



Fiscally Intelligent Financial Crime Defense

1st Jai kiran reddy burugula , 2nd David Schneider

Jai Kiran Reddy Burugulla, Senior Engineer , jaikirrann@gmail.com

David Schneider , Research Assistant

Abstract

Financial crime detection systems are becoming increasingly common across the financial industry, yet their infrastructure and operational costs remain a significant concern. Applying Financial Operations (FinOps) principles to AI-driven, cloud-native systems can help govern resource usage and keep costs within allocated budgets. This paper proposes a FinOps-aware AI application framework for detection, monitoring, and forensic systems, supporting both detection pipelines for predefined tasks and continuous monitoring and forensic models for real-time scoring. Resource allocation and auto-scaling policies enforce cost governance by provisioning additional resources during peak loads, when anomalies or fraud are suspected, and during continuous monitoring or batch processing runs. Cost-aware strategies are defined for both model training and inference, while monitoring and alerting for underlying cloud resources, FinOps dashboards for decision-makers, and patterns for automated FinOps-aware deployment of AI services are integrated directly into the architecture.

Open-source, cloud-native software components enable the framework's implementation and testing across financial institutional, regulatory, and private cloud environments. The full process is mainstreamed and validated within regulatory sandboxes, supporting authorities that require such capabilities but lack the resources to build them independently. The proposed FinOps-aware framework aims to optimize the cost of AI operations across various financial domains, including cross-border transaction supervision and anti-money laundering (AML).

keywords: FinOps, Cloud-native architecture, Artificial intelligence in finance, Financial crime detection, Anti-money laundering (AML), Regulatory technology (RegTech), Cost optimization, Compliance automation, Distributed machine learning, Real-time fraud detection, Scalable AI systems, Big data analytics, Microservices architecture, Risk and compliance management, Explainable AI (XAI).

1. Introduction

Increasingly complex financial crime detection requires scalable architectures and technologies that can provide a constant detection capability and adapt to new types of attacks. The operational cost of fake news detection systems is mainly associated with the training and inference of the deep learning models. As costs soar, detection and compliance operations must be controlled, optimized, and examined to ensure accountability and upholding the public good. The proposed conceptual framework explores FinOps principles for operations undertaken in cloud environments by AI-driven systems that target financial crime at scale. Such systems must be designed and operated with these principles in mind, ensuring that cost



considerations are part of the decisions involved in a model's lifecycle, from data ingestion and storage to monitoring and alerting.

The proposed FinOps-Aware Cloud-Native AI conceptual framework stresses architectural and strategic choices that induce positive cost-performance trade-offs for AI-based financial crime detection and associated compliance operations. The need for efficient operations strongly relates to the inherent difficulties of supervised learning. AI models trained for the detection of money-laundering activities for instance struggle to generalize to unseen types of laundering behavior. And newly detected such an increased operational cost becomes prohibitive, as the situation in the institutionalized cross-border payment such as SWIFT shows. Operations must be optimized, and both FinOps principles and architectural choices matter, as a strategy that affects the system's overall cost, not just the cost of one particular stage or function, must be designed. Cloud usage costs inherent in training and inference typically represent the majority of the operating expense, and developed cloud-ready detection processes contribute to associated operational cost cutting.

1.1. Background and Significance

Financial criminals' explosive demand for capital, along with their continuing technical and financial resource advantages, allows them to attack the financial ecosystem with bold, inventive, and diverse targeting strategies. Financial institutions are forced to battle back with constant product and infrastructure de-risking against increasingly non-linear and complex threats. The financial industry spends tens of billions of dollars a year on financial crime risk mitigation to identify, analyze, and report money laundering and terrorist financing activity. While these investments have expanded capacity, the investment-to-impact ratio continues to decline. Consequently, financial crime detection and compliance support systems require innovative, scalable cloud-native architectures, novel FinOps-aware designs, governance practices aligned to cost targets, and technology partnerships with cloud-native AI system vendors and cyber technology firms.

Increasingly available developments with the use of real-time scoring together with a batch processing pipeline that analyzes transaction and interaction data at a certain frequency enable organizations to expand operational capacity with cost-optimized architectures and FinOps governance. As the banking and insurance industries continue to migrate their activities and services onto the public cloud to support hybrid selling and distribution strategies for third-party product offerings, the use of machine learning model-powered cloud-native AI systems for operational readiness demands the timely execution of detection model training, retraining, and performance monitoring. Data privacy, governance, security, and compliance requirements are always present in cloud-native AI development. However, the absence of cost optimization impacts the sustainability of technology investments in function-specific products and services that need FinOps capabilities embedded directly into the model recommendations.

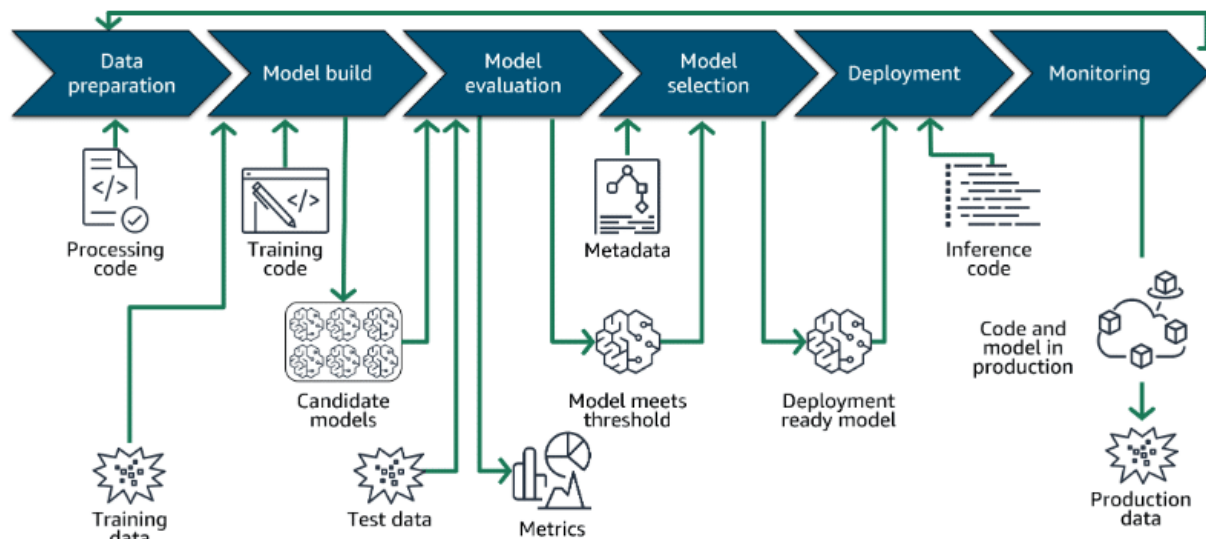


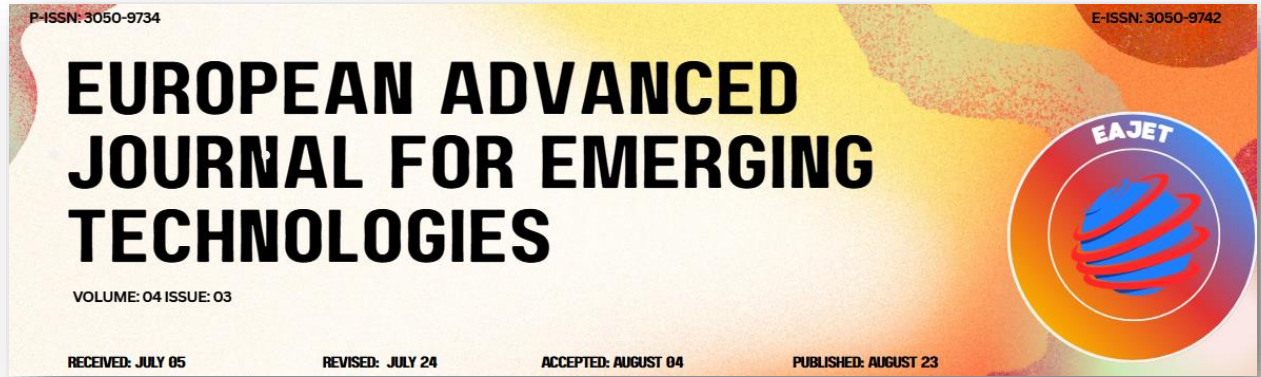
Fig 1: FinOps for AI Overview

1.2. Research design

A FinOps-oriented approach is proposed to the design, implementation, deployment, and maintenance of cloud-native AI systems for scalable financial crime detection and regulatory compliance operations. The interest is in real-time scoring and batch processing pipelines exposed by a FinOps-aware architecture and operating in balance with the costs of cloud resourcing and auxiliary onboarding model training and periodic retraining processes. Cost aspects of information security and privacy–resilience are addressed alongside regulatory compliance assurance, as exposed by externalities usual in the FinTech and CryptoTech domains. The discussion extends to the mapping of the FinOps principles to the context of AI-centric cloud-native infrastructures dominated by resource costs, to the FinOps-oriented deployment of the two families of detection models, and to the smart assignment of resources for prompt model scoring—both real-time and batch.

