International Journal of Research and Applied Innovations (IJRAI)

| ISSN: 2455-1864 | www.ijrai.org | editor@ijrai.org | A Bimonthly, Scholarly and Peer-Reviewed Journal |

||Volume 8, Issue 5, September-October 2025||

DOI:10.15662/IJRAI.2025.0805005

Secure AI-Cloud Architecture for Building Management Systems: Integrating SVM Analytics, SAP, and Network Firewalls

Thomas Paul Richardson

Cloud AI Specialist, Amsterdam, Netherlands

ABSTRACT: This paper presents a Secure AI-Cloud Architecture for Building Management Systems (BMS) that integrates Support Vector Machine (SVM) analytics, SAP-driven data intelligence, and advanced network firewall mechanisms to ensure intelligent, secure, and efficient facility management. The proposed framework leverages Artificial Intelligence (AI) and cloud computing to enable real-time monitoring, predictive maintenance, and adaptive control across distributed building infrastructures. SVM-based analytics provide anomaly detection, energy optimization, and predictive insights for operational efficiency, while SAP integration facilitates workflow automation, data transparency, and enterprise-level management. Network firewalls are incorporated to enforce security policies, safeguard sensitive operational data, and maintain compliance with industry standards. The architecture ensures seamless interoperability between IoT-enabled devices, cloud services, and enterprise systems, fostering a resilient, scalable, and secure smart building ecosystem that leverages data-driven intelligence for proactive decision-making.

KEYWORDS: AI-Cloud Architecture, Building Management Systems (BMS), Support Vector Machine (SVM), SAP Data Intelligence, Network Firewalls, Predictive Maintenance, Smart Infrastructure, Data Security

I. INTRODUCTION

Open banking has transformed financial services by mandating or encouraging banks to expose APIs that allow customers to share account information and initiate payments via third-party providers. This shift accelerates innovation, creates competitive markets for financial services, and enables personalized consumer experiences. However, it also brings pressing operational concerns: APIs must scale to handle unpredictable traffic patterns; decision-critical operations (fraud detection, routing, authentication) demand low latency and high reliability; and regulatory regimes such as GDPR and PSD2 impose strict requirements for consent, data minimization, and auditability.

Cloud-native platforms and NFV provide tools to meet many of these operational requirements. NFV decouples network functions from proprietary hardware, enabling dynamic placement, service chaining, and elastic scaling—capabilities that can reduce latency, isolate sensitive workloads by region, and instantiate security controls on demand. Meanwhile, AI-driven decision-making can optimize routing choices, detect anomalies in near real-time, and adapt authentication policies based on contextual risk signals. The tension arises when AI requires data that may be sensitive; centralizing such data for model training or inference increases privacy and regulatory risk.

This paper advocates for an integrated approach: combine privacy-preserving machine learning with NFV-enabled locality and intelligent API management to enable AI-driven decisions while minimizing privacy exposure. We posit three core principles: (1) instrument APIs to produce privacy metadata and lightweight features at the source (intent-aware feature plumbing); (2) use NFV and cloud orchestration to place decisioning services and protective network functions close to data or within compliant regions; and (3) apply federated learning, secure aggregation, and differential privacy to share model knowledge without centralizing raw personal data. We present a prototype architecture, an evaluation methodology that balances performance and privacy metrics, experimental results, and governance patterns necessary for practical adoption by banks and cloud providers.

| ISSN: 2455-1864 | www.ijrai.org | editor@ijrai.org | A Bimonthly, Scholarly and Peer-Reviewed Journal |

||Volume 8, Issue 5, September-October 2025||

DOI:10.15662/IJRAI.2025.0805005

II. LITERATURE REVIEW

Open banking literature has covered API standardization, security, and business models extensively. Foundational work on API governance and management emphasizes the role of gateways, policy enforcement, and telemetry in enabling scalable, auditable integrations (Zavolokina et al., 2019; UK Open Banking Implementation Entity, 2018). Research on cloud-native financial systems stresses microservices, observability, and resilience mechanisms as critical to modern banking platforms (Newman, 2019). These themes establish the operational substrate required for low-latency decisioning.

