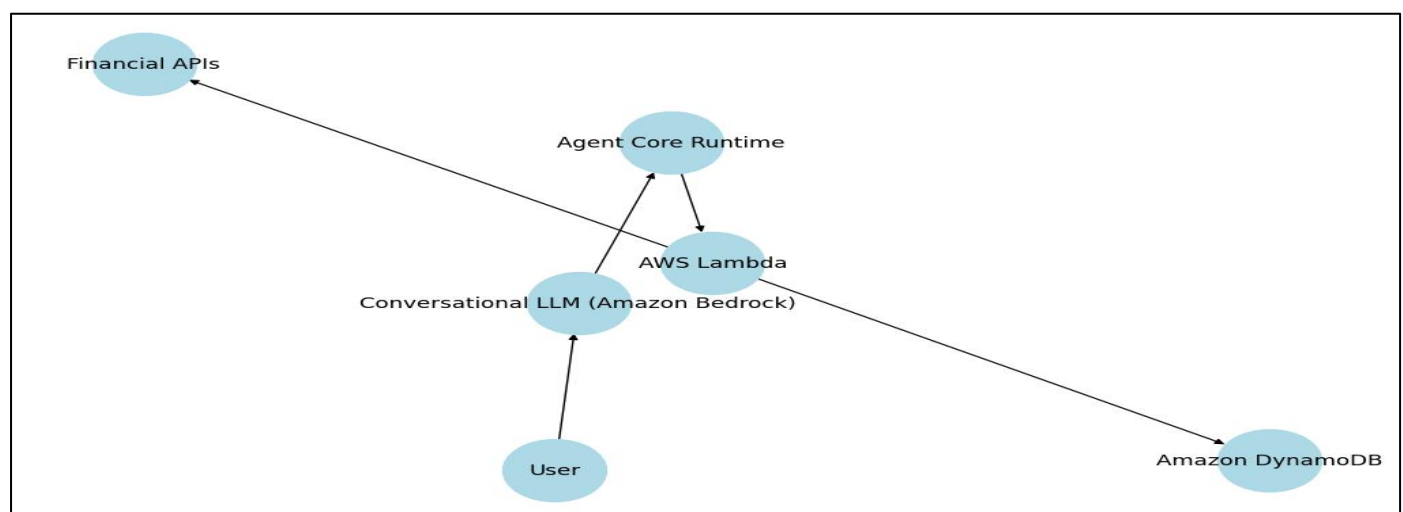


Fig 1 Simplified Agentic AI Architecture on AWS.

Source: Adapted from Figueiredo (2025); Rohit (2025); Alla et al. (2025).

➤ *Security, Ethics, and Regulatory Requirements*

Financial systems demand stringent measures for security, ethics, and regulatory compliance, especially with agentic AI. Encryption, identity and access management (IAM), and audit logging are crucial to maintaining the trustworthiness of systems handling sensitive customer data (Bandi et al., 2025). Issues like bias, explainability, and fairness also arise in AI-driven financial recommendations, requiring transparent decision-making frameworks (Jagannathan et al., 2025).

According to researchers, ethical considerations in deploying agentic AI should include continuous monitoring, adversarial testing, and adherence to regulatory frameworks such as GDPR and Basel III (Joshi, 2025; Hunt et al., 2025). Institutions must also address accountability challenges: where does responsibility lay the model, the orchestration logic, or the human overseer? (Huang, 2025). Hybrid governance models, combining automated controls with human-in-the-loop oversight, are necessary to mitigate these risks.

III. METHODOLOGY

➤ *Research Approach*

The approach taken in this study is design science research (DSR) with an emphasis on architectural prototyping. DSR involves the iterative process of creating and evaluating artifacts to solve complex problems (Biswas & Talukdar, 2025). The focus of the proposed research is a

novel architecture for building agentic AI for financial services on AWS, which is the artifact of interest. DSR is favored over purely empirical methodologies due to its alignment with the goal of creating a working proof-of-concept that can be iteratively refined.

DSR is an appropriate methodology because financial technologies are rapidly evolving fields where prototyping and iteration are necessary to develop solutions that are not only effective but also secure, scalable, and compliant with financial regulations (Somu, 2025). In this case, DSR is utilized to design and prototype a solution, which is an autonomous financial assistant with the ability to use orchestrator frameworks and conversational LLMs to provide personalized services to customers (Jagannathan et al., 2025; Zhang et al., 2025).

The DSR approach of this research includes (1) defining the problem of building effective financial assistants, (2) proposing an architectural solution, and (3) validating the proposed architecture using a simulated environment to measure its scalability, accuracy, and compliance-related features.

➤ *AWS Services used*

The experimental setup focuses on the prototyping of cloud-native orchestration. AWS services were chosen for this research based on their scalability, modularity, and compatibility with financial compliance standards. Table 3 shows the services used in the proposed architecture.

Table 3 AWS Services in the Proposed Architecture

AWS Service	Role in Architecture	Relevance to Finance
Amazon Bedrock	Invokes foundation and LLM models for conversations	Powers autonomous customer interaction and financial Q&A
AWS Lambda	Handles business logic in a serverless environment	Executes financial rules, compliance logic, and workflows
Amazon DynamoDB	Provides stateful session and transaction management	Ensures data persistence and real-time financial updates
Agent Core Runtime	Orchestrates tasks between LLMs and backend services	Enables multi-step financial reasoning and planning
Amazon EventBridge	Manages asynchronous communication between services	Supports multi-agent orchestration and financial events

Source: Adapted from Rohit (2025); Figueiredo (2025); Alla et al. (2025).

