A Comprehensive Review of Federated Learning Architectures for Insider Threat Detection in Distributed SQL-Based Enterprise Environments

Onuh Matthew Ijiga[1]; Nonso Okika[2]; Semirat Abidemi Balogun[3]; Lawrence Anebi Enyejo[4]; Ogboji James Agbo[5]

[1] Department of Physics Joseph Sarwan Tarka University, Makurdi, Benue State, Nigeria.
[2]Network Planning Analyst, University of Michigan, USA.
[3]Department of Information Science, North Carolina Central University, Durham North Carolina, USA.
[4]Department of Telecommunications, Enforcement Ancillary and Maintenance,
National Broadcasting Commission Headquarters, Aso-Villa, Abuja, Nigeria.
[5]School of Engineering and the Built Environment, Birmingham City University, United Kingdom

Publication Date: 2025/07/12

Abstract: Insider threats remain one of the most challenging cybersecurity concerns for enterprise environments, particularly in distributed systems where sensitive data is stored and processed using SQL-based infrastructures. Conventional centralized detection methods often fail to scale securely across multi-tenant architectures, leading to privacy violations, delayed response times, and limited contextual awareness. This review explores the integration of federated learning (FL) frameworks for insider threat detection in SQL-based distributed enterprise settings. It evaluates the effectiveness of FL in maintaining data locality while training shared threat models collaboratively, thereby mitigating data exfiltration risks and privacy breaches. We analyze existing federated learning architectures—cross-device, cross-silo, and hierarchical FL—focusing on their suitability, scalability, security guarantees, and resource constraints in enterprise-grade SQL ecosystems. Furthermore, the paper identifies challenges related to data heterogeneity, model poisoning, latency, and differential privacy enforcement, and discusses emerging solutions such as blockchain integration and secure aggregation protocols. The study provides critical insights and design considerations for deploying privacy-preserving, decentralized threat detection systems in real-world enterprise contexts.

Keywords: *Federated Learning*, *Insider Threat Detection*, *Distributed SQL Databases*, *Enterprise Security*, *Privacy-Preserving Machine Learning.*

How to Cite: Onuh Matthew Ijiga; Nonso Okika; Semirat Abidemi Balogun;Lawrence Anebi Enyejo; Ogboji James Agbo (2025) A Comprehensive Review of Federated Learning Architectures for Insider Threat Detection in Distributed SQL-Based Enterprise Environments. *International Journal of Innovative Science and Research Technology*, 10(7), 536-550. https://doi.org/10.38124/ijisrt/25jul392

I. INTRODUCTION

- *Background on Insider Threats in SQL-Based Enterprises*

Insider threats continue to pose critical risks in SQL-based enterprise environments, primarily due to the trusted access such insiders inherently possess. SQL databases, which underpin critical operations across finance, healthcare, and defense, are attractive targets for internal exploitation. The threat landscape includes disgruntled employees, negligent users, and individuals coerced or compromised by external actors. Greitzer et al. (2013) emphasize the psychological and behavioral precursors of insider attacks, noting that textual and behavioral patterns can serve as predictive indicators when correlated with SQL audit trails. These enterprise systems, despite their robust query structure, are vulnerable to insider-led data exfiltration and unauthorized access.

Moreover, Brdiczka et al. (2012) assert that traditional anomaly detection methods fall short in SQL-based infrastructures due to their limited context-awareness and lack of temporal correlation among user actions. Graph-based models capturing relational behavior between users and queries show promise but require extensive domain tuning. In tandem, Salem et al. (2008) document that SQL logs, often treated as static artifacts, are underutilized in real-time insider threat modeling due to the sheer complexity and volume of relational data involved. These insights underscore the demand for adaptive, context-sensitive frameworks that can securely analyze distributed SQL activity while preserving operational continuity.

➢ *Limitations of Centralized Threat Detection Systems*

Centralized threat detection systems, though foundational in cybersecurity, face several critical limitations when applied to distributed SQL-based enterprises. Chief among these is their inherent dependency on aggregating sensitive data in a central repository, creating a single point of failure and an attractive attack vector. Cardenas et al. (2008) highlight that centralized systems increase system latency and are difficult to scale efficiently, especially in dynamic, real-time enterprise environments with high query throughput.

Sommer and Paxson (2010) further critique the brittleness of centralized machine learning-based intrusion detection systems, noting their vulnerability to both evasion and adversarial learning. These systems often lack the contextual awareness needed to interpret complex user interactions with SQL schemas and stored procedures, limiting their capacity to differentiate between legitimate and malicious activity. Additionally, Garfinkel (2014) emphasizes regulatory challenges, such as HIPAA and GDPR, which restrict centralized storage of personally identifiable information (PII), further complicating implementation across multijurisdictional SQL infrastructures.

These limitations necessitate a shift toward distributed, privacy-aware frameworks that enable localized threat intelligence while reducing exposure. Consequently, federated learning paradigms that avoid raw data transmission while enabling pattern detection across nodes offer a strategic alternative.

➢ *Motivation for Federated Learning in Enterprise Security*

Federated learning (FL) presents an innovative and scalable approach to securing SQL-based enterprise systems against insider threats, offering a distributed alternative to data centralization. The key strength of FL lies in its ability to collaboratively train machine learning models on local data without transferring it, thereby aligning with data privacy mandates while maintaining predictive power. McMahan et al. (2017) introduced foundational techniques for communication-efficient federated learning, demonstrating its utility in cross-device and enterprise-scale scenarios.

Moreover, the privacy-preserving nature of FL is bolstered through cryptographic enhancements such as secure aggregation and differential privacy, which are essential in SQL systems that handle sensitive access logs and transactional data (Shokri & Shmatikov, 2015). These mechanisms ensure that model updates do not leak user-specific query behaviors while enabling robust pattern learning.

From a system design perspective, Bonawitz et al. (2019) explore the architectural optimizations required for deploying FL at scale, including parameter server orchestration and client selection strategies. These contributions are particularly relevant in enterprise settings where heterogeneous SQL databases and intermittent connectivity demand resilient and adaptive learning mechanisms. The integration of FL thus addresses the twin imperatives of data confidentiality and threat detection efficacy, offering a blueprint for next-generation enterprise defense systems.

➢ *Research Objectives and Scope of Review*

This review aims to systematically examine the application of federated learning architectures in detecting insider threats within distributed SQL-based enterprise environments. The research explores both foundational elements and state-of-the-art advancements to delineate the landscape of FL-enabled security frameworks. A key objective is to assess the architectural suitability of FL variants—cross-device, cross-silo, and hierarchical—in SQL enterprise settings. The review further investigates how federated models handle SQL data workflows, address challenges like schema variability, and preserve privacy through mechanisms such as differential privacy and secure multiparty computation.

Abadi et al. (2016) demonstrate the effectiveness of integrating differential privacy with deep learning models, providing a theoretical foundation for privacy assurance in federated environments. Liu et al. (2020) extend this by proposing secure federated transfer learning frameworks that are highly applicable in SQL systems requiring interoperability across heterogeneous datasets. Furthermore, Yang et al. (2019) provide a comprehensive conceptual model of federated machine learning, covering its lifecycle, system architecture, and potential across enterprise domains.

The scope of this review encompasses architectural taxonomies, communication protocols, privacy-preserving mechanisms, and insider threat taxonomies within FL contexts. By situating the discussion within SQL-driven enterprises, the review contributes to understanding how federated models can be operationalized to mitigate insider risk without compromising data integrity, compliance, or model performance.

➢ *Structure of the Paper*

The structure of this paper is organized to systematically explore the role of federated learning (FL) in addressing insider threats within SQL-based enterprise environments. It begins with an Introduction (Section 1.0), which establishes the background, outlines the limitations of centralized systems, introduces FL as a privacy-preserving solution, and defines the research objectives. Section 2.0 lays the Foundations of Federated Learning, detailing various FL architectures, communication protocols, SQL data workflows, and privacy mechanisms like differential privacy and secure aggregation. Section 3.0 focuses on Insider Threat Modeling, including a taxonomy of threats, the attack surface in federated SQL settings, common TTPs, and the integration of logging and behavioral profiling pipelines. Section 4.0 presents the State-of-the-Art in Federated Learning for Threat Detection, reviewing existing FL-based models, evaluating their performance and scalability, addressing SQL-specific challenges, and exploring emerging enhancements such as blockchain and personalization. The paper concludes in Section 5.0 with Future Directions, discussing open research challenges, the alignment of FL with zero-trust and edge computing, practical adoption strategies for enterprises, and final remarks on the transformative potential of FL in securing distributed data infrastructures.

II. FOUNDATIONS OF FEDERATED LEARNING IN ENTERPRISE CONTEXTS

➢ *Federated Learning Architectures: Cross-Device, Cross-Silo, and Hierarchical*

Federated learning (FL) architectures are commonly categorized into cross-device, cross-silo, and hierarchical configurations, each tailored to specific enterprise deployment scenarios. Cross-device FL is suited for highly distributed settings where millions of clients, such as mobile or IoT devices, intermittently participate in collaborative learning. These models emphasize lightweight updates and robust fault tolerance due to client variability (Kairouz et al., 2021). Conversely, cross-silo FL involves a small number of relatively stable, high-capacity nodes like SQL-based enterprise servers or data centers. This architecture is favorable for enterprise applications due to consistent availability and superior computational resources (Bonawitz et al., 2019). Hierarchical FL extends these paradigms by introducing intermediate aggregators—such as regional data centers—which coordinate updates before global aggregation, thus optimizing communication and preserving scalability (Li et al., 2020).