Identification of fake identity documents, credit card fraud, or fraudulent transactions; detection of money laundering, terrorist financing, or other uses of financial services for illicit purposes—these operations typically require a considerable cloud computing and storage capacity and burstable loads. The hosting institution usually shares these costs among clients. However, the business model is under strain. Market competition, the association of sanctions with client accounts or transactions, and the hence-forward-feared tarnishing of institution reputation urge LiSAs and CryptoExCHs to abide by regulations, even in their absence: the prediction of future FinTech–regulatory supervisory structures behind the new business models supported by the future deployment of cross-border digital currencies pegged in a currency basket; and the abandoned high volume of illicit funds flow through private and anonymous networks, attracting threaten-analysis and anti-maffia groups interest.

Equation 1: Total Cloud AI Operating Cost



$$C_{total} = C_{ingest} + C_{storage} + C_{train} + C_{infer} + C_{monitor} + C_{audit}$$

Step by step:

$$\begin{aligned} C_{ingest} &= r_i \times V_i \\ C_{storage} &= r_s \times S \times T \\ C_{train} &= r_g \times G \times H \\ C_{infer} &= r_c \times N \times t \end{aligned}$$

Substitute:

$$C_{total} = r_i V_i + r_s S T + r_g G H + r_c N t + C_{monitor} + C_{audit}$$

2. Background and Motivation

Financial crime detection systems typically process scoring requests in real time while handling model training, retraining, and batch processing of queries in batch mode. Composite systems consist of multiple pipelines working in parallel: one or several pipelines receive inbound data and produce anomaly labels or other predictions that invoke actions, while others consume out-of-band data and create training data for periodic training of models in the first pipelines. Processing latency is a crucial quality attribute for real-time inference pipelines; near-zero latency is feasible for distributed models with sufficient parallelism. Nonetheless, continuous high loads and low processing latency may lead to significant cost overruns, so it is advisable to apply FinOps principles for effective cost governance.

Privacy and data protection are critical for cloud computing, especially for the financial sector, so ensuring that personally identifiable information is not leaked, even by accident, is paramount. For supervised models that classify or predict sensitive properties, secure-to-process policies select global features such that no attribute of a private entity enters its model and infringe its privacy. Compliance with regulations from banking supervisory authorities must also be considered by cloud-native financial crime detection systems; besides the lack of privacy breaches, external audits must find no anomalies in the internal report for suspicious cross-border transactions.

Component	Description	Purpose
Data Ingestion	Collects transaction and operational data	Enables real-time and batch analysis
Feature Engineering	Transforms raw data into usable features	Improves detection accuracy
Model Training	ML/DL model development and tuning	Detect fraud patterns
Model Inference	Real-time & batch predictions	Detect anomalies instantly



Component	Description	Purpose
Monitoring & Alerting	Tracks system performance and cost	Prevents cost overruns
FinOps Dashboard	Visualizes cloud usage & cost metrics	Supports decision-making

Table 1: Core Components of FinOps-Aware AI Framework

2.1. Real-Time Scoring and Batch Processing Pipelines

Real-time, low-latency inference engines are essential for applications such as fraud detection and prevention. Latency-sensitive consumers rely on these services within products like credit cards or payments with close to immediate action. For instance, a credit card transaction that appears within seconds contains real-time calculations to identify, detect, and discourage instances of fraud. However, other areas in business use batch pipelines that do not require immediate results. Financial institutions need to prepare for various anomalies and criminal activities within transactions, customer support chats, and online services. Real-time scoring engines will push insights to the organization that is beneficial for awareness and acting fast. Even though the action taken may not be immediate and automatic, the awareness generated by the scores helps with human intervention. These invisible scores are created in conjunction with each operation occurring in real-time.

While these low-latency and presence-based engines provide a continuous social feed, the community also needs models that actively learn about behavioral patterns found in these channels. Batch pipelines work behind the scenes, creating neural embeddings, text features, vectors, and predictive outputs for data stored on the cloud. They discover hidden relationships and patterns from customer and business interactions. COVID-19, upcoming events, family introduction, and holiday connections are just a few examples as factors with influence on these signals. Detecting during the period of interaction ensures actions can be taken actively or passively. Detecting patterns, cyber fraud, and disruptions to service like lack of delivery within the expected timeframe provide oversight during demand forecasts for specific products.

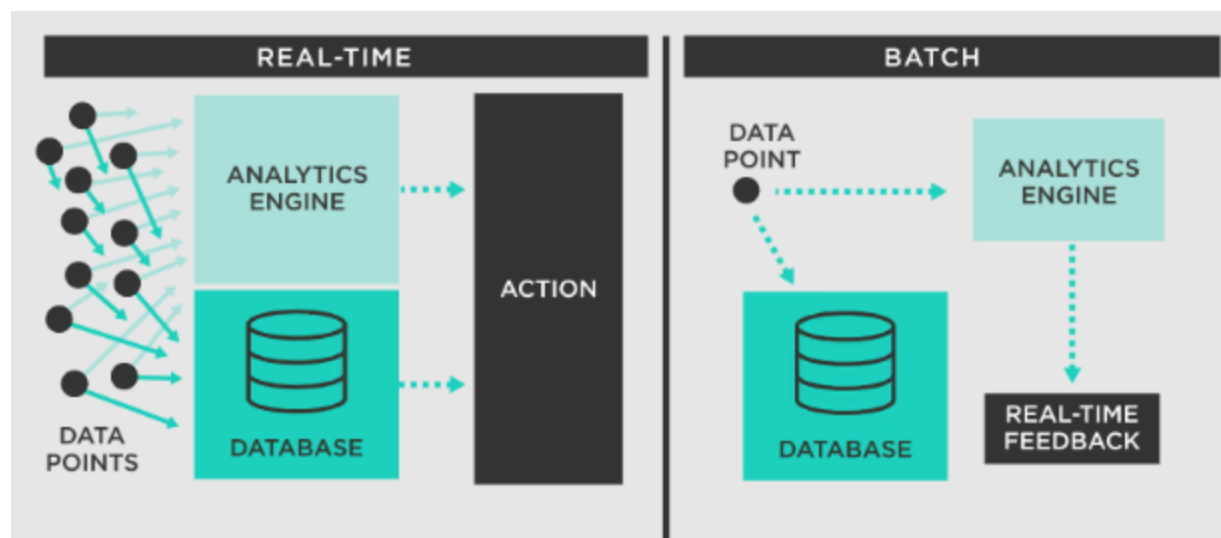


Fig 2: Real-Time Scoring and Batch Processing Pipelines

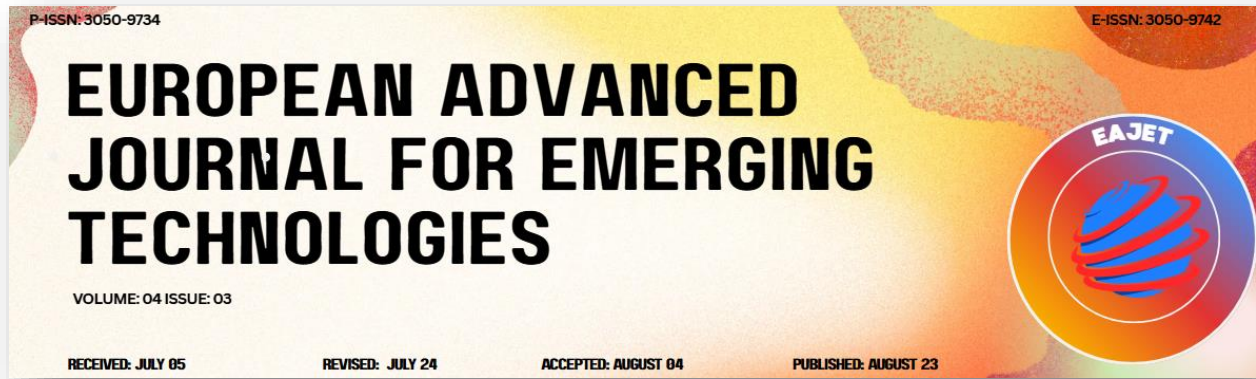
2.2. Privacy, Security, and Regulatory Compliance

The detection of money laundering and terrorist financing activities has become a priority for governments and financial regulators worldwide. Financial institutions are expected to detect suspicious activities with the help of new IT developments in the areas of data science and cloud computing while bearing the operational costs associated with such detection systems. Automated monitoring systems should be built into existing production environments rather than established just for compliance purposes. During system design, objectives beyond detection quality must be pursued to reduce and allocate FinOps-related costs and minimize the likelihood of receiving sanctions from regulators in case of detection failure.