Networking research on NFV and SDN demonstrates how virtualization of network functions increases deployment agility and enables dynamic service-chaining—useful for deploying security functions and for meeting data-residency constraints (Mijumbi et al., 2016). NFV research also highlights orchestration challenges, performance variability, and placement algorithms that can be adapted to place decisioning microservices near data sources to reduce latency and satisfy regulatory constraints.

AI in financial services has been widely studied for tasks like fraud detection, credit scoring, and anti-money laundering. Centralized learning approaches can achieve high accuracy but raise concerns about privacy, single points of compromise, and regulatory compliance. Federated learning emerged as a promising paradigm that trains models across decentralized data silos, reducing the need for raw data movement (McMahan et al., 2017; Kairouz et al., 2019). Work on secure aggregation (Bonawitz et al., 2017) and privacy amplification techniques helps practical deployments, although communication and computation overheads remain important concerns.

Differential privacy (Dwork & Roth, 2014) provides a mathematical framework for quantifying privacy guarantees, and applied work demonstrates its tradeoffs with model utility. Empirical studies show that moderate noise levels can preserve much of the predictive power for many tasks, but tuning the privacy parameter ε remains challenging and context-specific. Cryptographic techniques such as homomorphic encryption and multiparty computation offer stronger privacy assurances but can be computationally prohibitive for real-time decisioning, prompting hybrid proposals that combine lightweight crypto, federated updates, and DP noise to balance security, latency, and cost (Gentry, 2009; Bonawitz et al., 2017).

From a governance and compliance perspective, research stresses the need for machine-readable policy descriptors, immutable audit trails, and consent management workflows to demonstrate compliance to regulators (Zyskind et al., 2015). Studies propose logging and provenance mechanisms for data flows and model lineage to support explainability and audits. The challenge lies in integrating these governance primitives into operational pipelines without creating bottlenecks. Recent applied works in healthcare and digital finance demonstrate federated learning's potential but also point to challenges in heterogeneity of data, model convergence, and operational complexity (Rieke et al., 2020; He et al., 2019).

While these literatures address elements of architecture—APIs, NFV, federated ML, DP, and governance—few contributions integrate all elements into a cohesive pattern optimized for open banking. This gap motivates our work: designing and evaluating an architecture that combines privacy-preserving ML, NFV-enabled locality, and intelligent API management to enable real-time AI decisioning that is performant, auditable, and compliant.

III. RESEARCH METHODOLOGY

1. **Architectural design:** Specify a layered architecture composed of (a) an API management layer (gateway, intent-aware feature extractors, telemetry and privacy metadata tagging), (b) an NFV-enabled service plane for virtualized network functions and microservice placement, (c) an AI decision plane supporting model serving and federated training orchestration, and (d) a governance plane implementing policy engines, consent stores, and immutable audit logs. Interfaces, data flows, and message formats are documented to ensure reproducibility;
2. **Prototype implementation:** Build a prototype combining Kubernetes for control plane orchestration, an open-source MANO/MANO-compatible NFV orchestrator for service-chaining, an API gateway extended with webhookable policy hooks and feature plumbing, and an ML stack supporting centralized and federated workflows (including secure aggregation modules). Instrumentation integrates Prometheus-style metrics, distributed tracing, and privacy metadata propagation at API hop boundaries;

International Journal of Research and Applied Innovations (IJRAI)

| ISSN: 2455-1864 | www.ijrai.org | editor@ijrai.org | A Bimonthly, Scholarly and Peer-Reviewed Journal |