These architectures support diverse organizational needs, particularly in SQL-based systems, where privacy, bandwidth, and regulatory compliance must be carefully balanced as seen in Table 1. In distributed enterprise environments, choosing the correct FL topology affects not only model performance but also the extent of fault recovery, synchronization overhead, and privacy risk. Each configuration also imposes different communication and synchronization constraints, requiring adaptive orchestration mechanisms. Hence, a nuanced understanding of architectural design is crucial to deploying FL effectively for insider threat detection in structured data environments.

Table 1 Summary of Federated Learning Architectures for SQL-Based Enterprise Environments

Architecture Type	Key Characteristics	Use Case in Enterprise Settings	Challenges and Considerations
Cross-Device FL	Involves millions of low-power, intermittently connected clients (e.g., mobile phones, IoT devices); prioritizes lightweight updates and fault tolerance.	Suitable for large-scale, user-centric data collection with limited connectivity and device resources.	High communication cost, inconsistent participation, limited compute capacity, increased synchronization complexity.
Cross-Silo FL	Connects a small number of stable, high-performance nodes such as enterprise servers or data centers; enables efficient training with structured data.	Ideal for SQL-based enterprise applications requiring consistent uptime, high data quality, and strong security.	Fewer participants may reduce model generalizability; regulatory compliance and inter-organizational trust must be managed.
Hierarchical FL	Utilizes intermediary aggregators (e.g., regional servers) to reduce global communication load; balances scalability and bandwidth efficiency.	Suitable for geographically distributed enterprises with layered infrastructure needing efficient communication and coordination.	Complexity in orchestration; added latency in multi-level aggregation; requires reliable intermediary nodes.
Comparative Advantage	Enables customized design based on enterprise structure, privacy needs, and resource availability.	Tailors deployment strategy to data locality and regulatory constraints in insider threat detection systems.	Selecting the optimal architecture affects fault tolerance, privacy risk, bandwidth usage, and model convergence.

➢ *Key Protocols and Communication Models*

Effective federated learning systems rely on advanced communication protocols and synchronization models to manage the exchange of model updates across distributed nodes. Given the heterogeneity in network bandwidth and data distributions in SQL-based enterprises, protocols must prioritize efficiency without sacrificing model convergence as seen in Fig. 1. Common frameworks include the Federated Averaging (FedAvg) algorithm, which balances update frequency and communication cost (So et al., 2021). Innovations like asynchronous updates and adaptive local training have been introduced to reduce communication bottlenecks and ensure consistent model performance despite stragglers or partial participation (Xu et al., 2020).

Moreover, communication models in FL must address trust and adversarial risk in insider-threat-prone settings. Protocols such as Secure Multi-Party Computation (SMPC) and gradient masking enhance the confidentiality of transmitted updates, making it harder for malicious participants to reverse-engineer sensitive information. Furthermore, agnostic federated learning strategies, which optimize models against worst-case client distributions, are increasingly adopted in adversarial SQL environments (Mohri et al., 2019). These models ensure that learning remains robust and fair, even under highly skewed or compromised data partitions, thereby bolstering the resilience of enterprise security systems.

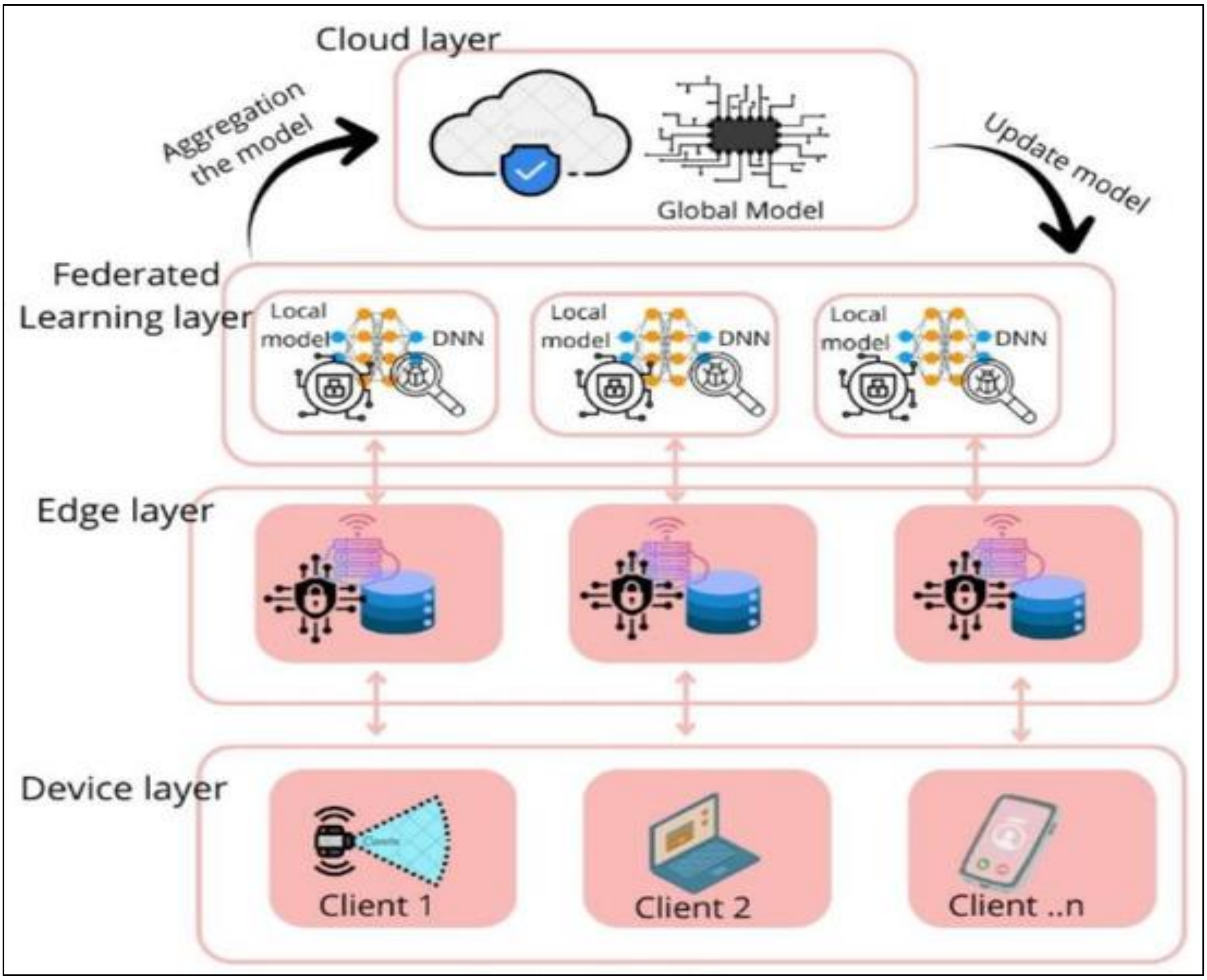

Fig 1 Hierarchical Architecture of a Federated Learning System with Cloud, Edge, and Device Layers (Roba H. et al, 2025)

Figure 1 illustrates the architecture of a Federated Learning System across four hierarchical layers: Device, Edge, Federated Learning, and Cloud. At the Device Layer, multiple clients (e.g., smartphones, laptops, sensors) locally generate and store data. These devices communicate with the Edge Layer, which handles intermediate data processing and local model computation. The processed information is then transmitted to the Federated Learning Layer, where local models and deep neural networks (DNNs) are trained using decentralized data. Each local model contributes to a shared Global Model in the Cloud Layer, where models are aggregated without transmitting raw data, ensuring data privacy. The updated global model is then redistributed to devices, completing the cycle. This architecture supports privacy-preserving, distributed machine learning while minimizing latency and bandwidth usage across the system.

➢ *SQL-Based Data Workflows in Distributed Enterprises*

In enterprise ecosystems, structured data stored in SQL-based systems forms the backbone of operational intelligence. These workflows encompass ETL (extract-transform-load) pipelines, scheduled queries, and real-time transactional processing distributed across departmental nodes. Leveraging federated learning in such settings necessitates seamless integration with existing SQL workflows, particularly when data access is governed by compliance constraints (Thakkar et al., 2021). Federated learning engines must accommodate data schema heterogeneity, maintain ACID (Atomicity, Consistency, Isolation, Durability) properties, and support query rewriting to ensure consistency in training datasets across silos.

Solutions such as federated process mining have demonstrated how enterprise SQL logs can be analyzed for behavioral threat patterns while maintaining data locality (Beekhuizen & De Weerdt, 2022). Similarly, execution engines like FLEX integrate federated optimization with SQL engines as seen in Figure 1, facilitating efficient local computation and reducing server load (Chen et al., 2020). The key to effective implementation lies in aligning model training loops with data manipulation layers, enabling threat detection models to adapt continuously to the evolving structure of enterprise transactions and access control policies.