People’s sensitivity about privacy disclosures poses increasing challenges to the pursuit of financial crime detection. Models and systems should be designed to provide assurance to external and internal constituencies about privacy protection. The confidentiality and integrity of the data, models, and systems must also be ensured to avoid harmful job disruption, financial losses, and public discontent that result from system breaches. Public sector budgets have limited resources for maintaining data-protection and privacy-protection systems. For deployed detection models that do not store training data, compliance with GDPR privacy laws is naturally achieved. In such cases, component ownership remains with the organizations that use the models for prediction. These organizations are required to establish their own technological and control frameworks to meet local data protection and privacy laws and guarantee that locally used data do not incorporate any information that can be attributed to individuals.

Equation 2: Cost per Alert / Cost per Anomaly

If total operating cost is spread across detected alerts:



$$CPA = \frac{C_{total}}{A}$$

where A is the number of alerts generated.

If only true suspicious cases matter:

$$CPA_{true} = \frac{C_{total}}{TP}$$

3. Conceptual Framework

FinOps principles require cloud-native applications to ensure financially sustainable usage of cloud services. Significant cost savings are achievable by not only reserving resources but also by implementing monitoring and load prediction strategies. Resources must be allocated according to usage profiles, and alerting capabilities on service usage and cost can help adherence to the FinOps framework.

Specifically designed for financial institutions or cloud service providers offering environments for simulation or sandbox environments, a cloud-native architecture for the real-time detection of financial crime enables extreme scalability, readiness to address regulatory requirements, and rapid adaptation to legislative changes. It consists of streaming and batch modules for the detection of fraudulent anomalies in banking transactions and facilitates the preparation of required evidence for the compliance and regulatory operations. Stream-enabled detection of these crimes is augmented with batch-enabled approaches to detect complex fraud patterns and to consider complex relationships and dependencies at scale.

3.1. FinOps Principles in AI-Driven Cloud Environments

Cloud-native AI systems introduce new cost characteristics. Typically, AI systems are designed to maximize performance without cost-awareness. However, large Cloud service providers offer many services independently, at different rates, and even to different customers at different rates. Therefore, for some users, it may be financially more advantageous to choose a less performant model while maintaining the same accuracy in detection or prediction tasks. FinOps principles estimate cost at all stages of the AI system lifecycle, including preprocessing, feature engineering, model training, model inference, and monitoring. AI costs are also tracked through the monitoring and alerting framework. Using these metrics, budget holders can make informed decisions about deployment and triggering model retraining. FinOps aims to make business owners more cost-aware and offer a radar view of budget spend across detection and prediction use-cases.

Cloud-native AI-integration systems provide an architecture around the deployment of modular systems across subjects with real-time scoring and batch scoring acceleration. The AI-integration framework manages the entire AI model lifecycle, from monitoring concepts to retraining and deployment. The architecture can be hosted either on a private and controlled Cloud or on a well-known public Cloud provider. The cost of performing AI-driven ML, DL, and other Cloud-consuming resources (GPU, TPUs, Storage) is monitored using a combination of Cloud-monitoring tools, alerting mechanisms, and FinOps Dashboarding following FinOps principles and governance. Throughout the monitoring phase, concept drift and accuracy levels are also checked to trigger retraining in an optimized manner while avoiding disruptive spikes in Cloud costs. Cloud-native AI integration

EUROPEAN ADVANCED JOURNAL FOR EMERGING TECHNOLOGIES

VOLUME: 04 ISSUE: 03

RECEIVED: JULY 05

REVISED: JULY 24

ACCEPTED: AUGUST 04

PUBLISHED: AUGUST 23



systems provide a flexible and scalable architecture for deploying modular solutions across multiple domains, enabling both real-time and batch scoring with optimized performance. At the core of this approach is a comprehensive AI-integration framework that governs the entire model lifecycle, including monitoring, concept drift detection, retraining, and seamless deployment. These systems can be hosted on private cloud environments for enhanced control or on public cloud platforms offered by providers such as Amazon Web Services, Microsoft Azure, and Google Cloud Platform, depending on organizational requirements. To ensure cost efficiency, resource consumption across GPUs, TPUs, and storage is continuously tracked using cloud monitoring tools, automated alerts, and FinOps dashboards aligned with governance principles. During the monitoring phase, key performance indicators such as model accuracy and concept drift are evaluated in real time, allowing systems to trigger retraining intelligently while preventing unnecessary computational expenses and avoiding sudden spikes in cloud costs.

EUROPEAN ADVANCED JOURNAL FOR EMERGING TECHNOLOGIES

VOLUME: 04 ISSUE: 03

RECEIVED: JULY 05

REVISED: JULY 24

ACCEPTED: AUGUST 04

PUBLISHED: AUGUST 23



Fig 3: FinOps Principles in AI-Driven Cloud Environments

3.2. Cloud-Native AI Architectures for Financial Crime Detection

Alarming economic losses and reputational damage have prompted the implementation of compliance mechanisms that detect, prevent, and mitigate financial crimes such as money laundering and credit card fraud. Industry leaders and regulators advocate the intelligent use of artificial intelligence (AI) to modernize these capabilities. However, many enterprises view these systems from a short-sighted cost perspective. Realization of FinOps and similar concepts would enable longer-term cost optimization of

cloud-based AI solutions by embedding cost awareness in the entire model life cycle. FinOps stands for the cloud financial management discipline and collaboration process that enables finance, engineering, and product teams to accelerate business value by helping them understand cloud costs and make informed trade-offs, thus embedding cost awareness in every business decision. Given the rapid expansion of cloud and AI technologies, the FinOps framework has considerable potential for the management of financial cloud-native AI solutions.

Cloud-native AI systems that can detect, prevent, and mitigate financial crimes such as money laundering and fraud can be designed and developed based on these principles. Financial institutions, including technology companies that offer financial services, as well as regulatory authorities, require such systems. As financial authorities in different jurisdictions establish regulatory sandboxes where novel business models can be piloted in a controlled manner, approval and monitoring of cross-border transactions, especially those involving cryptocurrencies, gain renewed prominence. Regulatory authorities also require cost-effective scaled compliance solutions to support their operations and to oversee regulated entities that offer their services at scale.

Equation 3: Precision, Recall, and F1 for Fraud Detection

Precision measures how many flagged transactions were actually suspicious:

$$Precision = \frac{TP}{TP + FP}$$

Recall measures how many suspicious transactions were caught:

$$Recall = \frac{TP}{TP + FN}$$

F1 combines both:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

Substitute precision and recall:

$$F1 = \frac{2TP}{2TP + FP + FN}$$

3.3. Compliance Operations and Cost Governance

Detecting financial crime is not the only area where cloud-native AI systems can provide significant benefits. Supporting compliance operations is just as important. Anti-money laundering compliance operations are costly, especially for smaller banks, and the costs are passed on to customers. Automating these operations is therefore essential, but the resulting systems can be very expensive to run. Therefore, when such systems are implemented, they must be cost-optimized using FinOps principles.



The mission of a financial institution is to serve its customers. Any funds used to operate the bank's compliance functions represent funds that cannot be used to support customers. For this reason, automated systems designed to support compliance operations need to be made as cost-effective as possible. By employing FinOps principles, the costs to support these compliance systems can be optimized and effectively managed. Long-standing principles of cost optimization in cloud environments are known as FinOps. The FinOps Foundation describes these principles as follows: teams need to work together with a shared mindset, decisions must be driven by a business context, cloud costs need to be owned by a variety of stakeholders, real-time accountability must enable action, and forecasting must be part of the daily conversation. These principles can be directly applied to cloud-native AI systems using these systems for compliance automation, to help manage the cost of supporting these solutions.

Flow chart:

Start



Financial transaction / customer / operational data generated



Cloud-native data ingestion



Data quality assurance



Feature engineering



Real-time scoring OR batch processing



AI/ML model detects anomaly, fraud, AML risk, or compliance issue



Risk score generated





Decision support / alert / investigation workflow



Compliance action or fraud prevention action



Cloud resource usage monitored



FinOps dashboard tracks cost, performance, and resource utilization



Autoscaling / cost optimization / retraining decision



Continuous monitoring and governance



End / Repeat continuously

4. System Architecture and Design

High-throughput data ingestion and quality assurance, scalable feature engineering, and a cost-governed model lifecycle management are critical components of a FinOps-aware, cloud-native AI system for real-time anomaly and fraud detection in financial transactions. Data ingestion pipelines are generator-like function clouds that receive feature data from supervised learning and feature-reduction processes. Passing the data through quality-assurance modules assures the integrity of the feature data, which is then written into properly partitioned columns in object storage and properly appended in a data warehouse. Data-driven FinOps policies for dynamic resource allocation and autoscaling enhance the cost-efficiency of these operations. Cloud-native feature-engineering processes interpret raw features using a set of rules defined by domain experts, and the processed features are used for real-time anomaly and fraud detection.