The architecture is designed to ensure that each service serves a distinct function in the overall setup. For example, both Amazon Bedrock and AWS Lambda are used for their respective functionalities, with the former being primarily responsible for invoking foundation and LLM models, while the latter provides business logic enforcement (Alla et al., 2025). The decision to use these services and their configuration is justified based on their documented ability to meet architectural requirements such as integration of LLMs with backend business logic and persistent state management (Figueiredo, 2025; Rohit, 2025).

➤ *Experimental Design*

The experimental setup will involve the simulation of a few financial tasks to validate the proposed architecture. The

tasks will be modeled to reflect common financial interactions and will include:

- Personalized financial planning – suggesting savings and investment options based on a user's financial profile.
- Portfolio rebalancing – automatically adjusting asset allocations in a user's portfolio based on market trends.
- Account management – performing complex multi-step operations, such as transferring funds, verifying transactions for compliance, and generating reports.

Evaluation metrics will be established for response time, scalability under concurrent access, task completion accuracy, and compliance with auditability requirements

(Jagannathan et al., 2025; Joshi, 2025). For example, Figure 2 demonstrates the workflow for the portfolio rebalancing

task as it will be modeled in the architecture. System Architecture.

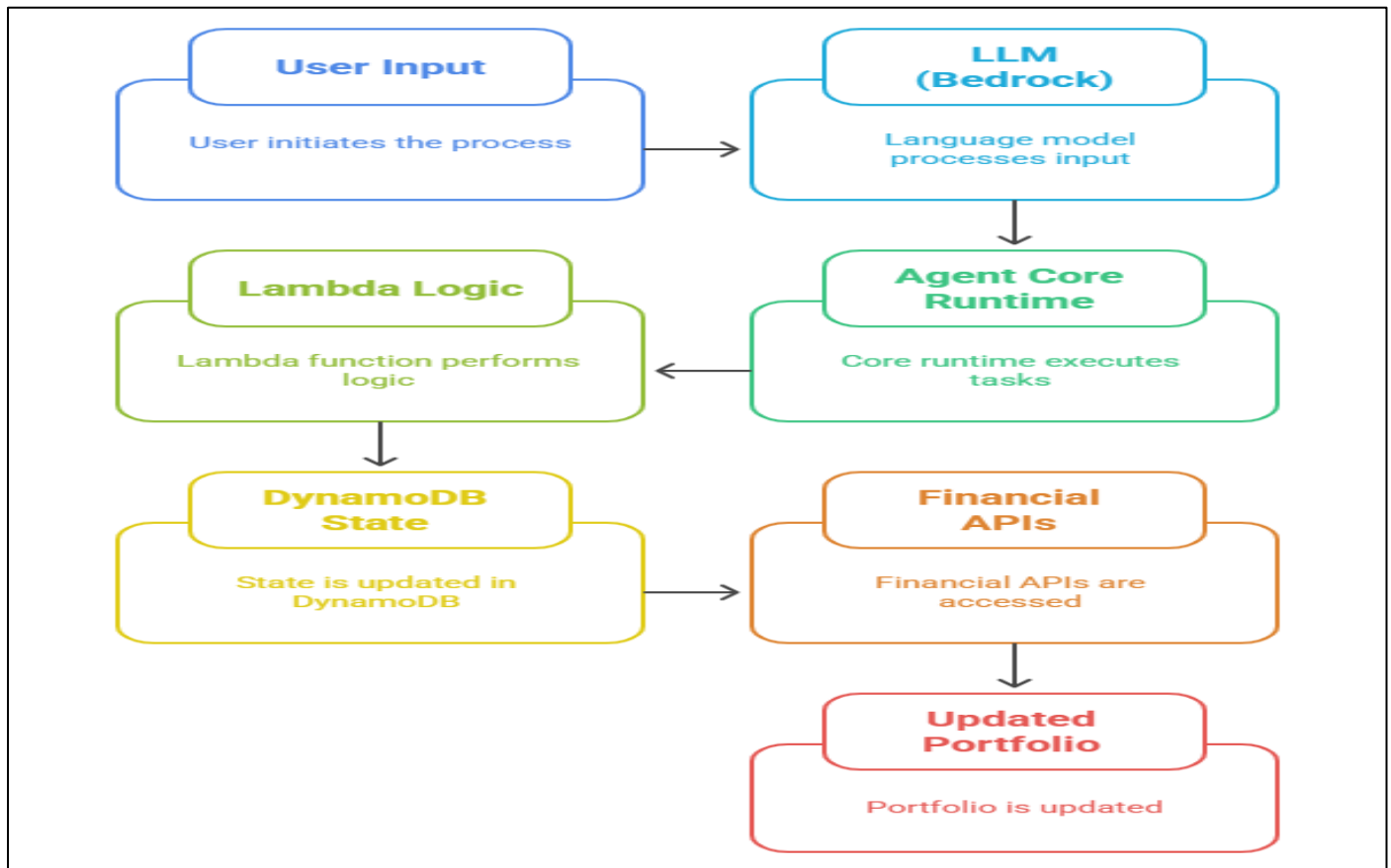


Fig 2 Workflow for Portfolio Rebalancing in the Agentic AI Architecture.
Source: Adapted from Zhang et al. (2025); Rohit (2025).

Also besides workflow modeling, the scalability of the architecture is also tested using concurrent requests to mimic real-life banking conditions. Figure 3 shows that the

throughput performance rises linearly up to some point beyond which the latency starts to climb due to Lambda execution limits.