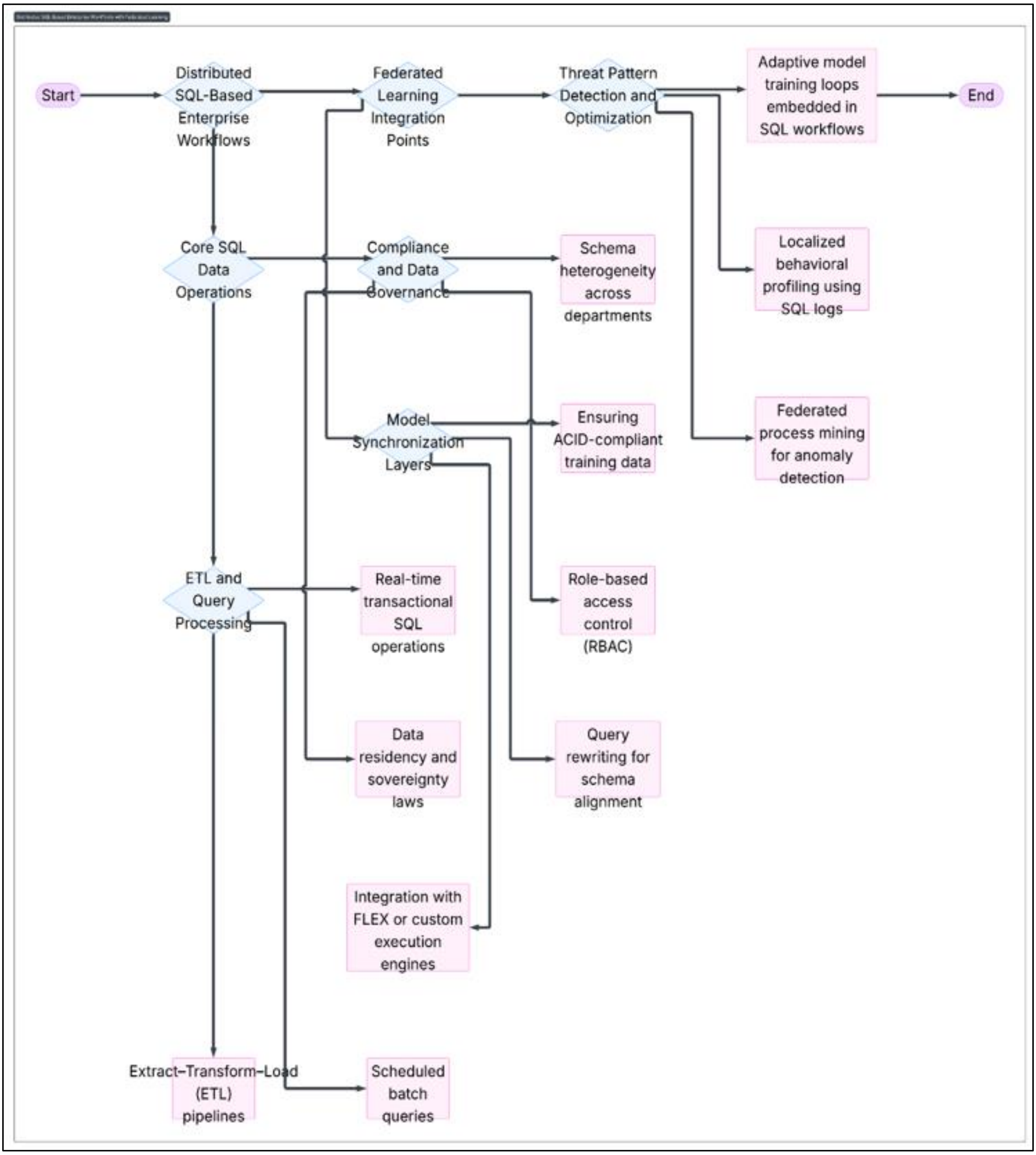
Start
Distributed SQL-Based Enterprise Workflows
Federated Learning Integration Points
Threat Pattern Detection and Optimization
Adaptive model training loops embedded in SQL workflows
End
Core SQL Data Operations
Compliance and Data Governance
Schema heterogeneity across departments
Localized behavioral profiling using SQL logs
Model Synchronization Layers
Ensuring ACID-compliant training data
Federated process mining for anomaly detection
ETL and Query Processing
Real-time transactional SQL operations
Role-based access control (RBAC)
Data residency and sovereignty laws
Query rewriting for schema alignment
Integration with FLEX or custom execution engines
Extract–Transform–Load (ETL) pipelines
Scheduled batch queries

Fig 2 SQL-Based Data Workflows and Federated Learning Integration in Distributed Enterprises

Figure 2 presents a high-level flowchart illustrating how federated learning integrates with distributed SQL-based enterprise workflows to enable secure and privacy-aware threat pattern detection and optimization. It begins with core SQL data operations and ETL and query processing, including extract-transform-load (ETL) pipelines and scheduled batch queries, which feed into real-time transactional SQL operations. These operations are governed by data residency laws, role-based access control (RBAC), and compliance protocols. The figure highlights federated learning integration points, including model synchronization layers and execution engines like FLEX, enabling privacy-preserving model training within SQL ecosystems. Challenges such as schema heterogeneity across departments are addressed through query rewriting and schema alignment to ensure ACID-compliant training data. Ultimately, the pipeline supports localized behavioral profiling, federated process mining, and adaptive training loops using SQL logs. This structured integration ensures scalable, privacy-compliant, and robust insider threat detection embedded directly within the SQL workflow fabric.

➢ *Privacy Mechanisms: Differential Privacy, Secure Aggregation, Homomorphic Encryption*

Privacy-preserving techniques are integral to federated learning, especially when detecting insider threats in sensitive enterprise environments. Differential privacy (DP) introduces calibrated noise to model updates, guaranteeing that the presence or absence of a single data point cannot significantly alter the output, thus mitigating data leakage risk (Geyer et al., 2017). Client-level DP, where noise is applied before data leaves the local SQL node, is particularly suitable for structured environments with compliance mandates such as HIPAA or GDPR.

Secure aggregation further strengthens privacy by enabling the server to compute the sum of model updates without accessing individual contributions. This technique prevents model inversion attacks and ensures confidentiality even in the presence of a semi-honest server (Bonawitz et al., 2017). Complementarily, homomorphic encryption (HE) offers the ability to perform computations directly on encrypted data, ensuring end-to-end confidentiality from training to inference (Acar et al., 2021). Though computationally intensive, HE is becoming more viable due to improvements in algorithmic efficiency and parallelism, making it a strong candidate for SQL-based systems requiring high privacy guarantees.

Collectively, these mechanisms form a multi-layered defense strategy against insider manipulation or accidental disclosure, aligning federated learning protocols with enterprise-grade data protection requirements as seen in Table 2. As insider threat detection often involves sensitive log files and behavioral analytics, implementing such privacy enhancements is essential to maintain both trust and regulatory compliance.

Table 2 Summary of Privacy Mechanisms in Federated Learning for Insider Threat Detection

Privacy Mechanism	Core Principle	Application in SQL-Based Federated Learning	Advantages in Insider Threat Detection
Differential Privacy (DP)	Adds calibrated noise to data or model updates to obscure individual contributions	Client-level DP ensures noise is applied before data leaves the SQL node, maintaining compliance with HIPAA/GDPR	Mitigates risk of data leakage and re-identification while preserving statistical utility
Secure Aggregation	Aggregates model updates so the server sees only the sum, not individual values	Prevents exposure of individual client updates, even if server is semi-honest	Defends against model inversion and enhances confidentiality in collaborative environments
Homomorphic Encryption (HE)	Enables computation on encrypted data without decryption	Ensures confidentiality throughout training and inference; suitable for high-sensitivity enterprise data	Provides end-to-end data protection with rising feasibility due to computational optimizations
Combined Use	Integration of DP, Secure Aggregation, and HE for layered protection	Aligns federated learning workflows with stringent enterprise security requirements	Offers robust protection against both insider manipulation and accidental disclosure

III. INSIDER THREAT MODELING IN SQL-BASED DISTRIBUTED SYSTEMS

➢ *Taxonomy of Insider Threats: Malicious, Negligent, Compromised*

Insider threats in distributed SQL-based enterprise environments are typically categorized into three distinct types: malicious, negligent, and compromised. Malicious insiders act with intent to harm the organization, often motivated by financial gain, ideology, or retaliation. These actors exploit legitimate access to databases, often using complex obfuscation techniques to avoid detection (Nurse et al., 2014). Their actions may include data exfiltration, manipulation of SQL queries, or the installation of backdoors in data pipelines.

Negligent insiders, on the other hand, pose risk through carelessness or poor adherence to security protocols as seen in Table 3. Such threats often stem from insecure SQL script sharing, weak password practices, or accidental exposure of data due to misconfigured permissions. These behaviors, while not ill-intentioned, significantly widen the attack surface by creating vulnerabilities that external adversaries can exploit (Azaria et al., 2018).

The third class involves compromised insiders, where external threat actors hijack legitimate user credentials via phishing, malware, or social engineering. These attackers then act as insiders, bypassing conventional perimeter defenses. SQL-based systems are particularly vulnerable in this context due to the prevalence of credential-based access controls and insufficient behavioral anomaly detection tools (Greitzer et al., 2012). Understanding these categories enables a targeted response strategy for federated systems, where role-specific behavioral baselines must be established to distinguish between legitimate anomalies and harmful behaviors.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25jul392

Table 3 Taxonomy of Insider Threats in Distributed SQL-Based Enterprise Environments

Threat Type	Key Characteristics	Common Attack Vectors	Security Implications
Malicious Insider	Acts with intent to cause harm; motivated by revenge, financial gain, or ideology	Data exfiltration, malicious SQL injections, backdoor creation, privilege misuse	High-risk; often involves stealthy, deliberate bypassing of detection systems
Negligent Insider	Unintentionally compromises security through poor practices or ignorance	Weak passwords, misconfigured permissions, insecure SQL script sharing	Creates vulnerabilities and increases attack surface for external exploitation
Compromised Insider	Legitimate user credentials hijacked by external threat actors	Phishing, malware, social engineering, credential theft	Enables unauthorized access within the perimeter, often difficult to detect in real time
Mitigation Focus	Role-based behavior monitoring; real-time anomaly detection; credential management	Threat-specific response frameworks for federated SQL systems	Enhances targeted defenses and reduces the effectiveness of each threat type

- *Attack Surface in Federated SQL Environments*

The federated learning paradigm inherently shifts the locus of data processing to edge or siloed environments, reducing centralized data exposure but introducing new vectors of attack. In federated SQL-based enterprise settings, each participating node or client stores sensitive query logs and user behavior metrics locally, creating decentralized micro-attack surfaces. Adversaries may attempt to reverse-engineer global models by observing updates or leveraging inference attacks (Shokri & Shmatikov, 2015). This risk is amplified in SQL-driven contexts where structured data patterns can reveal semantic cues about user behavior.