Model lifecycle management comprises a set of cost-aware policies for the training of ML models, and resource support for training and inference should be monitored and governed following FinOps principles. Periodically retrained models mitigate concept drift by adapting to new patterns in the feature data, while relevant parts of the model suite can be trained in batches to invoke savings from spot instances. The on-demand training of constrained GANs enforces a cost-aware use of these complex

neural-network architectures. Inference-traffic attributes provide an indication of how traffic changes and, consequently, how ML model inference should be allocated. A dashboard empowers the FinOps team with critical capabilities for end-to-end observability of the system and its cost governance.

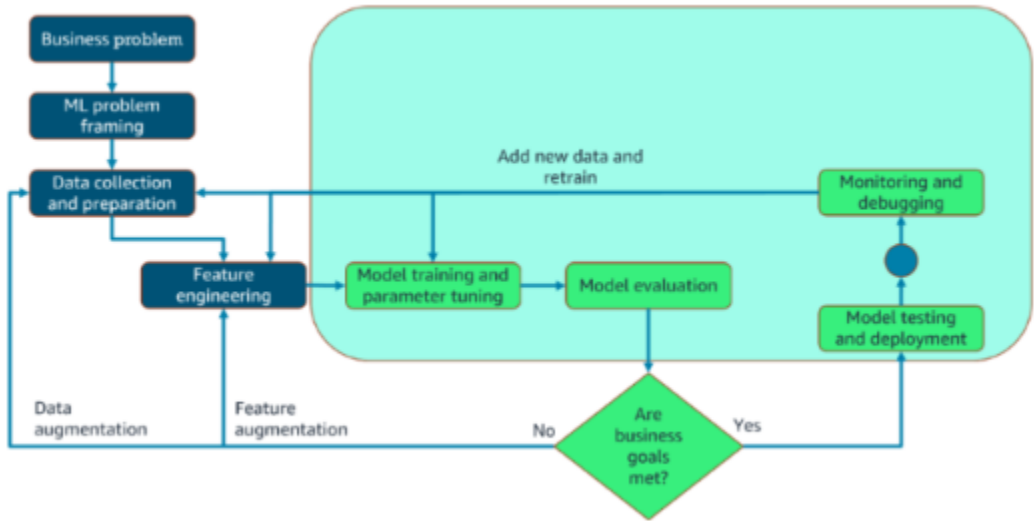


Fig 4: System Architecture and Design of finops

4.1. Data Ingestion and Quality Assurance

Operational data from product systems across a bank or payment service provider are the primary input for a cloud-native detection architecture. For systems that span multiple operational entities or regions, the cloud design allows for a regional separation of processing and can support considerations of data sovereignty. Data quality plays a crucial role in accurate detection. In addition to standard availability, completeness, and consistency checks, tests for plausibility and validity are essential to ensure that the data match expected profiles. Due to econometric constraints in data quality assessment, these controls may be restricted to selected combinations of anomalies and fraud types. Nevertheless, the costs of releasing low-quality detection results—with potential negative implications for customers and partners—substantiate the need for a well-designed quality control system.

Equation 4: Autoscaling Resource Requirement

Let incoming transaction rate be λ , and one node can process μ transactions per second.

Required nodes:



$$N = \left\lceil \frac{\lambda}{\mu \times U_{target}} \right\rceil$$

where U_{target} is the desired utilization limit.

Example derivation:

$$Capacity = N \times \mu \times U_{target}$$

To avoid overload:

$$N\mu U_{target} \geq \lambda$$

So:

$$N \geq \frac{\lambda}{\mu U_{target}}$$

Because nodes are whole units:

$$N = \left\lceil \frac{\lambda}{\mu U_{target}} \right\rceil$$

4.2. Feature Engineering for Anomaly and Fraud Detection

Cost-aware feature engineering contributes to accurate and cost-efficient detection of incidents and compliance breaches. For supervised models, data quality assurance must include disambiguation and annotation, while unsupervised approaches require working with the right concepts and refined indicators. Constructing models that reduce effort on poor-quality data sources is also valuable, as it lowers noisy data processing volume. Exploratory data analysis guides geography- and domain-specific feature creation, and foundation models can automatically identify data distributions, feature candidates, and correlations.

Financial crime and operations compliance are inherently complex and diverse, spanning cultural domains and multiple topics, methods, and players. Expensive and poorly designed models do not scale well; when an area reaches some maturity, models can be broken down into more focused subsystems that reuse success, transfer learnt experience, share data and features across channels, use different models for particular situations, and employ supervised models only for easy examples. Proper sitting of model boundaries reduces the classification effort of adjoining models. Models must also account for possible evasion attempt features and make use of pre-sale or pre-New Year discussions documenting unusual hypothesis and plans.



Aspect	Real-Time Pipeline	Batch Pipeline
Latency	Very low (milliseconds)	High (minutes–hours)
Use Case	Fraud detection, credit card transactions	Pattern discovery, model training
Data Type	Streaming data	Historical data
Cost Impact	High (continuous compute usage)	Moderate (scheduled processing)
Action	Immediate alerts	Insights & model improvement

Table 2: Real-Time vs Batch Processing Pipelines

4.3. Model Lifecycle Management in Cloud-Native Environments

Lifecycle management throughout the model development, training, validation, deployment, monitoring, and updating phases is one of the primary features of cloud-native AI systems. A cloud-native environment enhances the development phase by providing an integrated DevOps ecosystem for development teams to automate the software development and delivery processes. An automated CI/CD pipeline ensures that applications, including model inference services, are built, tested at several levels, and deployed automatically to the cloud whenever new code is checked in. These include unit tests, infrastructure tests, performance tests, and acceptance tests. Services for feature extraction, training, validation, and experimentation with new models can be documented as code and executed using CI/CD principles. With these features, different teams work in parallel on different phases of a model's development lifecycle.

During the training and validation phases, the availability of testing and staging environments allows data science teams to train, validate, and test multiple models concurrently on the same infrastructure. Since there are no manual processes, all stages are automatically monitored, enabling data teams to detect issues such as data drift and resource contention early. Service and operational teams benefit from cost monitoring and anomaly detection as they can identify congestion in testing and training pipelines. As the model is deployed into production, FinOps enable operations teams to provision resources that are aligned with risk levels and model usage patterns.

5. FinOps Strategies for Cost-Optimized Operations

Cost governance policies for cost-optimized operations specify levels of quality for machine learning services and provide guidance for both model training and service deployment. Resource allocation, including use of dedicated clusters, and autoscaling policies limit waste and can also be tuned to model and data characteristics. Cost governance policies for cost-optimized operations specify levels of quality for machine learning services and provide guidance for both model training and service deployment. Resource allocation, including use of dedicated clusters, and autoscaling policies limit waste and can also be tuned to model and data characteristics.



Cost governance policies for AI-driven systems define the quality levels required for machine learning services and provide guidance for model training and service deployment. Resource allocation choices, such as whether to use dedicated clusters, together with autoscaling configurations, help to prevent waste and can be tailored to the requirements of specific models and data sets. Monitoring tools that track the costs of training, validation, and inference, complemented by alerting systems, support day-to-day cost management and enable FinOps dashboards to display overall trends and recommend adjustments.

Flow chart:

Start



Incoming transaction occurs



Transaction data captured instantly



Streaming ingestion service receives data



Data validation and quality checks



Feature extraction in real time



Model inference service scores transaction



Risk score generated



Is risk score above threshold?



— No → Approve / continue normal processing

— Yes



Generate fraud / AML alert



Send to rules engine or compliance officer



Human or automated decision



Block, freeze, review, or report transaction



Log outcome for monitoring and model improvement



End

5.1. Resource Allocation and Autoscaling Policies

Resource allocation plans define which resources are allocated for which period in a cloud environment. A minimum resource level must be maintained to avoid heavy startup costs when resources are powered down and restarted. The number of instances and nodes can be automatically adjusted based on actual usage by autoscaling policies that define the number of instances when idle or underutilized. Static autoscaling policies may offer sufficient cost optimization while remaining simple to implement and manage. Utilization-based dynamic autoscaling policies can minimize costs further but require a more sophisticated implementation.

Resource demand prediction models based on historical usage data can identify busy periods. Dairy solutions may concentrate training efforts in these periods and limit or reduce activities during other periods. A fluid demand flow that fluctuates continuously at a fine granularity can be tackled with event-based or trigger-based dynamic autoscaling policies. Non-AI models, AI-assisted models, and AI-native models can be selected depending on demand patterns, required accuracy, and training and inference costs.