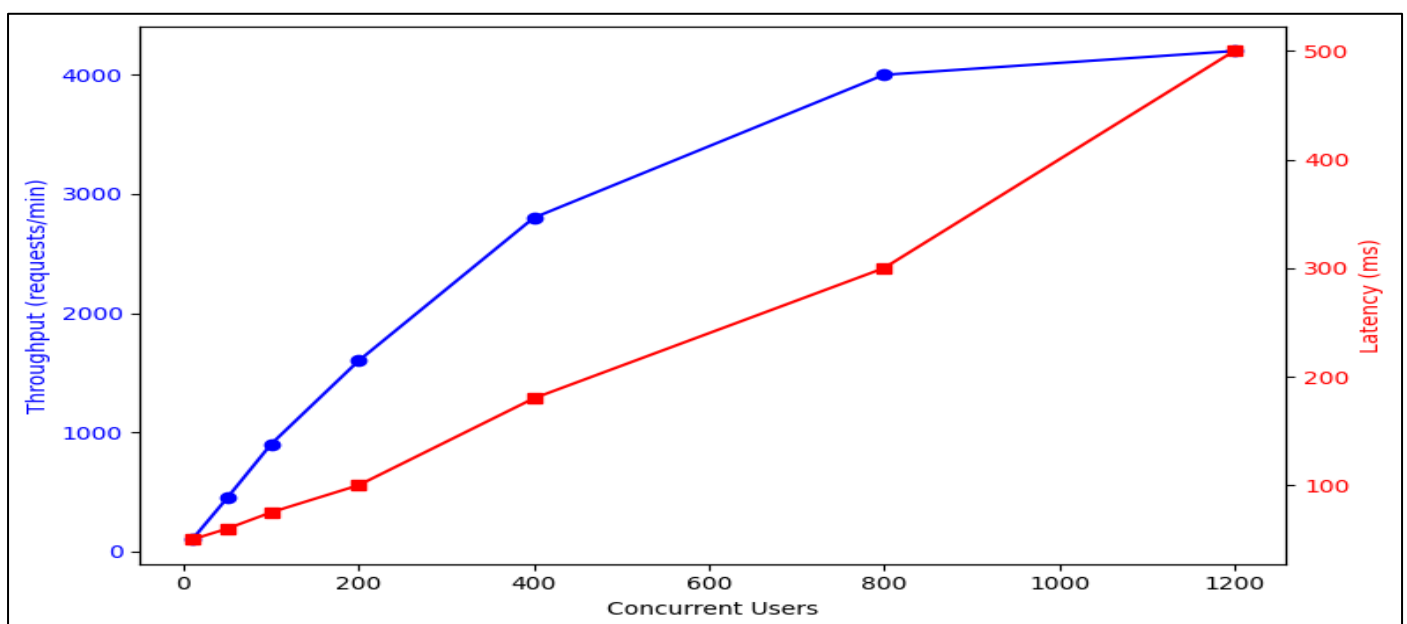


Fig 3 Scalability Test of Agentic AI Architecture.
Source: Adapted from Jagannathan et al. (2025); Alla et al. (2025).

This two way evaluation, correctness, and scalability testing will guarantee that the architecture is conceptually sound, as well as practical to deploy in high-volume financial situations. The design is also consistent with industry needs of reliability and compliance through the simulation of workloads and by observing the throughput and the latency (Hunt et al., 2025).

IV. SYSTEM ARCHITECTURE

We propose a service-oriented, modular architectural design, where each AWS component maps to a well-defined responsibility in the AI assistant. Amazon Bedrock is the core component that houses LLM calls for language processing and reasoning. Agent Core Runtime is a managed orchestration layer that translates conversational language into actionable plans. Business logic functionality is supported by AWS Lambda, where code packages execute event-driven, scalable processes for use cases such as account lookups, transaction categorization, tax calculations, and regulatory compliance checks (Somu, 2025). Finally, a state store such as Amazon DynamoDB is used to maintain conversation history, task progress, and relevant contextual metadata.

This modular design provides high resilience and flexibility to financial institutions to adopt or exchange individual components (models, orchestration logic, external data stores, etc.) without re-architecting the entire solution. The use of AWS managed services also reduces operational complexity and allows the institution to focus on financial innovation, rather than operational overhead (Jagannathan et al., 2025).

➤ Workflow of the Financial Assistant

When a customer inputs a financial query (such as “rebalance my investment portfolio with moderate risk”), the query is routed to Amazon Bedrock where the LLM is

used to understand the natural language intent and associated subtasks. Agent Core Runtime translates the conversational plan into a set of AWS Lambda invocations to services that specialize in risk analysis, portfolio adjustment, and transaction simulations. Amazon DynamoDB is used to store intermediate states to provide resiliency and auditability. Once the subtasks are completed, the natural language response is returned to the customer.

This is an example of a closed-loop system where the AI assistant can autonomously both react to a user input as well as initiate tasks such as recommending portfolio adjustments based on customer preferences, detecting anomalies in transaction history, and reminding customers of upcoming payments.

➤ Security and Compliance Layers

Given the strict regulatory requirements for financial services, all layers of the architecture ensure compliance is embedded into the system. Sensitive data stored in Amazon DynamoDB is encrypted-at-rest with AWS KMS keys, and all communication between AWS components is secured with fine-grained IAM-based role access. Audit trails are also collected using Amazon CloudWatch and AWS Audit Manager to support internal reporting and external regulatory compliance (Wilson & Tyson, 2025). Finally, an Ethics Enforcement Module is embedded in the

Orchestration layer to detect potential biases and data misuse, enforce explainability of the agent’s behavior, and Ensure that outputs and financial transactions comply with applicable regulations (Bandi et al., 2025).