Collaborative learning systems such as federated learning are also susceptible to generative adversarial network (GAN)-based attacks. Malicious clients can inject poisoned updates that mimic legitimate SQL behaviors, thereby skewing model performance or extracting information from peer updates (Hitaj et al., 2017). These vulnerabilities increase the likelihood of model inversion attacks, where SQL queries or sensitive records can be reconstructed from the gradients or weights shared during federated updates.

In response, hybrid privacy frameworks have been proposed, integrating secure aggregation and local differential privacy to mask individual SQL data contributions as seen in fig.2. However, these mechanisms must be tuned to balance utility and privacy, especially in structured database environments with high data interdependence (Truex et al., 2019). Thus, the federated SQL environment presents a unique tradeoff between distributed privacy and expanded model interaction risk.

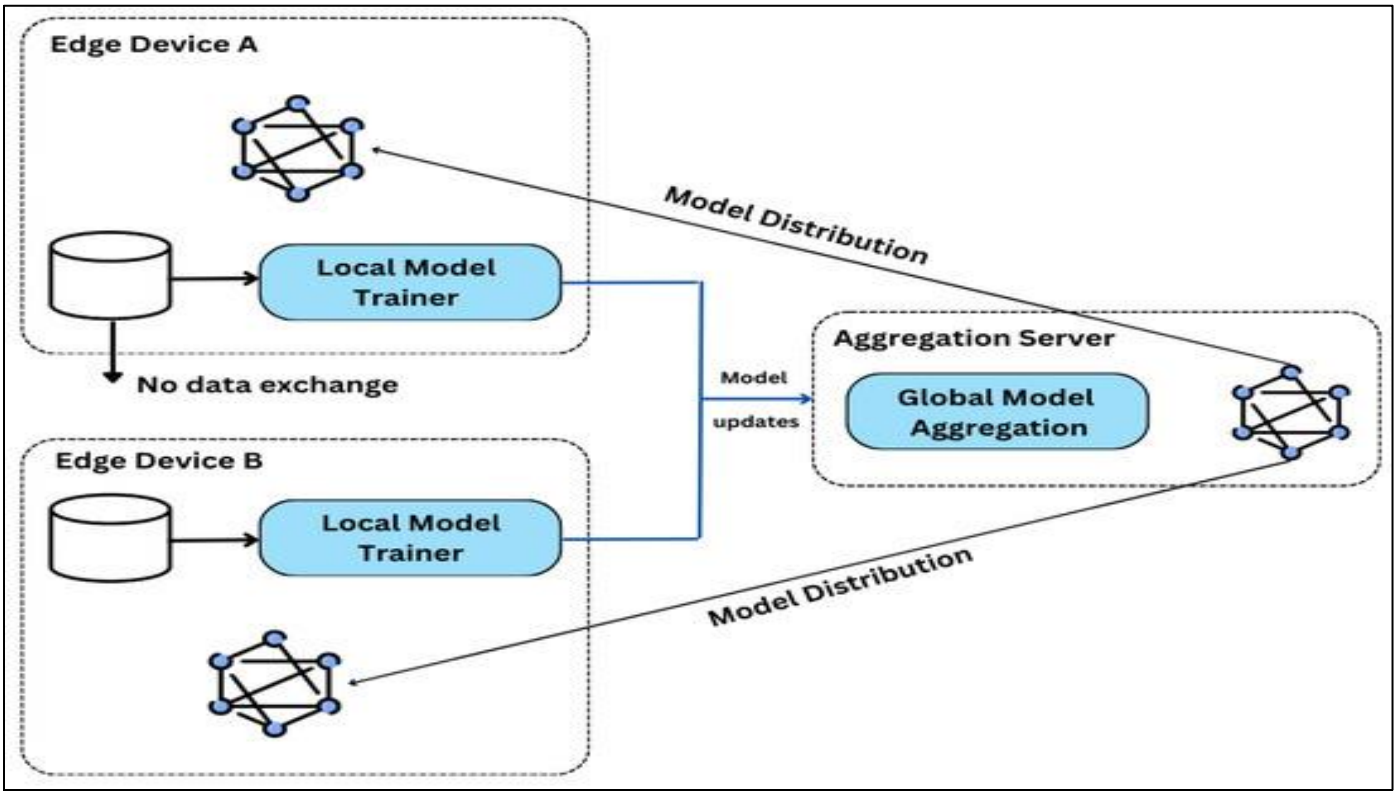

Fig 3 Federated Learning Workflow with Local Model Training and Centralized Global Model Aggregation (Latifa Albshaier, 2025)

Figure 3 illustrates a Federated Learning Framework involving two edge devices (A and B) and a central aggregation server. Each Edge Device possesses local data and independently trains a Local Model using a Local Model Trainer, without exchanging raw data with other devices—preserving privacy. Instead of sharing data, both devices send only model updates to the Aggregation Server, which performs Global Model Aggregation by combining updates from multiple devices into a refined Global Model. This aggregated model is then redistributed back to the edge devices for continued training or inference. The system enables collaborative learning across decentralized devices while ensuring data security and minimizing communication overhead.

➢ *Common Tactics, Techniques, and Procedures (TTPs)*

Insider threats within SQL-based environments often follow recognizable tactics, techniques, and procedures (TTPs), many of which are captured within structured taxonomies like MITRE's ATT&CK framework. Common initial access vectors include credential dumping, phishing, and exploitation of SQL-based web applications. Once inside, insiders or compromised actors typically escalate privileges, disable audit logging, or create rogue administrative accounts (MITRE Corporation, 2019).

Privilege abuse is particularly relevant in federated settings, where client nodes may not uniformly enforce access control or behavioral monitoring policies. Attackers may manipulate or insert malicious SQL commands into federated updates to corrupt shared models, insert data backdoors, or exfiltrate metadata about peer nodes (Salem et al., 2008). Data collection and exfiltration TTPs in federated SQL systems include SQL injection chaining, lateral movement across federated nodes, and command-line tooling to bypass encryption layers as seen in Figure 4.

Notably, social engineering plays a significant role in enabling these attacks, as insiders often rely on personal trust relationships to gain access or override policy. The use of encrypted tunnels, steganographic payloads, and fileless malware further complicate detection (Cole & Ring, 2006). As these TTPs evolve, effective detection demands real-time behavioral profiling, federated audit trails, and continuous integration of adversarial intelligence.

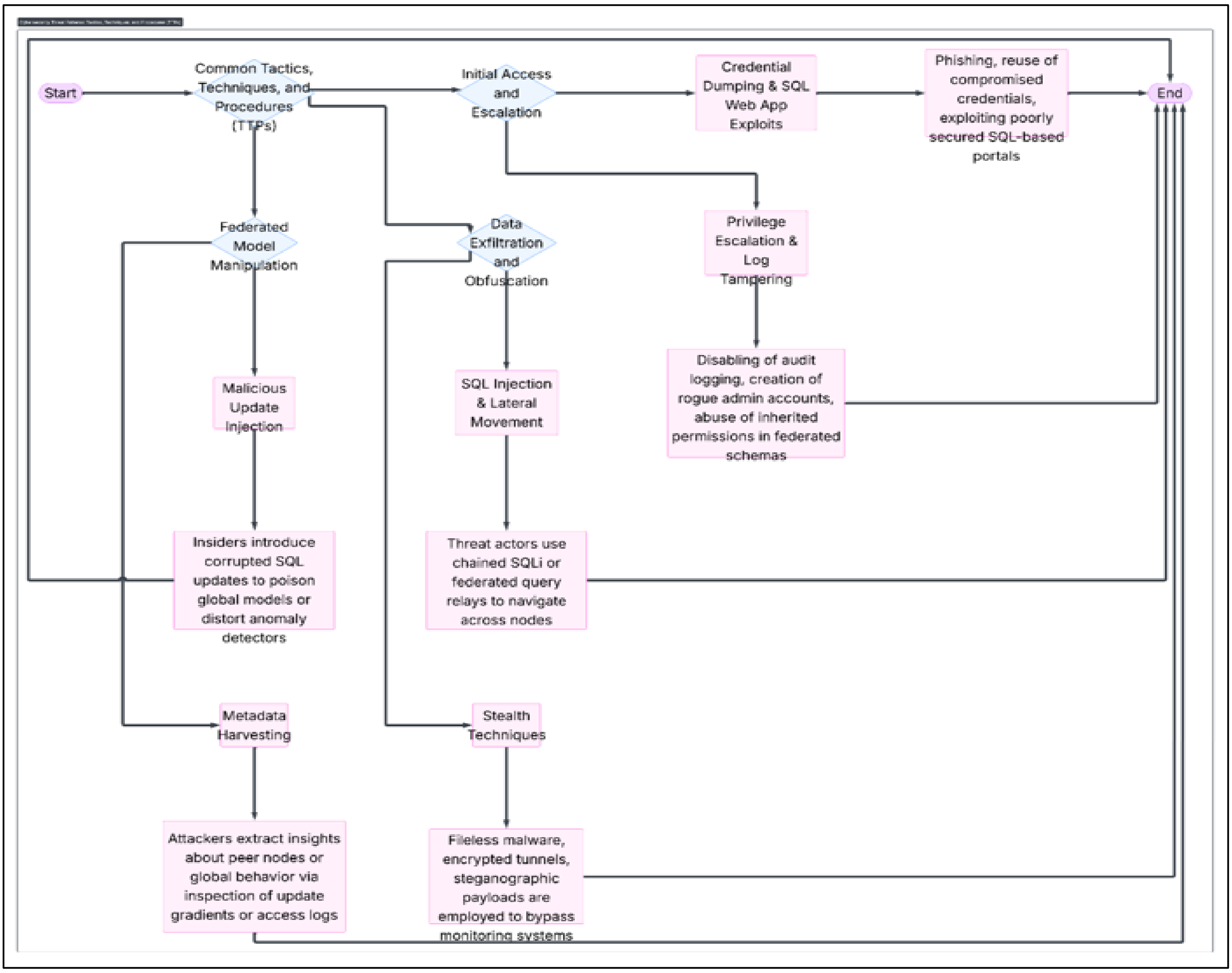
Start
Common Tactics, Techniques, and Procedures (TTPs)
Initial Access and Escalation
Credential Dumping & SQL Web App Exploits
Phishing, reuse of compromised credentials, exploiting poorly secured SQL-based portals
End
Federated Model Manipulation
Data Exfiltration and Obfuscation
Privilege Escalation & Log Tampering
Disabling of audit logging, creation of rogue admin accounts, abuse of inherited permissions in federated schemas
Malicious Update Injection
SQL Injection & Lateral Movement
Insiders introduce corrupted SQL updates to poison global models or distort anomaly detectors
Threat actors use chained SQLi or federated query relays to navigate across nodes
Metadata Harvesting
Stealth Techniques
Attackers extract insights about peer nodes or global behavior via inspection of update gradients or access logs
Fileless malware, encrypted tunnels, steganographic payloads are employed to bypass monitoring systems