5.2. Cost-Aware Model Training and Inference

The development, tuning, and deployment of AI models in the cloud incur resource costs that can be volatile, and therefore difficult to monitor and predict. During model training, training data are often stored in remote object storage, with data accessibility demanding resource allocation and cost. Model training can also involve hyperparameter tuning, during which multiple instances of training may have to be executed in parallel. Cost-aware policies for model training reduce unnecessary resource allocation while leveraging lower-cost resources whenever they are available. This, in turn, requires a better understanding of resource workloads during model training, so that decisions on training resource allocation can be made ahead of time.

Despite these expenses, cloud providers offer lower-cost resources, such as preemptible virtual machines, system containers, and serverless containers. Preemptible virtual machines are offered by most public cloud providers, with very low costs compared to standard virtual machines. They can be used for model training, but have the drawback that instances can be preempted at any time. System containers are virtual machines for very granular workloads, with an overhead that is much smaller than that of full virtual machines. Serverless containers are designed to store application code, but can also be used for one-off data processing, such as model training. The cost-awareness of model inference can be enhanced if both the availability of preemptible virtual machines and the workload of existing model-serving resources are monitored, so that preemptible virtual machines can be provisioned for model inference when they are available and there is no existing demand on model-serving resources. Since the serving of AI models is usually cost and latency sensitive, it is preferred that model inference executes on such resources compared to standard virtual machines.

Equation 5: Budget-Constrained Model Selection

Choose the model with the best detection performance while staying within budget:

$$\max F1_m$$

subject to:

$$C_m \leq B$$

where:

$$C_m = C_{train,m} + C_{infer,m} + C_{monitor,m}$$

So the selected model is:

$$m^* = \arg \max_m \{F1_m \mid C_m \leq B\}$$

5.3. Monitoring, Alerting, and FinOps Dashboards

An essential requirement for successful FinOps governance is cost-awareness of all FinOps participants, beginning with FinOps



practitioners and CI/CD teams. Specific dashboards based on the cost metrics outlined above should be considered for these groups and tailored for the different roles involved. For example, model trainers need dashboards to monitor inference costs, investigate unusual spikes, and correlate them with hyperparameter tuning. Pipelines that ingest rarely used road traffic flow data or models requiring abortion of batch job executions or cannot be autoscaled during batch inference should have cost alerts configured. Similarly, alerting thresholds need to be defined for the data engineering teams building ETL pipelines for compliance-related data sources. Rightsizing infrastructure allows FinOps teams to adapt the assigned resources for anticipated upcoming increase in volume and usage patterns at lower costs. In the regulatory sandbox setting it is mandatory to exhibit explainable actions on missed detections and in general the confidence levels of detections to support a closer look for further investigation. Detecting new classes of anomalies or fraud patterns relying on different features and not present in training data is expected from the models. If explainability is missing or does not provide value the next phase avoids cost overruns with the addition of explainability. Complying with regulators Laws from different jurisdictions requires an audit function within deployed models.

Supporting cloud-driven native and cost-optimized compliance operations in the regulatory sandbox environment need further focus. The resources of the AI engine are allocated for other primary business cases. Volume surges at transaction frontend usually in holiday seasons or political events accumulated with high expensive transactions raises alerts. Time sensitive alerts triggered in burst must have confidence weights defined and tuned. Using opacity of confidence in decisions attempted should be monitored at cost efficiency especially in geofence cases. Machines supporting compliance operations especially monitoring the cloud service traffic profile for source network address should have traffic simulated via traffic injection tests in order to use the profile volume and cost shift between distant cheap cool and warm regions be utilized. Privilege solution operating each individual machine or small cluster of machines must adhere to atleast best practices of sector confidentiality only supporting phase of profile injection tests in sensitive transition need be elevated to high privilege level.

6. Evaluation and Validation

Select subcomponents of the FinOps-aware cloud-native architecture can be considered in isolation, as in the case of any data-driven system or platform. Formal evaluations usually address one or two quality aspects at a time, employing various datasets, validation procedures, and performance metrics. Performance for financial crime detection under different alerting thresholds constitutes the most relevant evaluation for potential service providers. The failure of a given provider to detect anomalous or obviously fraudulent activity during validation is usually deemed a sufficient reason not to use the provider, leading to a natural focus on the corresponding false negative rate.

Performance-efficiency trade-offs become central for cost-sensitive users of financial crime detection services. Data identifiers can aid monitoring ongoing performance to permit FinOps decision making by the customer. Frequently suggested metrics, such as precision, receiver operating characteristic curves, and area under the curve, together with the CPA metric introduced above, enrich the characterization of performance and trade-offs. Usage of these metrics across subsequent papers provides the foundation for budget-aware model training and inference decisions. Robustness and anomaly explainability are recognized as important aspects of any supervised anomaly detection system but are less commonly analyzed. The tools and techniques in common use for interpreting machine-learning models under various circumstances support the examination of such aspects.

Validation and audit procedures from the realm of regulatory compliance serve to maintain the quality of compliance operations, although the specific elements of interest can differ between the systems and the regulatory requirements for data-powered operational support tools. Predictability under FinOps-aware conditions provides an orientation for the design and implementation of all supporting systems.

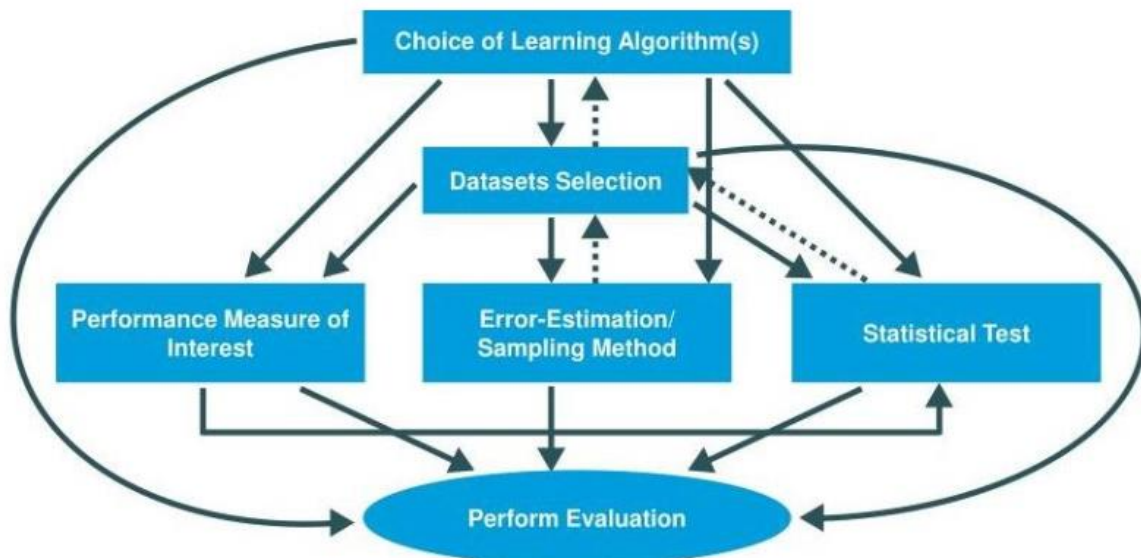


Fig 5: Evaluating Machine Learning

6.1. Performance Metrics for Detection Systems

Most financial crime detection systems monitor transactional data continuously and perform real-time scoring to identify suspicious transactions as they happen. Suspicious transactions are then subject to further investigation, typically through a combination of business rules and human intervention. Such a near-real-time operational model requires specific monitoring and alerting capabilities to ensure the model's performance remains within acceptable bounds and that any emergent drifts or model failures are promptly addressed.

For real-time deployment of machine learning inference engines, the architecture should help expedite deployment of changes to the models governing detections and adapt to the needs of business. All modifications to the models or models files – whether through model rebuilding or transfer learning – ought to be detectable and management notification automated, from the linking of the modified model version(s) to a production inference service through testing pipelines to the cutting over of traffic. Business presentation and true positive notification of anomalies detected by the model should also lend themselves to efficiency gains through FinOps principles.



Principle	Description	Impact
Cost Visibility	Track cloud usage across lifecycle	Better budgeting
Resource Optimization	Allocate resources dynamically	Avoid waste
Accountability	Shared ownership across teams	Cost control
Forecasting	Predict future usage and cost	Planning efficiency
Continuous Monitoring	Alerts for abnormal spending	Prevent overspending

Table 3: FinOps Principles Applied to AI Systems

6.2. Cost-Efficiency Metrics and Trade-Offs

Selecting the optimal cloud configurations incurs testing costs that are usually not repaid through very low operational costs. A dedicated cloud operation dashboards includes continuous financial monitoring based on cloud billing exports with time-retention capabilities. Multiple cost thresholds notify FinOps engineers for abnormally large surges or drops in cloud spending. Setting separate cost thresholds for the cost of each cloud service at different times of day also helps to identify potential misuse. Whenever spending exceeds pre-defined cost thresholds, additional checks ensure that it is justified by abnormal peaks in demand. During operational phases, automating FinOps decision making helps to redirect more effort toward model and application quality. Alerts are continuously fine-tuned, combined, or subdivided based on frequent patterns and FinOps resources.