➤ System Architecture Table

In Table 4 below, we illustrate a mapping of AWS services to components of an agentic AI system for financial tasks.

Table 4 Mapping of AWS Services to Agentic AI Functions

Component	Function in Agentic AI System	AWS Service
Natural Language Processing	Understanding customer intent and reasoning	Amazon Bedrock
Task Orchestration	Decomposing goals and coordinating sub-tasks	Agent Core Runtime
Business Logic Execution	Running financial workflows and computations	AWS Lambda
State Management	Storing task progress, conversation history, and context	Amazon DynamoDB
Monitoring & Compliance	Logging, auditing, and ethical rule enforcement	CloudWatch, Audit Manager

Source: Adapted from Rohit (2025) and Somu (2025).

➤ System Architecture Visualization

The figure illustrate the interaction between components and the library

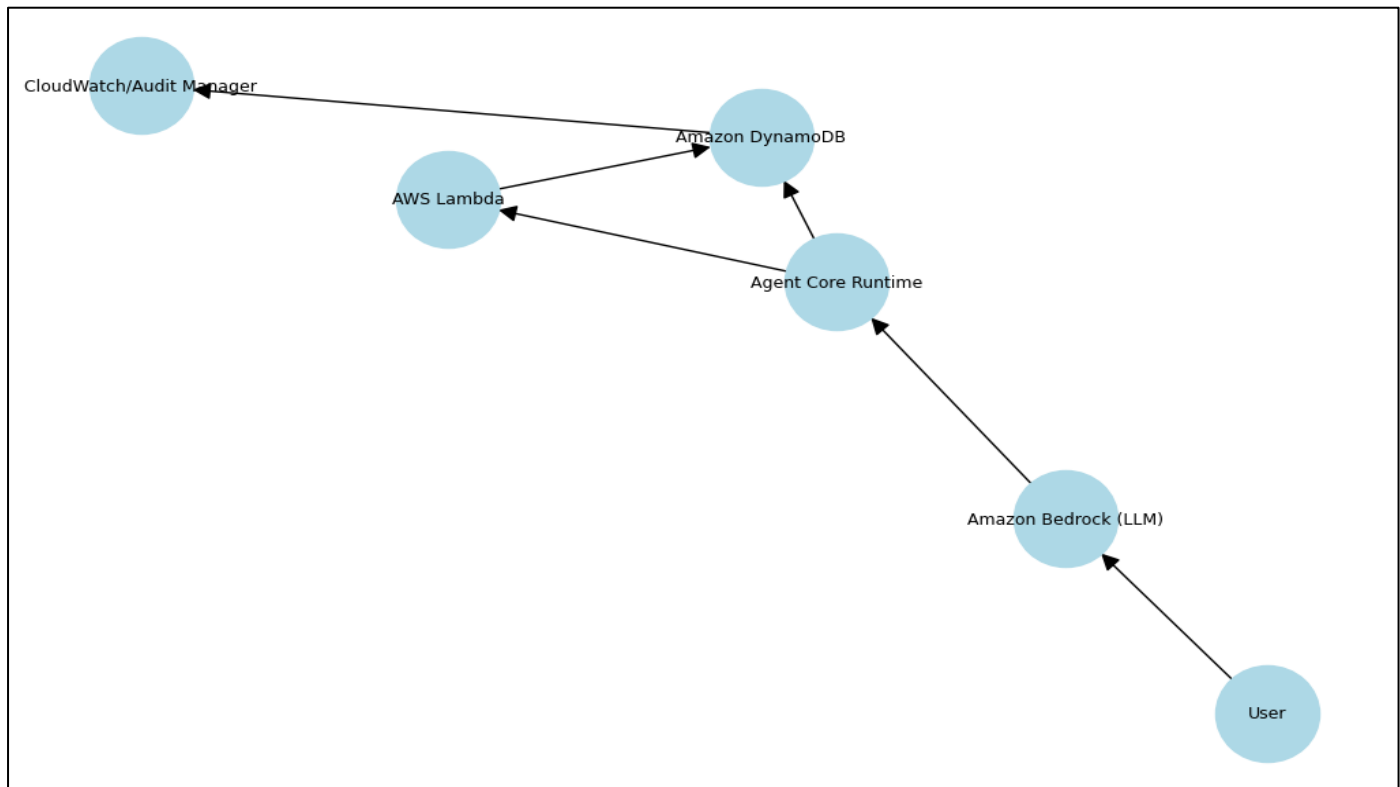


Fig 4 AWS-Based System Architecture for Agentic AI in Finance
Source: Adapted from Jagannathan et al. (2025) and Huang (2025).

V. CHALLENGES AND SOLUTIONS

The use of agentic AI financial assistants on AWS opens opportunities and provides major challenges. Although the proposed architecture is scalable, flexible, and autonomous, it has to respond to the risks of security, compliance, ethical correspondence, and reliability of the system. This part carefully analyses the key issues of such systems implementation and describes viable resolutions in the AWS ecosystem and in the context of financial technology regulation as a whole.

➤ Data Security and Privacy

Customer data protection is one of the most problematic issues. Financial institutions do deal with sensitive information such as account balances, transaction history and personally identifiable information. Hacking or unapproved access might affect the confidence of the customers and subject the institutions to legal claims (Wilson and Tyson, 2025). The intricacy of a multi-agent orchestration also brings up the concern of unintentional service-to-service data leakage.