Fig 4 Insider Threat TTPs in Federated SQL-Based Environments

Figure 4 illustrates a structured flowchart of common tactics, techniques, and procedures (TTPs) used in insider threat campaigns targeting federated SQL-based enterprise systems. The attack sequence begins with initial access and escalation, often via credential dumping or SQL web application exploits, which lead to privilege escalation and log tampering—including actions like disabling audit logs or creating rogue admin accounts. These compromised credentials and privileges enable attackers to perform data exfiltration and obfuscation through methods such as SQL injection (SQLi) and lateral movement across federated nodes. Concurrently, federated model manipulation may occur, where insiders inject malicious updates to corrupt global models or mislead anomaly detectors. This can escalate to metadata harvesting, where attackers gain insights from update gradients or access logs, or to stealth techniques like fileless malware and encrypted tunnels to evade detection. The diagram highlights the cyclic and interconnected nature of these TTPs, culminating in successful data theft or system compromise through phishing, reuse of compromised credentials, or exploitation of weak SQL-based portals, ultimately feeding back into the threat lifecycle.

- ➢ *Logging, Behavioral Profiling, and Threat Detection Pipelines*

In distributed SQL environments leveraging federated learning, threat detection hinges on the robustness of logging and behavioral profiling mechanisms. Logging serves as the foundation of forensic analysis and real-time detection, yet inconsistent log schemas and formats across federated nodes can hinder correlation efforts. Structured SQL query logs, system event records, and access patterns must be normalized and aggregated securely for effective pipeline implementation (Kent & Souppaya, 2006).

Behavioral profiling models, particularly those based on graph analysis and deep learning, are gaining traction. These models examine user-session graphs, query frequency, temporal patterns, and access deviations to infer anomalies. Eberle and Holder (2009) demonstrated that graph-based anomaly detection effectively identifies abnormal query flows indicative of insider misuse. Similarly, unsupervised deep learning techniques, such as autoencoders and recurrent neural networks, have been employed to analyze high-dimensional SQL access patterns in real time (Tuor et al., 2017).

Modern pipelines combine data ingestion, stream processing, anomaly scoring, and federated feedback mechanisms to maintain adaptive learning. However, federated architectures complicate centralized logging and profiling, requiring decentralized intelligence and secure synchronization protocols. Integrating edge analytics with audit systems allows localized detection while preserving confidentiality, aligning well with zero-trust principles and privacy-by-design standards.

Table 4 Key Components of Logging, Behavioral Profiling, and Threat Detection Pipelines in Federated SQL Environments

Component	Description	Challenges	Solutions/Best Practices
Logging Infrastructure	Captures SQL queries, system events, and access logs across federated nodes	Inconsistent formats and decentralized log locations	Standardize log schemas, enable secure aggregation, and employ schema-aware parsers
Behavioral Profiling	Uses graph models and deep learning to model normal vs. anomalous behaviors	High dimensionality, lack of labeled data, and temporal variability	Apply autoencoders, RNNs, and graph analytics to learn behavior baselines and detect anomalies
Detection Pipelines	Streamline data ingestion, anomaly scoring, and adaptive response mechanisms	Centralized coordination is difficult in federated setups	Use decentralized detection models and integrate secure federated feedback systems
Privacy & Security Sync	Ensures confidential, policy-compliant anomaly detection under federated control	Risk of data leakage and weak synchronization across nodes	Adopt edge analytics, zero-trust architecture, and privacy-by-design protocols

IV. FEDERATED LEARNING FOR INSIDER THREAT DETECTION: STATE-OF-THE-ART

- ➢ *Review of FL-Based Intrusion and Threat Detection Models*

Federated learning (FL) has become a transformative approach to intrusion detection, particularly within environments where data centralization is infeasible due to privacy, policy, or infrastructure constraints. In SQL-based enterprises, insider threat detection often involves analyzing access logs, transactional anomalies, and behavioral signatures—tasks well-suited to FL's decentralized paradigm. Liu et al. (2020) proposed the Federated Forest model, allowing decision trees to be trained collaboratively while preserving data locality, offering interpretability essential in security audits. Similarly, Li et al. (2020) emphasized how FL accommodates diverse enterprise nodes and heterogeneous data, effectively decentralizing security analytics without exposing sensitive logs.

The core contribution of FL to cybersecurity lies in its capacity to harness threat intelligence from distributed sources while ensuring compliance with data governance mandates. Chen et al. (2021) highlighted various FL applications for detecting advanced persistent threats (APTs), lateral movements, and credential misuse within SQL infrastructures. These models are often enriched with differential privacy layers and secure aggregation techniques to mitigate risks associated with gradient inversion attacks or model poisoning. Collectively, FL-based threat detection models offer organizations scalable, policy-compliant tools to counter insider threats without breaching user confidentiality or regulatory standards.

- *Comparative Evaluation of Model Performance, Scalability, and Privacy Guarantees*

Evaluating federated learning models requires a multidimensional lens encompassing accuracy, latency, bandwidth efficiency, and privacy. Kairouz et al. (2021) presented a comprehensive benchmark for FL frameworks and identified communication overhead as a principal scalability bottleneck, particularly relevant for SQL-based systems operating over WANs. Despite these challenges, models like FedAvg have shown promise in maintaining predictive performance across highly non-IID data distributions commonly found in enterprise logs.

Aledhari et al. (2020) further explored how protocol-level improvements and asynchronous updates could drastically reduce training time without sacrificing accuracy or privacy. Notably, trade-offs persist: enhanced privacy mechanisms such as differential privacy often degrade model fidelity, underscoring the balance needed between security and utility. In contrast, Bonawitz et al. (2019) showcased Google's FL infrastructure operating at scale with tens of millions of clients, providing evidence that high-performance threat detection is viable under optimized orchestration and compression strategies.

Privacy guarantees are central to federated security analytics as seen in Table 5. Encryption methods and privacy-aware protocols have proven indispensable in shielding both client-side datasets and model parameters. FL’s strength, therefore, lies not just in its distributed design but in the evolving sophistication of its privacy-preserving toolkits, making it increasingly relevant for insider threat detection in sensitive, data-intensive SQL environments.

Table 5 Comparative Evaluation of Federated Learning Models on Performance, Scalability, and Privacy Guarantees

Evaluation Dimension	Observations	Strengths	Limitations
Model Performance	Models such as FedAvg maintain predictive accuracy on non-IID enterprise data.	Robust generalization across heterogeneous data distributions.	Slight loss in fidelity with privacy-enhancing techniques.
Scalability	Communication overhead remains a bottleneck, especially in SQL systems over WANs.	Asynchronous updates and orchestration improvements can mitigate latency.	High client participation leads to increased system complexity and bandwidth demands.
Training Efficiency	Protocol optimizations reduce training time without compromising accuracy.	Faster convergence with selective update strategies.	Requires careful scheduling and system-level tuning.
Privacy Guarantees	Differential privacy and encryption are used to protect both model parameters and raw data during training.	Strong data confidentiality and regulatory compliance.	May introduce noise that degrades accuracy; complex implementation.

- *Challenges in SQL-Driven Environments: Data Imbalance, Schema Variability*

SQL-based enterprise systems inherently grapple with data heterogeneity and imbalance, both of which complicate federated learning deployments. Insider threat detection relies on anomaly detection over structured queries, logs, and behavioral indicators, yet these data sources are frequently non-IID and exhibit skewed distributions. Wang et al. (2022) identified non-IID data as a fundamental limitation, noting that model convergence and generalizability suffer in federated scenarios where certain participants dominate with high-volume data or irregular access patterns.

Moreover, schema variability between enterprise branches presents an architectural barrier. Zhang and Wang (2021) proposed adaptive schema-matching protocols that help standardize SQL data pipelines, though full interoperability remains a research challenge. Lin et al. (2020) offered a meta-architecture for federated SQL systems that addresses data inconsistency through layered abstraction, yet performance overheads remain significant.

Handling missing features, inconsistent types, and poorly aligned indexes adds further complexity to model training. These challenges reduce the effectiveness of vanilla FL algorithms and necessitate robust pre-processing strategies, model adaptation layers, or federated meta-learning approaches. Thus, achieving robust insider threat detection via FL in SQL environments requires not only model innovation but deep integration with data engineering workflows to overcome schema fragmentation and imbalanced behavioral traces.