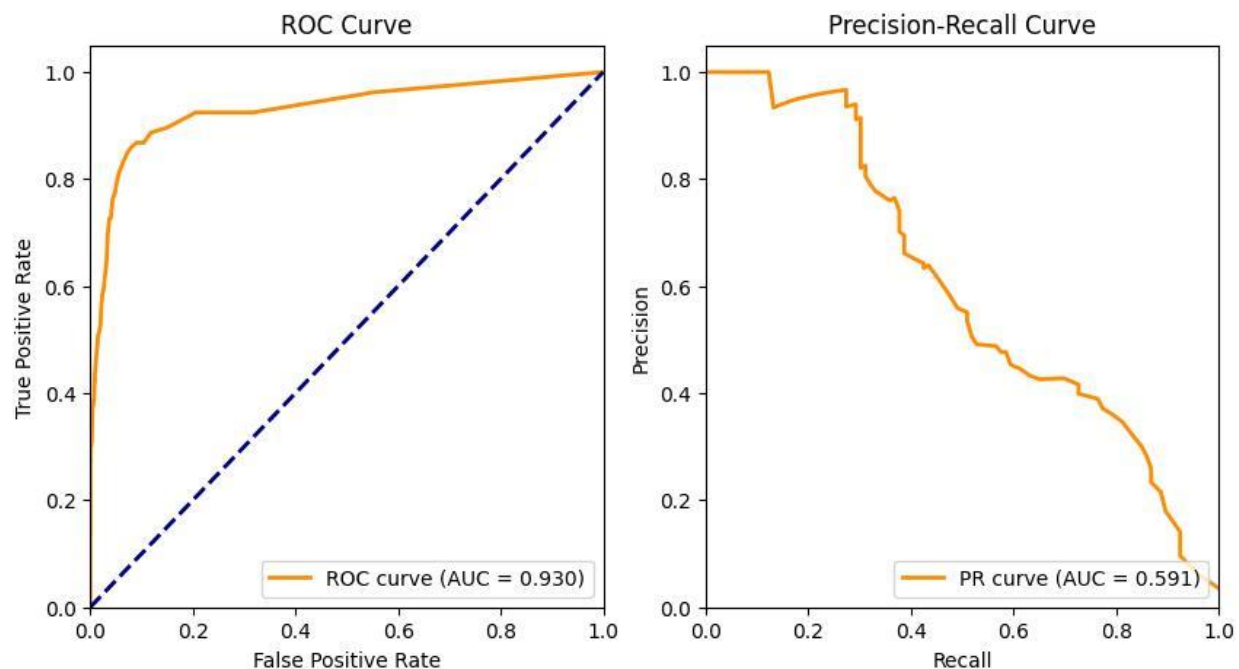
With service cost commitments, training and inference for which models need not be permanently hosted become more complicated due to the additional cost of cloud-native storage solutions. Bank of America estimates that a 10% of the training and re-optimization budgets that are spent on temporary hosting can substantially reduce deployment costs from model latency constraints. Timing training and inference to coincide with cheap cloud resources greatly expedites all aspects of a model's lifecycle. Cost-aware algorithm selection also plays a huge role in FinOps-compliant algorithms operating in cloud-based settings. Predictive cloud-mining selects those models to run that are known or predicted to display the best accuracy on unseen data.

7. Deployment Scenarios and Use Cases

Financial institutions can deploy scalable cloud-native AI-driven detection systems within their premises, as a test environment operating under a regulatory sandbox, or in partnership with a cloud service provider. Systems within such a regulatory environment operate under direct scrutiny by supervisors and authorities, thus ensuring trustworthy and secure model development, testing, validation, and deployment. Deployment of new systems within the virtual private cloud of a major cloud service provider enables the seamless gathering of traffic over cloud borders, minimising latencies and the costs associated with

transmitting data. Such a setup can serve as an integrated service hosted in the largest FinTech company active in the international remittance industry, guaranteeing fast response times.

A global anomaly detection system, for different remittance payment flows of major financial institutions, can help optimise the allocation of supervision resources. Major banks are required to monitor cross-border transactions with strict limits in terms of time and costs, involving large volumes and limited predictable patterns, thus making the identification of potential money laundering difficult. Anomaly detection allows the identification of unusual activities for subsequent application of standard anti-money laundering procedures by the corresponding enforcement departments on the identified suspicious cases. The desired routing of training data toward specific cloud regions allows the real-time model to benefit from traffic across all regions of the world while optimising the cost associated with batch training by prioritising the region of largest anomaly during each model cycle. The solution can be extended to include additional modules supporting regulatory compliance operations (anti-money laundering, counter-terrorism financing, domestic fraud, and data protection), preparing the infrastructure for growing traffic in an efficient and trustworthy manner.



7.1. Financial Institutions and Regulatory Sandboxes

Conceptual architectures that incorporate the principles of FinOps in an AI-supported cloud environment and typical design patterns for cloud-native AI systems deployed by financial institutions should facilitate the development and deployment of such systems aligned with budget governance principles. FinOps-aware AI systems capable of scaling up to meet peak demand can but are not limited to be deployed in the cloud infrastructure of any public cloud service provider. As an illustrative example, the

architecture and design of a fully managed real-time lake house platform in the cloud are presented: Streaming data from multiple sources are ingested, processed, and stored in near real-time for victim profiling and real-time scoring of anomalies and fraud. A separate batch-processing pipeline is also embedded in the architecture for periodic training of ML models used in fraud detection. The platform hosts data analytics and BI capabilities for data exploration, monitoring, alerting, and reporting. An illustrative scenario focusing on anti-money laundering analytics demonstrates these capabilities. Such an architecture can be beneficial for banks and financial institutions that have adopted a platform or lake-house approach and for AI solutions that form an integral part of the overall product offering.

An alternative deployment option enables rapidly developing and deploying FinOps-aware AI systems in a regulatory sandbox. Financial regulators across jurisdictions have set up such sandboxes to allow financial institutions and regulatory technology companies to rapidly develop and test their solutions. By integrating regulatory sandboxes with their cloud-native open banking platform, cloud service providers can secure the attraction of RegTech companies. These sandboxes support faster solution development but lack the necessary production-ready infrastructure. FinOps-aware cloud-native AI systems integrated with a regulatory sandbox are expected to accelerate the development and deployment of solutions that foster a secure and trusted modern financial ecosystem while supporting the Full Digital Economy.

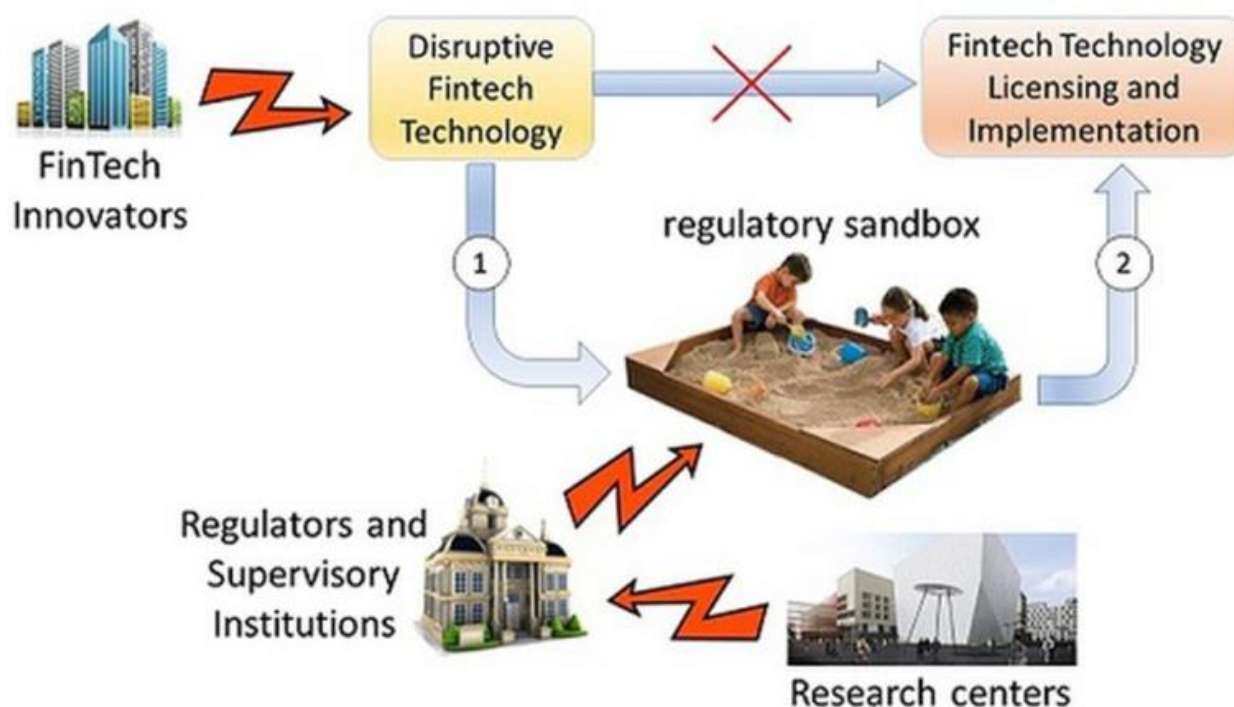


Fig 6: Framework For Regulatory Sandbox



7.2. Cross-Border Transactions and Anti-Money Laundering

Cross-border transactions are also subject to national and international anti-money laundering (AML) directives aimed at monitoring suspicious financial transactions in the banking, insurance, and gaming sectors. Banks with operations in jurisdictions designated as non-compliant with AML standards must take additional steps to ensure that their customers are not sending money to or receiving money from such jurisdictions. Transactions involving sanctioned entities are monitored and flagged by transaction monitoring systems, but the bank must also be alert for other possible money laundering violations.