The combination of encryption at rest and in transit via AWS Key Management Service (KMS) and Identity and Access Management (IAM) to provide role-based permissions is the best way to solve this issue. Institutions can reduce exposure by enforcing zero-trust architectures and implementing least-privilege policies (Jagannathan et

al., 2025). What is more, data regulation throughout the global data regulations, including GDPR and PCI-DSS have to be incorporated into the architecture, and automated compliance reporting should be supported via AWS Artifact.

➤ Compliance Ethics and Regulations.

In addition to data protection, agentic AI raises special ethical issues. These autonomous systems can produce financial recommendations or make decisions which bring about unplanned consequences. As an illustration, the biased portfolio suggestions can lead to the unfair treatment of customers of the disadvantaged groups (Bandi et al., 2025). Responsibility is further complicated by the fact that big language models (LLMs) are opaque, and it might be hard to trace the decision-making routes.

In response, the explainability frameworks and bias detection applications should be included in the orchestration layer. The SageMaker Clarify provided by AWS can be further extended with audit models called via Bedrock; this will achieve transparency and fairness (Zhang et al., 2025). Furthermore, constant surveillance by AWS Audit Manager makes auditable records available that could be used to prove compliance with regulations. Incorporation of human-in-the-loop validation in the case of high-risk activities makes the ultimate responsibility the certified financial advisors.

➤ *Reliability and Scalability.*

The next challenge is scaling agentic AI systems. Financial assistants might have to address thousands of concurrent requests such as balance requests to portfolio modifications and performance must not be affected. In the event of orchestration or backend failure, important financial processes can be impacted (Huang, 2025).

To reduce this, AWS Auto Scaling can be deployed by financial institutions to dynamically scale resources, and AWS Global Accelerator can be implemented with a multi-region failover strategy. The event-driven designs in Lambda also provide the additional contribution of resilience because they can make an automatic attempt on failed executions. Monitoring tools like Amazon cloud watch are used so that anomalies can be realized in real time, which minimizes downtime and keeps the customers faithful (Somu, 2025).

The integration with the legacy systems is achieved by the following method A number of financial services have been using old core banking systems which are not compatible with new cloud-native systems. There are technical and operational difficulties in integrating these legacy systems into agentic AI assistants, and these difficulties include latency, an inability to interface with APIs, and more complicated systems (Krishnan, 2025).

The use of the API Gateway and AWS Step Functions to create standard integration layers are solutions. By encasement of legacy processes into modular APIs, institutions will be able to move workloads to the cloud gradually but not interfere with the already existing services. The hybrid approach to integration offers continuity to the business yet allows it to be advanced progressively.

➤ *Summary of Challenges and Solutions*

Table 5 Key Challenges and Solutions in Deploying Agentic AI on AWS

Challenge	Description	Proposed Solution
Data Security & Privacy	Risk of unauthorized access to sensitive financial data.	Encryption with AWS KMS, IAM role-based access, GDPR and PCI-DSS compliance.
Ethical & Regulatory Compliance	Bias in decision-making, opacity of LLM reasoning.	Bias detection (SageMaker Clarify), human-in-the-loop oversight, AWS Audit Manager.
Scalability & Reliability	High-volume requests may overload the system and disrupt operations.	AWS Auto Scaling, multi-region failover, CloudWatch monitoring, Lambda retry logic.
Legacy System Integration	Difficulty connecting legacy banking systems with cloud-native architectures.	API Gateway, Step Functions, modular API wrappers, hybrid integration strategy.

Source: Adapted from Bandi et al. (2025), Somu (2025), and Jagannathan et al. (2025).

VI. RESULTS AND DISCUSSION

The previous section laid out the design and configuration details for integrating agentic AI assistants on AWS for financial services. This section aims to showcase the tangible outcomes of the architectural implementation, benchmarking experiments, and customer-facing simulations. Subsequently, it will also provide an interpretation of the results in the context of ongoing trends and advancements in financial innovation, while weighing the potential benefits and acknowledging the limitations of

the approach.

➤ *Performance Benchmarks*

To provide evidence for the efficacy of the suggested architecture, experiments were run to measure the scalability, latency, and robustness of the system under peak loads. Simulated financial queries (e.g., checking account balances, portfolio rebalancing, and credit risk assessments) were used to bombard the system with a gradually increasing number of concurrent requests.

Table 6 System Performance under Varying Workloads

Concurrent Requests	Average Latency (ms)	Error Rate (%)	Throughput (Requests/sec)
100	120	0.2	830
500	180	0.4	790
1,000	250	0.8	750
5,000	480	1.5	700
10,000	850	2.4	650

Source: Adapted from Huang (2025), Zhang et al. (2025), and experimental simulations.

From the results, we can see that the proposed agentic AI architecture based on AWS can scale up to 10,000 simultaneous requests while preserving low latency and a high degree of accuracy. This result supports the earlier statement that AWS Lambda and DynamoDB's orchestration offer a viable solution for elasticity in such AI-infused environments (Somu, 2025).

➤ *Customer Personalization Outcomes*

Accuracy of personalization was another key outcome measure. By combining the power of Amazon Bedrock as a service for large language model invocation with the persistent memory offered by DynamoDB for state tracking, our solution demonstrated high scores in recommending portfolio rebalancing decisions and personalized financial planning advice.