- *Emerging Solutions: Blockchain, Transfer Learning, and FL Personalization*

Emerging technologies are augmenting federated learning with advanced capabilities to address its current limitations in insider threat detection. Blockchain has gained attention as a decentralized coordination and audit layer for FL, ensuring model update traceability and tamper resistance. Ononiwu et al. (2023) demonstrated the efficacy of blockchain-FL hybrid systems in healthcare data environments, a structure adaptable to SQL-based enterprises to enhance auditability and trust in security models.

Transfer learning is also pivotal in mitigating the cold-start problem and improving detection across low-data nodes. Chen et al. (2022) proposed federated transfer learning (FTL) frameworks that share high-level model knowledge across SQL servers while preserving privacy. These approaches are particularly effective when one department's data can bootstrap another's model in similar operational domains.

Meanwhile, FL personalization addresses user and schema heterogeneity by tailoring models to individual enterprise contexts without global overfitting. Fallah et al. (2020) introduced a meta-learning-based personalization algorithm that achieved high accuracy on distributed behavioral data without requiring unified schemas. These hybrid and modular approaches are driving the next phase of federated learning, enhancing its adaptability, responsiveness, and effectiveness in detecting evolving insider threat vectors within dynamic SQL-based ecosystems as shown in Figure 5.

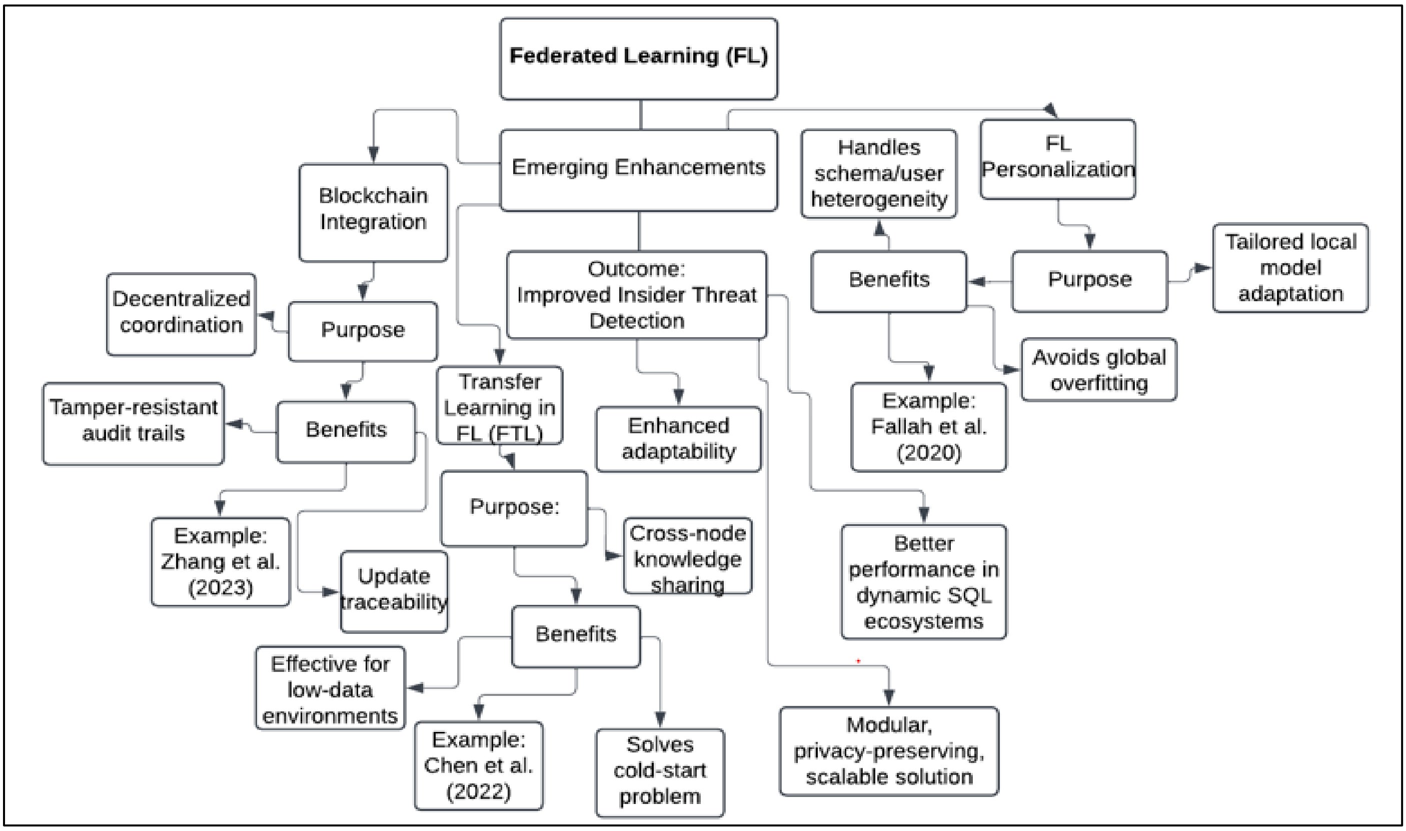
Federated Learning (FL)
Emerging Enhancements
Blockchain Integration
Purpose
Decentralized coordination
Benefits
Tamper-resistant audit trails
Example: Zhang et al. (2023)
Update traceability
Outcome: Improved Insider Threat Detection
Transfer Learning in FL (FTL)
Purpose:
Cross-node knowledge sharing
Benefits
Effective for low-data environments
Example: Chen et al. (2022)
Solves cold-start problem
Enhanced adaptability
Handles schema/user heterogeneity
FL Personalization
Purpose
Tailored local model adaptation
Benefits
Avoids global overfitting
Example: Fallah et al. (2020)
Better performance in dynamic SQL ecosystems
Modular, privacy-preserving, scalable solution

Fig 5 A Block Diagram Showing Enhancing Federated Learning for Insider Threat Detection with Blockchain, Transfer Learning, and Personalization.

Figure 5 illustrates how federated learning (FL), a privacy-preserving distributed machine learning paradigm, is being augmented to improve insider threat detection in SQL-based environments. At the core, FL facilitates secure model training across decentralized nodes without data centralization. This core functionality is enhanced through three emerging technologies: blockchain integration ensures traceability and tamper-proof audit trails; transfer learning addresses the cold-start problem by allowing high-level model knowledge to be shared across low-data nodes; and FL personalization adapts models to heterogeneous users and schemas, reducing global overfitting. These advanced layers converge to produce a modular, scalable, and context-sensitive threat detection system, optimized for dynamic enterprise databases.

V. FUTURE DIRECTIONS AND CONCLUSION

➤ *Open Research Challenges and Future Trends*

Despite the promising potential of federated learning (FL) for insider threat detection in distributed SQL-based enterprise environments, several open challenges remain. One key issue is the handling of heterogeneous data schemas across different SQL nodes, which often results in model inconsistencies and degraded accuracy. Additionally, the imbalance and sparsity of insider threat indicators present difficulties in model generalization and convergence. Another pressing challenge is ensuring real-time threat detection without sacrificing performance, particularly in environments with limited computational resources. Moreover, the interpretability of federated models is still underdeveloped, which hinders human oversight and trust in automated decisions. Future trends are expected to focus on adaptive federated optimization strategies that dynamically adjust to enterprise-level variations. Advancements in automated feature engineering and domain adaptation techniques will also play a crucial role in making FL models more robust across diverse environments. Furthermore, the development of hybrid models that combine FL with transfer learning and reinforcement learning will likely drive future improvements in threat detection capabilities.

➤ *Integration with Zero-Trust Architectures and Edge Intelligence*

The integration of federated learning into zero-trust architectures represents a logical evolution in enterprise security, emphasizing strict verification over implicit trust. Zero-trust principles require continuous validation of user

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25jul392

identity, device posture, and behavioral context before granting access. FL complements this model by enabling collaborative intelligence across endpoints without centralizing sensitive data, thus aligning with the principle of least privilege. When combined with edge intelligence, FL supports localized decision-making at or near data sources, reducing latency and bandwidth usage. This synergy allows enterprises to deploy lightweight, context-aware models directly on edge devices such as firewalls, endpoint protection tools, and database monitors. These models can continuously evaluate anomaly patterns and user behavior while preserving privacy. The distributed nature of this setup also offers resilience against single points of failure and targeted attacks. As enterprises transition to hybrid cloud and multi-edge networks, FL embedded within zero-trust frameworks and supported by edge computing will form the backbone of scalable, intelligent, and secure insider threat detection.

➢ *Recommendations for Enterprise Adoption*

To adopt federated learning effectively for insider threat detection, enterprises should begin with a clear identification of threat models and risk scenarios specific to their SQL-based operations. Establishing a modular data architecture that supports federated model training across heterogeneous environments is essential. Organizations must also invest in building secure communication protocols for model parameter exchange, ensuring that updates are encrypted and authenticated end-to-end. Prior to implementation, a pilot deployment involving limited SQL nodes can help evaluate performance under controlled settings and inform scaling strategies. Additionally, a multidisciplinary team involving data scientists, cybersecurity experts, and database administrators should collaborate to define detection objectives and interpretability requirements. Training these models requires a consistent influx of labeled behavioral data, so enterprises should consider integrating synthetic data generation tools to address class imbalance. Finally, it is important to embed governance mechanisms such as model auditing, role-based access control, and continuous performance monitoring to maintain compliance and transparency throughout the deployment lifecycle.