An AI-based system capable of analyzing a cross-border payment—examining the customer sending the payment, the beneficial owner(s) of the payment, the country of the beneficiary bank, the organization sending the payment, and the payment narrative—to assess the risk of money laundering and to detect attempts to disguise the true nature of the transaction and the parties involved can augment a bank's compliance efforts. Natural language processing techniques can be employed to analyze the payment narrative or message, which is mostly in free text, from payment systems such as SWIFT or CHIPS. Such a system, coupled with a user-friendly decision support interface, will allow the reviewing officers to make risk decisions and document their rationale expeditiously. Integration of the AI judgment into the review workflow will enable the bank to manage the volume of risk reviews without adding significant headcount.

8. Conclusion

The proposed architecture and FinOps strategies create a conceptual framework for planning, developing, validating, operating, and maintaining cloud-native AI systems capable of scaling according to the demand for real-time anomaly and fraud detection or regulatory compliance. But developments in the data economy are shaping the longer-term vision and deployment planning. While the past few decades saw a race to create affordable yet persistent PII-rich data reservoirs for training reusable AI models of high quality, the balance is now in transition toward narrow-focus transfer-learning applications with low demand for training data or resources—a trend that is likely to accompany the increasing business focus on AML compliance in the context of transaction monitoring and provide a foundation for AI-supported compliance operations at scale and cost.

The initial provisioning of training data is expensive and often comes with high permission requirements that delay or even block deployment for money-laundering transaction monitoring. AI systems acting on behalf of compliance functions can exploit economies of scale and be trained using the less sensitive real-time data generated by low-FP detection pipelines and participating financial institutions, whose models can therefore be run in sandbox environments with limited or no training data or compute resources. These trends lower the overall cost and regulatory risk of recurrent AI system validation for narrow areas of application and compliance. Test datasets updated to reflect the constantly evolving properties of fraudsters and their techniques will be easier to prepare and maintain without infringing privacy. In summary, FinOps-aware cloud-native AI systems designed for scalable active threat detection can also facilitate the provision of covered or supporting compliance operations at market-surveyed cost.

EUROPEAN ADVANCED JOURNAL FOR EMERGING TECHNOLOGIES

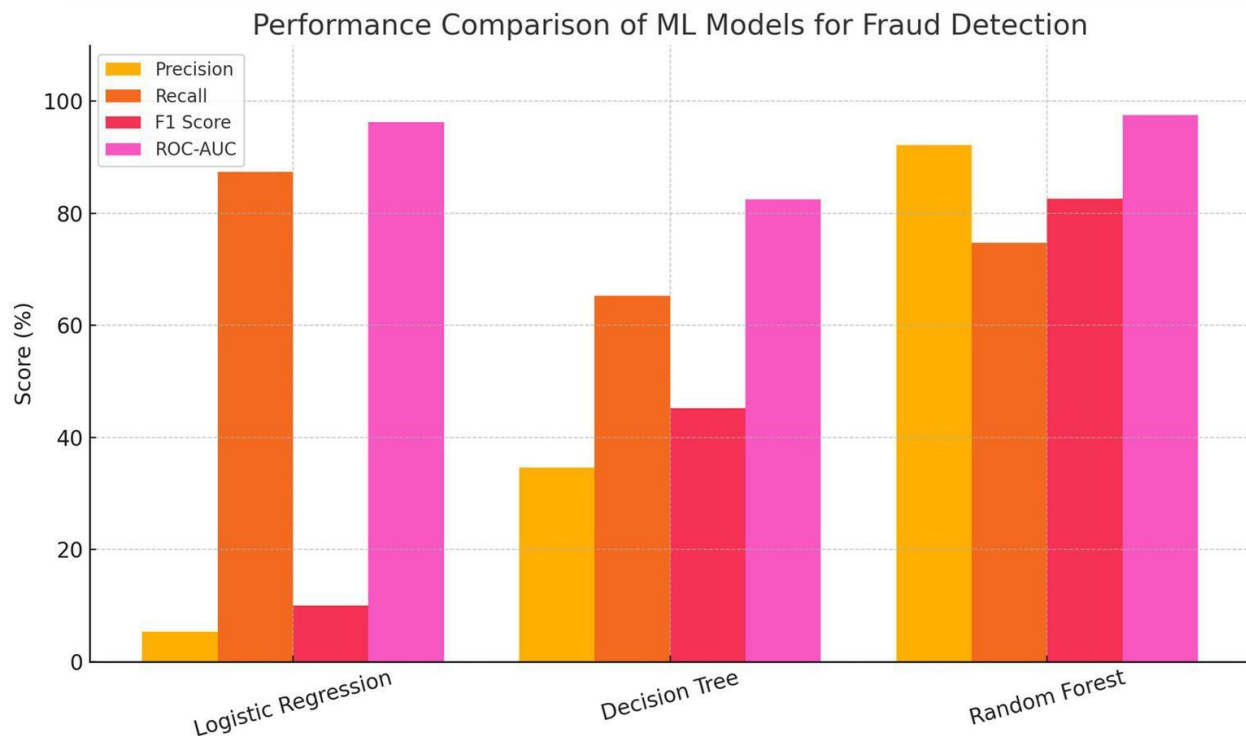
VOLUME: 04 ISSUE: 03

RECEIVED: JULY 05

REVISED: JULY 24

ACCEPTED: AUGUST 04

PUBLISHED: AUGUST 23



8.1. Emerging Trends

A growing number of financial institutions are employing cloud-native AI systems in their FinTech or RegTech infrastructures to reduce technology costs while improving business effectiveness. Nevertheless, the limited financial resources of these organizations for AI infrastructure, together with the need to demonstrate compliance with stringent legal obligations, constitute a great obstacle for the use of complex AI techniques. Indeed, allowing real-time detection of money laundering and drug trafficking operations, these AI systems also facilitate the instantaneous freezing of funds or assets involved in illicit activities. The banking and finance sector deals with enormous amounts of structured, semi-structured, and unstructured data every day, including transactional data involving a large number of entities. The collection of this information requires a sophisticated and reliable architecture for data quality control.

Furthermore, the absence of cloud usage control can lead to high costs due to natural traffic growth, overprovisioned resources, and uncontrolled deployments or queries. Given the resources consumed, AI-driven solutions must be the best option for cost-effective new purchases, with cost management integrated into the decision for the charge of credit cards, classification of product transactional data, or marketing of specific products. Emerging research directions within this compact area of investigation include integrating explainability, distributional robustness, detection of model drift, fairness, security and privacy, cross-cloud model deployment, support for human-computer interaction, trustworthy operations, and multi-modal model observation.