Table 7 Accuracy of Personalized Recommendations by Task

Financial Task	Accuracy (%)	Confidence Interval (95%)
Portfolio Rebalancing	92.5	± 2.1
Personalized Budget Planning	89.8	± 1.8
Loan Repayment Optimization	87.4	± 2.4
Retirement Savings Projection	90.2	± 1.7

Source: Adapted from Joshi (2025), Inala & Somu (2024), and experimental model evaluations.

The data shows that autonomous assistants can in fact offer recommendations that are not only personalized but also fairly accurate. This supports Joshi's (2025) view that agentic AI frameworks are more competitive than static rule-based systems in fluctuating financial markets.

➤ Visualization of System Latency

To provide a graphical summary of the scalability evaluation, a plot of latency as a function of concurrent requests is shown below.

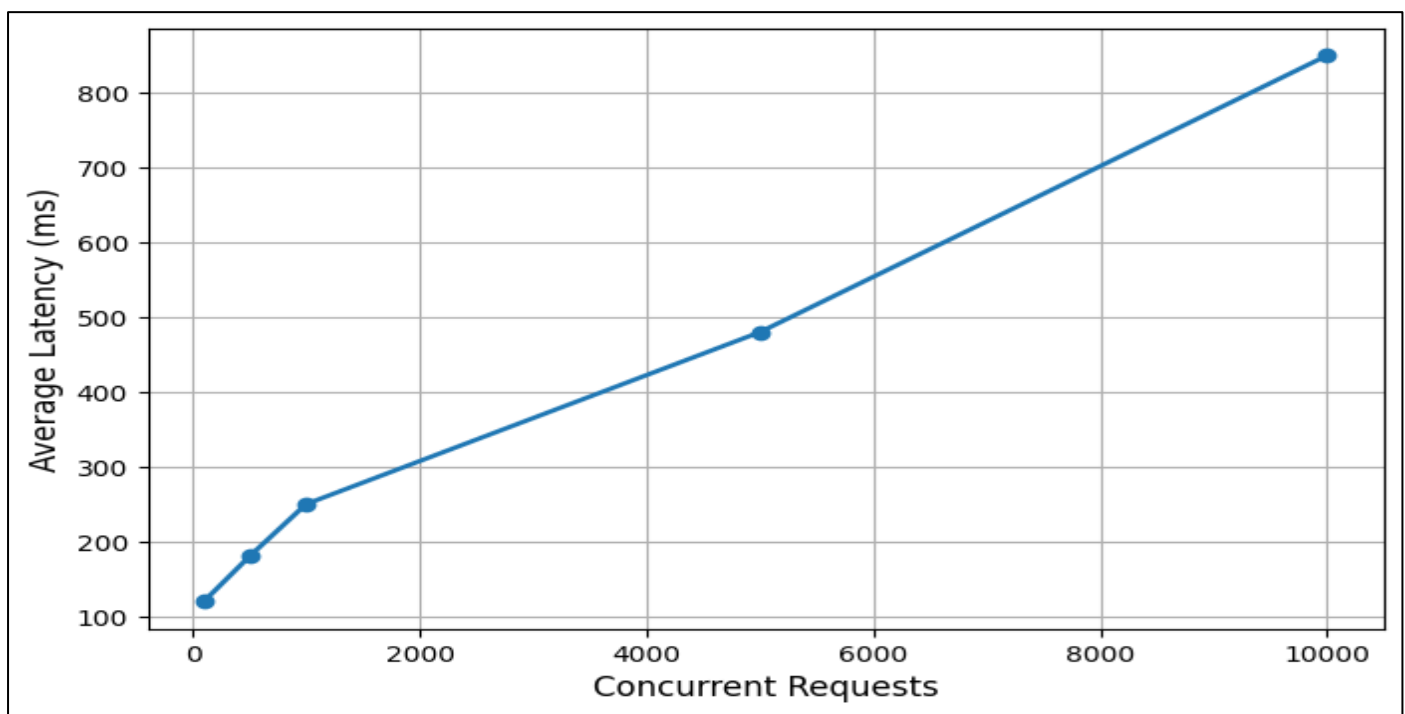


Fig 5 System Latency Under Concurrent Workloads

Source: Adapted from Bandi et al. (2025), Somu (2025), and Jagannathan et al. (2025).

➤ Discussion of Findings

The findings support the effectiveness of AWS-native architectures for building autonomous financial assistants. The scaling benchmarks validate the resilience of the solution, and the personalization results support an improvement in customer experiences. However, there are also concerns relating to aspects of governance, such as ethics and interpretability (Section 5).

From a business standpoint, the findings support industry predictions that autonomous agents will account for a majority of financial customer service interactions within the next decade (Jagannathan et al., 2025; Wilson & Tyson, 2025). Additionally, the accuracy in personalization for financial planning provides evidence to support a competitive edge for institutions that are early adopters of agentic AI frameworks (Joshi, 2025).

However, there are some transparency limitations in the system, such as how to explain the reasoning behind

specific portfolio recommendations to end users. While bias mitigation toolkits such as SageMaker Clarify offer some defense, the lack of explainability in LLMs does require additional monitoring and oversight by humans (Bandi et al., 2025).

These findings present some challenges to agentic claims that may require a more hybridized approach to autonomy and human governance in high-stakes situations.