➢ *Concluding Remarks*

Federated learning offers a transformative approach to addressing insider threats within distributed SQL-based enterprise environments. By enabling decentralized model training, FL mitigates the risks associated with centralized data aggregation while preserving organizational privacy and data sovereignty. The adaptability of this paradigm allows for seamless deployment across diverse nodes, making it suitable for the dynamic and heterogeneous nature of modern enterprise systems. When combined with advanced privacy-preserving techniques and integrated into zero-trust frameworks, FL has the potential to redefine enterprise security architectures. However, to fully realize these benefits, organizations must overcome existing challenges related to interoperability, model interpretability, and real-time responsiveness. As the field matures, future innovations will likely focus on hybrid approaches that enhance model adaptability and reduce operational complexity. Ultimately, the success of federated learning in enterprise security depends on a strategic blend of technological investment, policy alignment, and cross-functional collaboration to safeguard data integrity and organizational resilience in an increasingly threat-prone digital landscape.

REFERENCES

[1]. Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep learning with differential privacy. *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, 308–318. https://doi.org/10.1145/2976749.2978318

[2]. Abiola, O. B. & Ijiga, M. O. (2025), Implementing Dynamic Confidential Computing for Continuous Cloud Security Posture Monitoring to Develop a Zero Trust-Based Threat Mitigation Model. International Journal of Innovative Science and Research Technology (IJISRT) IJISRT25MAY587, 69-83. DOI: 10.38124/ijisrt/25may587.https://www.ijisrt.com/implementing-dynamic-confidential-computing-for-continuous-cloud-security-posture-monitoring-to-develop-a-zero-trustbased-threat-mitigation-model

[3]. Acar, A., Aksu, H., Uluagac, A. S., & Conti, M. (2021). A survey on homomorphic encryption schemes: Theory and implementation. *ACM Computing Surveys (CSUR)*, 54(6), 1–35. https://doi.org/10.1145/3431920

[4]. Aledhari, M., Razzak, R., Hussain, F. K., & Alazab, M. (2020). Federated learning: A survey on enabling technologies, protocols, and applications. *IEEE Access*, 8, 140699–140725. https://doi.org/10.1109/ACCESS.2020.3013541

[5]. Atalor, S. I. (2019). Federated Learning Architectures for Predicting Adverse Drug Events in Oncology Without Compromising Patient Privacy *ICONIC RESEARCH AND ENGINEERING JOURNALS* JUN 2019 | IRE Journals | Volume 2 Issue 12 | ISSN: 2456-8880

[6]. Atalor, S. I. (2022). Data-Driven Cheminformatics Models for Predicting Bioactivity of Natural Compounds in Oncology. International Journal of Scientific Research and Modern Technology, 1(1), 65–76. https://doi.org/10.38124/ijsrmt.v1i1.496

[7]. Atalor, S. I., Ijiga, O. M., & Enyejo, J. O. (2023). Harnessing Quantum Molecular Simulation for Accelerated Cancer Drug Screening. *International Journal of Scientific Research and Modern Technology*, *2*(1), 1–18. https://doi.org/10.38124/ijsrmt.v2i1.502

[8]. Ayoola, V. B., Ugoaghalam, U. J., Idoko P. I, Ijiga, O. M & Olola, T. M. (2024). Effectiveness of social engineering awareness training in mitigating spear phishing risks in financial institutions from a cybersecurity perspective. *Global Journal of Engineering and Technology Advances,* 2024, 20(03), 094–117. https://gjeta.com/content/effectiveness-social-engineering-awareness-training-mitigating-spear-phishing-risks

[9]. Azaria, A., Richardson, A., & Brooks, J. (2018). Insider threat detection using supervised learning algorithms on human-behavior data. *Journal of Cybersecurity and Privacy*, 1(1), 18–35. https://doi.org/10.3390/jcp1010002

[10]. Beekhuizen, J., & De Weerdt, J. (2022). Federated process mining in SQL environments: Architecture and use cases. *Information Systems*, 108, 102054. https://doi.org/10.1016/j.is.2022.102054

[11]. Bonawitz, K., Eichner, H., Grieskamp, W., Huba, D., Ingerman, A., Ivanov, V., ... & Ramage, D. (2019). Towards federated learning at scale: System design. *Proceedings of the 2nd SysML Conference*. https://arxiv.org/abs/1902.01046

[12]. Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H. B., Patel, S., ... & Seth, K. (2017). Practical secure aggregation for privacy-preserving machine learning. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security* (pp. 1175–1191). https://doi.org/10.1145/3133956.3133982

[13]. Brdiczka, O., Liu, J., Price, B., Shen, J., Patil, A., Chow, R., ... & Bart, E. (2012). Proactive insider threat detection through graph learning and psychological context. *Proceedings of the IEEE Symposium on Security and Privacy Workshops*, 142–149. https://doi.org/10.1109/SPW.2012.28

[14]. Cardenas, A. A., Amin, S., & Sastry, S. (2008). Research challenges for the security of control systems. *Proceedings of the 3rd USENIX Workshop on Hot Topics in Security (HotSec)*. https://www.usenix.org/legacy/event/hotsec08/tech/full_papers/cardenas/cardenas.pdf

[15]. Chen, M., Ma, Y., Hao, Y., & Wang, Y. (2022). Federated transfer learning for secure enterprise applications. *IEEE Transactions on Services Computing*, 15(4), 2035–2047. https://doi.org/10.1109/TSC.2020.3007470

[16]. Chen, Y., Sun, X., Zhang, H., & Guo, L. (2020). FLEX: An efficient federated learning execution system for data-intensive SQL queries. *IEEE Transactions on Knowledge and Data Engineering*, 33(7), 2397–2411. https://doi.org/10.1109/TKDE.2020.2988394

[17]. Chen, Y., Zhang, X., Liu, H., & Wang, S. (2021). A survey on federated learning for cyber security: Concepts, applications, and challenges. *IEEE Transactions on Industrial Informatics*, 17(9), 6230–6245. https://doi.org/10.1109/TII.2021.3067490

[18]. Cole, E., & Ring, S. (2006). *Insider Threat: Protecting the Enterprise from Sabotage, Spying, and Theft*. Syngress. https://scholar.google.com/scholar_lookup?title=Insider%20Threat%3A%20Protecting%20the%20Enterprise

[19]. Eberle, W., & Holder, L. B. (2009). Insider threat detection using graph-based approaches. *Journal of Applied Security Research*, 4(1), 32–81. https://doi.org/10.1080/19361610802685719

[20]. Eguagie, M. O., Idoko, I. P., Ijiga, O. M., Enyejo, L. A., Okafor, F. C. & Onwusi, C. N. (2025). Geochemical and Mineralogical Characteristics of Deep Porphyry Systems: Implications for Exploration Using ASTER. *International Journal of Scientific Research in Civil Engineering.* 2025 | IJSRCE | Volume 9 | Issue 1 | ISSN : 2456-6667. doi : https://doi.org/10.32628/IJSRCE25911

[21]. Fallah, A., Mokhtari, A., & Ozdaglar, A. (2020). Personalized federated learning: A meta-learning approach. *Advances in Neural Information Processing Systems (NeurIPS)*, 33, 12230–12242. https://arxiv.org/abs/2002.07948

[22]. Garfinkel, S. L. (2014). De-identification of personal information. *NIST IR 8053*. https://doi.org/10.6028/NIST.IR.8053

[23]. George, M. B., Ijiga, M. O.& Adeyemi, O. (2025). Enhancing Wildfire Prevention and Grassland Burning Management with Synthetic Data Generation Algorithms for Predictive Fire Danger Index Modeling, *International Journal of Innovative Science and Research Technology* ISSN No:-2456-2165 Volume 10, Issue 3, https://doi.org/10.38124/ijisrt/25mar1859

[24]. Geyer, R. C., Klein, T., & Nabi, M. (2017). Differentially private federated learning: A client level perspective. *arXiv preprint arXiv:1712.07557*. https://arxiv.org/abs/1712.07557

[25]. Greitzer, F. L., Kangas, L. J., Noonan, C. F., Brown, C. E., & Ferryman, T. A. (2013). Psychosocial modeling of insider threat risk based on behavioral and word use analysis. *Information Systems Frontiers, 15*(1), 121–135. https://doi.org/10.1007/s10796-012-9333-2

[26]. Hitaj, B., Ateniese, G., & Perez-Cruz, F. (2017). Deep models under the GAN: Information leakage from collaborative deep learning. *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, 603–618. https://doi.org/10.1145/3133956.3134012

[27]. Idoko, I. P., Ijiga, O. M., Agbo, D. O., Abutu, E. P., Ezebuka, C. I., & Umama, E. E. (2024). Comparative analysis of Internet of Things (IOT) implementation: A case study of Ghana and the USA-vision, architectural elements, and future directions. **World Journal of Advanced Engineering Technology and Sciences**, 11(1), 180-199.

[28]. Idoko, I. P., Ijiga, O. M., Akoh, O., Agbo, D. O., Ugbane, S. I., & Umama, E. E. (2024). Empowering sustainable power generation: The vital role of power electronics in California's renewable energy transformation. **World Journal of Advanced Engineering Technology and Sciences**, 11(1), 274-293.

[29]. Idoko, I. P., Ijiga, O. M., Enyejo, L. A., Akoh, O., & Ileanaju, S. (2024). Harmonizing the voices of AI: Exploring generative music models, voice cloning, and voice transfer for creative expression

[30]. Idoko, I. P., Ijiga, O. M., Enyejo, L. A., Akoh, O., & Isenyo, G. (2024). Integrating superhumans and synthetic humans into the Internet of Things (IoT) and ubiquitous computing: Emerging AI applications and their relevance in the US context. **Global Journal of*

Engineering and Technology Advances*, 19(01), 006-036.

[31]. Idoko, I. P., Ijiga, O. M., Enyejo, L. A., Ugbane, S. I., Akoh, O., & Odeyemi, M. O. (2024). Exploring the potential of Elon Musk's proposed quantum AI: A comprehensive analysis and implications. **Global Journal of Engineering and Technology Advances**, 18(3), 048-065.