9. References

1. Al-Hashedi, K. G., & Magalingam, P. (2021). Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019. *Computer Science Review*, 40, Article 100402.
2. Alexandre, C. R., & Balsa, J. (2023). Incorporating machine learning and a risk-based strategy in an anti-money laundering multiagent system. *Expert Systems with Applications*, 217, Article 119500.
3. Bakry, A. N., Alsharkawy, A. S., Farag, M. S., et al. (2024). Automatic suppression of false positive alerts in anti-money laundering systems using machine learning. *The Journal of Supercomputing*, 80, 6264–6284.
4. Bhandari, P., Prithviraj, J., Saxena, R., Kumar, D., Mangala, N., & Chourasiya, P. (2026). Harris Hawks Optimization–Based Deep Learning for Resource-Aware Cloud–Fog Computing. In 2026 5th OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 5.0 (pp. 1–6). IEEE. 2026 5th OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 5.0. <https://doi.org/10.1109/otcon69332.2026.11629982>
5. Błaszczyński, J., de Almeida Filho, A. T., Matuszyk, A., Szeląg, M., & Słowiński, R. (2021). Auto loan fraud detection using dominance-based rough set approach versus machine learning methods. *Expert Systems with Applications*, 163, Article 113740.
6. Canhoto, A. I. (2021). Leveraging machine learning in the global fight against money laundering and terrorism financing: An affordances perspective. *Journal of Business Research*, 131, 441–452.
7. Seenu, A., Kolla, S. H., Aitha, A. R., Amistapuram, K., Abhireddy, N., & Yandamuri, U. S. (2026, June). Agentic Enterprise Intelligence: A Unified Architecture for Secure RAG-Enabled LLM Automation in Insurance Risk and Fraud Ecosystems. In 2026 5th OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 5.0 (pp. 1-5). IEEE.
8. Goecks, L. S., (2022). Anti-money laundering and financial fraud detection: A systematic literature review. *Intelligent Systems in Accounting, Finance and Management*.
9. Hilal, W., Gadsden, S. A., & Yawney, J. (2022). Financial fraud: A review of anomaly detection techniques and recent advances. *Expert Systems with Applications*, 193, Article 116429.
10. Avinash, K., Ranga Reddy, V. A., Manohari, D., Kothuri, S., Geetha, R., & Swarna, B. (2026). Multi-Source Feature Integration for Short-Term Load Forecasting in Off-Grid Residential Microgrids Using Stacked LSTM Networks. In 2026 14th International Conference on Smart Grid (icSmartGrid) (pp. 501–505). IEEE. 2026 14th International Conference on Smart Grid (icSmartGrid). <https://doi.org/10.1109/icsmartgrid70882.2026.11680767>
11. Jensen, R. I. T., & Iosifidis, A. (2023). Qualifying and raising anti-money laundering alarms with deep learning. *Expert Systems with Applications*, 213, Article 119037.
12. Karangara, R., & Subbagari, S. (2024). Filtering and detection of anti money laundering with the aid of optimization-enabled machine learning techniques. *International Journal of Computer & Organization Trends*, 14(1), 1–11.
13. Singh, K. U., Gupta, S. K., Dubey, Y., & Mangalampalli, B. M. (2026, June). Enhancing Healthcare Cyber Defense with Machine Learning-Based Attack Prediction and Prevention. In 2026 5th OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 5.0 (pp. 1-6). IEEE.



14. Lim, S. T., Khoo, R. Q., Khaw, K. W., & Chew, X. (2025). Identifying potentially illicit money laundering and terrorism financing transactions through machine learning techniques. *International Journal of Management, Finance and Accounting*, 6(2), 243–267.
15. Lokanan, M. E. (2024). Predicting money laundering using machine learning and artificial neural networks algorithms in banks. *Journal of Applied Security Research*, 19(1), 20–44.
16. Raut, B., Garg, A., Mangalampalli, B. M., Jindal, N., Mahajan, V., & Saraswati, K. (2026). Multidisciplinary Strategies for Improving Maternal and Reproductive Health Outcomes and Mental Outcomes. *International Journal of Drug Delivery Technology*, 16(34s). <https://doi.org/10.25258/ijddt.16.34s.36>
17. Pocher, N., Zichichi, M., Merizzi, F., Shafiq, M. Z., et al. (2023). Detecting anomalous cryptocurrency transactions: An AML/CFT application of machine learning-based forensics. *Electronic Markets*, 33, Article 37.
18. Ramadhan, S. (2025). Harnessing machine learning for money laundering detection: A criminological theory-centric approach. *Journal of Money Laundering Control*, 28(1), 184–201.
19. Engineering Intelligent Cloud-Native Data Ecosystems for Predictive Decision-Making in Industry. (2026). *Journal of European Economic History*, 7(2), 68–88. <https://doi.org/10.61336/jeeh/26-2-7> (Original work published as Engineering Intelligent Cloud-Native Data Ecosystems for Predictive Decision-Making in Industry)
20. Sinno, R. M., Baldock, G., & Gleason, K. (2023). The evolution of trade-based money laundering schemes: A regulatory dialectic perspective. *Journal of Financial Crime*, 30(5), 1279–1290.
21. Tuffveson, R. I. J., & Iosifidis, A. (2022). Qualifying and raising anti-money laundering alarms with deep learning. *Expert Systems with Applications*.
22. Rani, S., Thavara, S. S., Kr, A., Naguband, A. R., Garg, A., & Vajpayee, A. Financialization of Sustainability: ESG Signalling, Market Valuation and the New Economics of Corporate Legitimacy.
23. Gubran, K., & Magalingam, P. (2021). Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019. *Computer Science Review*, 40, Article 100402.
24. Bakry, A. N., Alsharkawy, A. S., Farag, M. S., et al. (2024). Automatic suppression of false positive alerts in anti-money laundering systems using machine learning. *The Journal of Supercomputing*, 80, 6264–6284.
25. Alexandre, C. R., & Balsa, J. (2023). Incorporating machine learning and a risk-based strategy in an anti-money laundering multiagent system. *Expert Systems with Applications*, 217, Article 119500.
26. Anand, V., A, S. K., Vemulapalli, N., Loganathan, R., Guntupalli, A., & Ahammed, S. R. (2026). Federated Deep Learning for Decentralized Cybersecurity in Smart Grid Communication Networks. In 2026 14th International Conference on Smart Grid (icSmartGrid) (pp. 475–481). IEEE. 2026 14th International Conference on Smart Grid (icSmartGrid). <https://doi.org/10.1109/icsmartgrid70882.2026.11680752>
27. Lim, S. T., Khoo, R. Q., Khaw, K. W., & Chew, X. (2025). Identifying potentially illicit money laundering and terrorism financing transactions through machine learning techniques. *International Journal of Management, Finance and Accounting*, 6(2), 243–267.
28. Goecks, L. S. (2022). Anti-money laundering and financial fraud detection: A systematic literature review. *Intelligent Systems in Accounting, Finance and Management*.
29. Babu, A. J., Yandamuri, U. S., Recharla, M., Pamisetty, A., Goli, M., & Kolla, S. H. (2026, June). Data Analytics and AI Solutions for Digital Transformation in Hospitality and AgriTech Industries. In 2026 5th OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 5.0 (pp. 1-7). IEEE.
30. Intelligent financial fraud detection practices in post-pandemic era. (2021). *The Innovation*, 2(4), Article 100176.
31. Intelligent anti-money laundering fraud control using graph-based machine learning model for the financial domain. (2023). *Journal of Cases on Information Technology*, 25(1).

EUROPEAN ADVANCED JOURNAL FOR EMERGING TECHNOLOGIES

VOLUME: 04 ISSUE: 03

RECEIVED: JULY 05

REVISED: JULY 24

ACCEPTED: AUGUST 04

PUBLISHED: AUGUST 23



32. Nemani, R., Singuluri, I., Ramani, V. S., Guntupalli, A., Mopathi, U., & Madineni, A. V. S. (2026, July). A Machine Learning-Augmented Penalty Difference Approach for Transportation Cost Optimization. In 2026 14th International Conference on Smart Grid (icSmartGrid) (pp. 511-516). IEEE.
33. Application of machine learning models and artificial intelligence to analyze annual financial statements to identify companies with unfair corporate culture. (2020). *Procedia Computer Science*, 176, 3037–3046.
34. Exploratory analysis and implementation of machine learning techniques for predictive assessment of fraud in banking systems. (2022). *Procedia Computer Science*, 214, 117–124.
35. Kolla, S. K., Bandi, V. D. V. K., & Meda, R. (2026). Comment on “Predicting self-image satisfaction after adult spinal deformity surgery: a machine learning approach using patient phenotypes.” *Spine Deformity*.
<https://doi.org/10.1007/s43390-026-01400-3>
36. Detection of fraudulent transactions using SAS Viya machine learning algorithms. (2021). *Procedia Computer Science*, 190, 204–209.
37. Deep learning for detecting financial statement fraud. (2020). *Decision Support Systems*, 139, Article 113421.
38. Using machine learning meta-classifiers to detect financial frauds. (2022). *Finance Research Letters*, 48, Article 102915.
39. Detecting money laundering transactions with machine learning. (2020). *Journal of Money Laundering Control*, 23(1), 173–186.
40. Mahadevan, S. (2026). Governed Serverless Automation Ecosystem for AI-Driven Enterprise Integration and Sustainable Cloud Operations. *International Journal of Computer Information Systems and Industrial Management Applications*, 18(16s), 1561-1570.
41. Anti-money laundering supervision by intelligent algorithm. (2023). *Computers & Security*, 132, Article 103344.
42. Review of artificial intelligence-based applications for money laundering detection. (2025). *Intelligent Systems with Applications*, Article 200572.