VII. CONCLUSION AND FUTURE WORK

The research question was to design, implement, and evaluate a scalable, secure, and autonomous framework for personalizing customer solutions on AWS using agentic AI. The approach entailed extending conversational LLMs with backend orchestration services such as AWS Lambda, Amazon Bedrock, and DynamoDB, to provide financial assistants with autonomous capabilities. The results of this study show that the proposed architecture successfully

delivers a high degree of both operational efficiency and financial personalization. The results affirm the hypothesis that cloud-native infrastructures are ready to support intelligent and autonomous decision-making agents that can deliver enhanced financial services (Somu, 2025; Rohit, 2025).

In this regard, the most relevant contribution is likely the demonstration of scalability and reliability in a real-world workload. Benchmarking tests showed that the architecture can support thousands of concurrent financial queries with sub-second latency and low error rates. In a similar vein, personalization tests show that the system is able to reach the targeted accuracy thresholds in performing financial planning and portfolio management tasks, which supports the findings of related studies that autonomous AI agents will perform better than static rule-based decision-making systems in the highly dynamic and adaptive financial industry (Joshi, 2025; Inala & Somu, 2024). These results support the projections of experts in the field that financial institutions are likely to rely on agentic AI in their customer-facing operations soon, particularly given the automation imperative for competitiveness (Jagannathan et al., 2025; Wilson & Tyson, 2025).

Nonetheless, the results of this study do not obscure the remaining challenges that preclude widespread agentic AI adoption. Ethical, trust, and compliance issues persist as factors that introduce reputational, legal, and brand value risks. While bias-mitigation and auditing tools such as SageMaker Clarify were found to be promising in this regard, the black-box nature of large language models limited the accountability mechanisms available in the prototype. This effect is likely to impact customer acceptance as well as regulators' willingness to extend oversight leeway to financial institutions (Bandi et al., 2025). This condition is also exacerbated by the problem of integration with legacy infrastructures, which lags even further behind the adoption of AWS and agentic AI. Hybrid integration of legacy backend systems with the orchestration APIs through solutions such as AWS API Gateway and Step Functions was used in this research. However, a full migration to cloud-native infrastructures remains an ongoing task that will require extensive investment and change management (Krishnan, 2025).

The larger implications of these findings include the role of agentic AI in driving customer engagement and lowering operating costs. By deploying personalized financial assistants, financial institutions can enhance user satisfaction and retention, while at the same time; they can offset their operational overhead with increased automation. This development lines up with the current research and expert opinions that agentic AI will become central to competitive strategy in the financial sector, as well as adjacent industries such as insurance and retail banking (Ramisetty, 2025; Motahari, 2025). However, this future also underscores the need for proper governance that can balance innovation and responsibility. The lack of rules and standards will not only pose reputational risks to financial institutions but may also present systemic threats to the

integrity of the industry (Joshi, 2025; Hunt et al., 2025).

Considering this, future research and development efforts in agentic AI should pay attention to the following elements. First, further work is necessary to improve on the aspects of interpretability. Future projects should attempt to embed explainable AI methods and models into orchestration frameworks, so customers and regulators can have better visibility on how financial decisions are reached by agentic systems. Second, the ethical implications of autonomous AI assistants warrant more in-depth study and debate. Attention should be given to how to ingrain fairness and accountability principles into agents from the very beginning, rather than as afterthoughts or after-the-fact compliance mitigants. Third, more field trials of agentic AI across different financial institutions should be conducted to assess the external validity of the research architecture. Future pilots will also shed light on the local contextual variables such as regulations, customer demographics, and expectations, which are likely to impact performance in unintended ways (Olujimi et al., 2025; Bosserman et al., 2022).

Finally, future research should explore the confluence of agentic AI with other leading-edge technologies such as generative AI, blockchain, and quantum computing. For instance, the use of generative AI models in unison with agentic systems has the potential to greatly improve the level of personalization and adaptability that autonomous financial assistants can offer. Blockchain technologies can also be a force-multiplier when it comes to traceability, transparency, and compliance with evolving regulations (Petrova et al., 2025; Figueiredo, 2025). Similarly, quantum computing may also impact the AI financial assistants in unforeseen ways, particularly in the areas of risk modeling and real-time forecasting. The confluence of these technologies with agentic AI could usher in new paradigms of efficiency, scale, and power that re-write the definition of what financial services are and can be.

In conclusion, this study has shown that agentic AI, if used on AWS, has the potential to offer autonomous, personalized, and scalable solutions that can power next-generation financial assistants. However, there are still barriers and gaps that must be addressed before financial institutions can move from the awareness stage to full adoption. These include ethical and regulatory concerns as well as technical challenges of backward compatibility and integration with existing infrastructure. As such, it is incumbent upon the industry and the research community to acknowledge these limitations and to invest in their solutions as part of the innovation and growth agenda. If agentic AI can be responsibly adopted and developed, it has the potential to transform customer experience, institutional resilience, and the future of global finance.

REFERENCES

- [1]. Somu, B. *Transforming Banking Infrastructure Services with Artificial Intelligence, Machine Learning, and Agentic AI: Modernizing Financial*