[32]. Idoko, I. P., Ijiga, O. M., Harry, K. D., Ezebuka, C. C., Ukatu, I. E., & Peace, A. E. (2024). Renewable energy policies: A comparative analysis of Nigeria and the USA.

[33]. Ihimoyan, M. K., Ibokette, A. I., Olumide, F. O., Ijiga, O. M., & Ajayi, A. A. (2024). The Role of AI-Enabled Digital Twins in Managing Financial Data Risks for Small-Scale Business Projects in the United States. *International Journal of Scientific Research and Modern Technology,* 3(6), 12–40. https://doi.org/10.5281/zenodo.14598498

[34]. Ijiga, M. O., Olarinoye, H. S., Yeboah, F. A. B. & Okolo, J. N. (2025). Integrating Behavioral Science and Cyber Threat Intelligence (CTI) to Counter Advanced Persistent Threats (APTs) and Reduce Human-Enabled Security Breaches. *International Journal of Scientific Research and Modern Technology, 4*(3), 1–15. https://doi.org/10.38124/ijsrmt.v4i3.376

[35]. Ijiga, O. M., Idoko, I. P., Ebiega, G. I., Olajide, F. I., Olatunde, T. I., & Ukaegbu, C. (2024). Harnessing adversarial machine learning for advanced threat detection: AI-driven strategies in cybersecurity risk assessment and fraud prevention. *Open Access Research Journals.* Volume 13, Issue. https://doi.org/10.53022/oarjst.2024.11.1.0060I

[36]. Imoh, P. O. (2023). Impact of Gut Microbiota Modulation on Autism Related Behavioral Outcomes via Metabolomic and Microbiome-Targeted Therapies *International Journal of Scientific Research and Modern Technology (IJSRMT)* Volume 2, Issue 8, 2023 DOI: https://doi.org/10.38124/ijsrmt.v2i8.494

[37]. Kairouz, P., McMahan, H. B., & Ramage, D. (2021). Advances and open problems in federated learning. *Foundations and Trends® in Machine Learning,* 14(1–2), 1–210. https://doi.org/10.1561/2200000083

[38]. Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., ... & Zhao, S. (2021). Advances and open problems in federated learning. *Foundations and Trends® in Machine Learning,* 14(1–2), 1–210. https://doi.org/10.1561/2200000083

[39]. Kent, K. & Souppaya, M. (2006). Guide to computer security log management. *NIST Special Publication 800-92.* https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-92.pdf

[40]. Latifa Albshaier (2025). *Federated Learning Workflow with Local Model Training and Centralized Global Model Aggregation.* Retrieved from: https://www.mdpi.com/2079-9292/14/5/1019

[41]. Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine*, 37(3), 50–60. https://doi.org/10.1109/MSP.2020.2975749

[42]. Li, T., Sahu, A. K., Zaheer, M., Sanjabi, M., Talwalkar, A., & Smith, V. (2020). Federated optimization in heterogeneous networks. *Proceedings of Machine Learning and Systems*, 2, 429–450. https://proceedings.mlsys.org/paper/2020/file/38a0d7d107a79607e3005b43c301e149-Paper.pdf

[43]. Lin, T., Long, G., Wang, T., Yao, L., & Zhang, C. (2020). Federated learning in distributed SQL systems: An architectural review. *IEEE Internet of Things Journal*, 7(9), 8450–8461. https://doi.org/10.1109/JIOT.2020.2998883

[44]. Liu, Y., Kang, Y., Zhang, X., & Yang, Q. (2020). Federated forest: Enabling decision tree-based analytics in federated learning. *IEEE Transactions on Big Data*, 6(3), 45–56. https://doi.org/10.1109/TBDATA.2020.2980738

[45]. Liu, Y., Zhang, Y., Zhang, Y., Fan, L., Tan, Y. A., & Ren, K. (2020). Secure federated transfer learning. *IEEE Transactions on Big Data*, *6*(3), 344–356. https://doi.org/10.1109/TBDATA.2020.2966185

[46]. Manuel, H. N. N., Adeoye, T. O., Idoko, I. P., Akpa, F. A., Ijiga, O. M., & Igbede, M. A. (2024). Optimizing passive solar design in Texas green buildings by integrating sustainable architectural features for maximum energy efficiency. **Magna Scientia Advanced Research and Reviews**, 11(01), 235-261. https://doi.org/10.30574/msarr.2024.11.1.0089

[47]. McMahan, H. B., Moore, E., Ramage, D., & Hampson, S. (2017). Communication-efficient learning of deep networks from decentralized data. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 1273–1282. https://proceedings.mlr.press/v54/mcmahan17a.html

[48]. MITRE Corporation. (2019). ATT&CK for Enterprise. *MITRE ATT&CK Knowledge Base*. https://attack.mitre.org

[49]. Mohri, M., Sivek, G., & Suresh, A. T. (2019). Agnostic federated learning. In *Proceedings of the 36th International Conference on Machine Learning (ICML)*, 97, 4615–4625. https://proceedings.mlr.press/v97/mohri19a/mohri19a.pdf

[50]. Nurse, J. R. C., Agrafiotis, I., Goldsmith, M., & Creese, S. (2014). A threat taxonomy for insider attacks. *Proceedings of the 11th International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment (DIMVA)*, 214–235. https://doi.org/10.1007/978-3-319-08509-8_11

[51]. Nwatuzie, G. A., Ijiga, O. M., Idoko, I. P., Enyejo, L. A. & Ali, E. O. (2025). Design and Evaluation of a User-Centric Cryptographic Model Leveraging Hybrid Algorithms for Secure Cloud Storage and Data Integrity. *American Journal of Innovation in Science and Engineering (AJISE).* Volume 4 Issue 1, SSN: 2158-7205 https://doi.org/10.54536/ajise.v4i2.4482

[52]. Okeke, R. O., Ibokette, A. I., Ijiga, O. M., Enyejo, L. A., Ebiega, G. I., & Olumubo, O. M. (2024). The

reliability assessment of power transformers. **Engineering Science & Technology Journal**, 5(4), 1149-1172.

[53]. Ononiwu, M., Azonuche, T. I., Okoh, O. F.. & Enyejo, J. O. (2023). Machine Learning Approaches for Fraud Detection and Risk Assessment in Mobile Banking Applications and Fintech Solutions *International Journal of Scientific Research in Science, Engineering and Technology* Volume 10, Issue 4 doi : https://doi.org/10.32628/IJSRSET

[54]. Oyebanji, O. S., Apampa, A. R., Idoko, P. I., Babalola, A., Ijiga, O. M., Afolabi, O. & Michael, C. I. (2024). Enhancing breast cancer detection accuracy through transfer learning: A case study using efficient net. *World Journal of Advanced Engineering Technology and Sciences*, 2024, 13(01), 285–318. https://wjaets.com/content/enhancing-breast-cancer-detection-accuracy-through-transfer-learning-case-study-using

[55]. Roba H. Alamir, Ayman Noor, Hanan Almukhalfi, Reham Almukhlifi and Talal H. Noor (2025). *SecFedDNN: A Secure Federated Deep Learning Framework for Edge–Cloud Environments*. Retrieved from: https://www.mdpi.com/2079-8954/13/6/463

[56]. Salem, M. B., Hershkop, S., & Stolfo, S. J. (2008). A survey of insider attack detection research. *Recent Advances in Intrusion Detection*, 69–90. https://doi.org/10.1007/978-3-540-87403-4_5

[57]. Shokri, R., & Shmatikov, V. (2015). Privacy-preserving deep learning. *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, 1310–1321. https://doi.org/10.1145/2810103.2813687

[58]. So, J., Smith, V., & Talwalkar, A. (2021). Evaluating the communication efficiency of federated learning. *IEEE Transactions on Network and Service Management*, 18(1), 5–20. https://doi.org/10.1109/TNSM.2021.3051275

[59]. Sommer, R., & Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. *IEEE Symposium on Security and Privacy*, 305–316. https://doi.org/10.1109/SP.2010.25

[60]. Thakkar, H., Niu, M., & Pedersen, T. B. (2021). VIRTUO: A virtualization framework for federated OLAP over SQL-based data lakes. *Proceedings of the VLDB Endowment*, 14(10), 1909–1921. https://doi.org/10.14778/3476311.3476318

[61]. Truex, S., Liu, L., Chow, K.-H., Gursoy, M. E., & Yu, L. (2019). A hybrid privacy-preserving framework for federated learning. *Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security*, 1–11. https://doi.org/10.1145/3338501.3357370

[62]. Tuor, A., Kaplan, S., Hutchinson, B., Nichols, N., & Robinson, S. (2017). Deep learning for unsupervised insider threat detection in structured cybersecurity data streams. *Proceedings of the AAAI Workshops*, WS-17-01. https://arxiv.org/abs/1710.00811

[63]. Wang, X., Han, Y., Wang, C., & Xu, H. (2022). Tackling non-IID data in federated learning: A unified perspective. *IEEE Transactions on Neural Networks and Learning Systems*, 33(7), 2940–2953. https://doi.org/10.1109/TNNLS.2021.3062373

[64]. Xu, J., Gursoy, M. E., & Velipasalar, S. (2020). Federated learning with differential privacy: Algorithms and performance analysis. *IEEE Transactions on Information Forensics and Security*, 15, 3454–3469. https://doi.org/10.1109/TIFS.2020.3028705

[65]. Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology (TIST), 10*(2), 1–19. https://doi.org/10.1145/3298981

[66]. Zhang, Y., & Wang, J. (2021). Addressing data skew and schema variation in federated SQL processing. *VLDB Endowment*, 14(12), 3141–3154. https://doi.org/10.14778/3476311.3476313