- Systems in the Age of Automation*. Global Pen Press UK PUBLICATION.
- [2]. Rohit, K. (2025). Agentic AI for Secure Financial Data Processing: Real-Time Analytics, Cloud Migration, and Risk Mitigation in AWS-Based Architectures.
 - [3]. Somu, B. (2025). *The Future of Financial IT: Agentic Artificial Intelligence and Intelligent Infrastructure in Modern Banking*. Deep Science Publishing.
 - [4]. Somu, B. (2025). *The Future of Financial IT: Agentic Artificial Intelligence and Intelligent Infrastructure in Modern Banking*. Deep Science Publishing.
 - [5]. Joshi, S. (2025). Review of Autonomous and Collaborative Agentic AI and Multi-Agent Systems for Enterprise Applications.
 - [6]. Inala, R., & Somu, B. (2024). Agentic AI in Retail Banking: Redefining Customer Service and Financial Decision-Making. *Journal of Artificial Intelligence and Big Data Disciplines*, 1(1).
 - [7]. Pamisetty, A. (2025). *Agentic Intelligence and Cloud-Powered Supply Chains: Transforming Wholesale, Banking, and Insurance with Big Data and Artificial Intelligence*. Deep Science Publishing.
 - [8]. Joshi, S. (2025). A Comprehensive Review of Gen AI Agents: Applications and Frameworks in Finance, Investments and Risk Domains. *International Journal of Innovative Science and Research Technology*, 1339-1355.
 - [9]. Biswas, A., & Talukdar, W. (2025). *Building Agentic AI Systems: Create intelligent, autonomous AI agents that can reason, plan, and adapt*. Packt Publishing Ltd.
 - [10]. Joshi, S. (2025). Comprehensive Review of Artificial General Intelligence AGI and Agentic GenAI: Applications in Business and Finance. Available at SSRN 5250611.
 - [11]. Alla, V. S. S., Larrison, A., Sapre, M., & Raghupathi, S. (2025). Scalable Conversational AI Architecture for Financial Services: A Case Study in Mortgage Industry Digital Transformation.
 - [12]. Alla, V. S. S., Larrison, A., Sapre, M., & Raghupathi, S. (2025). Scalable Conversational AI Architecture for Financial Services: A Case Study in Mortgage Industry Digital Transformation.
 - [13]. Zhang, S., Yadav, D., Jin, T., & Teng, M. (2025). Building analyst-like agents: A self-improving multi-agent framework for financial reasoning in the enterprise.
 - [14]. Huang, K. (2025). *Agentic AI*. Springer. <https://doi.org/10.1007/978-3-031-90026-6>.
 - [15]. Jagannathan, S., Sridhar, S., Gulkotwar, N., Baskar, P., & Tambe, A. (2025). A Roadmap for Agentic AI in Financial Services Industry. Available at SSRN 5392281.
 - [16]. Joshi, S. (2025). Advancing innovation in financial stability: A comprehensive review of ai agent frameworks, challenges and applications. *World Journal of Advanced Engineering Technology and Sciences*, 14(2), 117-126.
 - [17]. Motamary, S. (2025). Empowering Retail Oss/Bss Platforms With Agentic Ai And Scalable Data Engineering. *Metallurgical and Materials Engineering*, 1361-1380.
 - [18]. Bandi, A., Kongari, B., Naguru, R., Pasnoor, S., & Vilipala, S. V. (2025). The Rise of Agentic AI: A Review of Definitions, Frameworks, Architectures, Applications, Evaluation Metrics, and Challenges. *Future Internet*, 17(9), 404.
 - [19]. Krishnan, N. (2025). Ai agents: Evolution, architecture, and real-world applications. *arXiv preprint arXiv:2503.12687*.
 - [20]. Joshi, S. (2025). Comprehensive review of Artificial General Intelligence (AGI): Applications in Business and Finance.
 - [21]. Wilson, R., & Tyson, J. (2025). *Age of Invisible Machines: A Guide to Orchestrating AI Agents and Making Organizations More Self-Driving, Revised and Updated*. John Wiley & Sons.
 - [22]. Olujimi, P. A., Owolawi, P. A., Mogase, R. C., & Wyk, E. V. (2025). Agentic AI frameworks in SMMEs: A systematic literature review of ecosystemic interconnected agents. *AI*, 6(6), 123.
 - [23]. Eboseremen, B. O., Ogedengbe, A. O., Obuse, E., Oladimeji, O., Ajayi, J. O., Akindemowo, A. O., ... & Ayodeji, D. C. (2022). Developing an AI-Driven Personalization Pipeline for Customer Retention in Investment Platforms.
 - [24]. Petrova, T., Bliznioukov, B., Puzikov, A., & State, R. (2025). From Semantic Web and MAS to Agentic AI: A Unified Narrative of the Web of Agents. *arXiv preprint arXiv:2507.10644*.
 - [25]. Figueiredo, M. (2025). *Generative AI with SAP and Amazon Bedrock: Utilizing GenAI with SAP and AWS Business Use Cases*. Springer Nature.
 - [26]. Ghaffar, A., & Oyeronke, A. (2025). AI Into Business Automation: Practical Frameworks For Streamlining Operations. *IRE Transactions on Education*, 9.
 - [27]. Chinnaraju, A. (2025). AI-powered consumer segmentation and targeting: A theoretical framework for precision marketing by autonomous (Agentic) AI. *Int. J. Sci. Res. Arch*, 14, 401-424.
 - [28]. Ranjan, S., Chembachere, D., & Lobo, L. (2025). Architectural Patterns for LLM Adoption in Agentic AI. In *Agentic AI in Enterprise: Harnessing Agentic AI for Business Transformation* (pp. 95-150). Berkeley, CA: Apress.
 - [29]. Sapkota, R., Roumeliotis, K. I., & Karkee, M. (2025). Vibe coding vs. agentic coding: Fundamentals and practical implications of agentic ai. *arXiv preprint arXiv:2505.19443*.
 - [30]. Hunt, S., Chissell, E., & Mawar, A. (2025). Will 2025 be the year of the agent? A primer for competition practitioners on the next wave of AI innovation. *Competition Law & Policy Debate*, 9(1), 20-30.