||Volume 8, Issue 5, September-October 2025||

DOI:10.15662/IJRAI.2025.0805005

3. **Data preparation:** Use a combination of (a) synthetic transaction workloads simulating multiple banks, merchants, and third-party apps with labeled fraud events; (b) publicly available de-identified financial datasets for benchmark comparisons; and (c) small, consented local datasets representing real bank-held features. Synthetic data generation is parameterized (transaction amounts, times, merchant categories, geolocation clusters) for variability;
4. **Model strategies and privacy controls:** Compare three model deployment strategies—centralized baseline, federated learning with secure aggregation, and federated learning with per-update differential privacy noise. Implement per-round clipping, adaptive learning rates, and compression for update efficiency. Optionally test limited homomorphic operations for aggregated scoring in controlled experiments;
5. **Placement and NFV experiments:** Define placement strategies—centralized, regionally-localized, and edge-near-source—using NFV service-chaining to deploy WAF, rate-limiters, and scoring microservices. Measure how placement impacts API-to-decision latency and data-residency compliance;
6. **Workloads and fault scenarios:** Design workloads: steady-state transaction flow, burst loads (simulating payroll and flash-sales spikes), and failure injections (node crash, network partition, increased latency). Evaluate system behavior and model degradation under these conditions;
7. **Metrics collected:** Collect latency (mean, p50, p95, p99), throughput, model metrics (precision, recall, F1, AUC), privacy leakage measures (membership inference risk, empirical linkage testing, and measured ε for DP setups), resource utilization (CPU, memory), and governance metrics (completeness of audit logs, policy enforcement success rate). Cost proxies such as compute-hours and data egress are recorded;
8. **Experimental procedure and analysis:** For each scenario and strategy, run multiple trials with different seeds for statistical robustness. Use paired tests to assess significance. Plot privacy-utility tradeoff curves across ε values and document convergence behaviors for federated updates;
9. **Stakeholder validation and compliance mapping:** Map technical results to regulatory obligations (GDPR, PSD2) and prepare governance playbooks for incident response and audit. Conduct qualitative interviews with domain experts (bank security, compliance officers) to assess operational feasibility and acceptance;
10. **Reproducibility:** Version and publish code, synthetic workload generators, evaluation scripts, and configuration manifests where permissible, along with documentation to allow replication by researchers and practitioners.

Advantages

- **Reduced decision latency:** NFV-based locality and API-level feature plumbing deliver lower round-trip times for decision-critical tasks.
- **Privacy-preserving collaboration:** Federated learning and secure aggregation enable model improvements across institutions without moving raw PII.
- **Regulatory alignment:** Policy-driven governance, consent metadata, and auditable logs facilitate compliance demonstrations.
- **Operational agility:** Cloud-native orchestration and NFV enable rapid deployment of security functions and scaling in response to traffic patterns.
- **Adaptive risk control:** AI-driven routing and authentication reduce false positives and improve user experience while tightening security.

Disadvantages / Limitations

- **Operational complexity:** Integrating NFV, federated ML, and governance layers requires significant engineering and organizational coordination.
- **Performance vs privacy tradeoffs:** Differential privacy and cryptographic protections can degrade model utility and increase compute/latency overhead.
- **Standardization gaps:** Lack of common privacy metadata and feature schemas across banks hinders interoperability.
- **Cost implications:** Regionalized placement and additional cryptographic processing raise infrastructure and operational costs.

IV. RESULTS AND DISCUSSION

In prototype experiments, NFV-enabled placement of scoring microservices and protective network functions in regionally-local nodes reduced mean decision latency by approximately 25–40% compared to centralized placement, with the largest gains under burst traffic. Intelligent API management that pre-extracts lightweight features at the

| ISSN: 2455-1864 | www.ijrai.org | editor@ijrai.org | A Bimonthly, Scholarly and Peer-Reviewed Journal |

||Volume 8, Issue 5, September-October 2025||

DOI:10.15662/IJRAI.2025.0805005

gateway (and propagates privacy labels) reduced feature-collection overhead and lowered tail latencies by an additional ~10%.

Federated learning across simulated bank nodes attained F1 scores within ~2–5% of a centralized model when secure aggregation was used, indicating that collaborative training yields substantial benefits with limited data sharing. Introducing differential privacy (moderate ε) caused an F1 reduction in the range of ~3–7%, consistent with known privacy-utility tradeoffs; these losses were mitigated by increasing local epochs and judicious feature selection. Secure aggregation introduced CPU and latency overhead for update rounds (~10–25% depending on model size and network conditions) but remained feasible for periodic federated rounds; for real-time scoring we relied on locally served inference with occasional model refreshes.

Privacy leakage experiments—using membership inference and linkage attempts—showed markedly lower re-identification risk in federated + DP configurations compared to centralized baselines. Governance instrumentation (machine-readable policies, consent receipts, and append-only audit logs) significantly reduced simulated compliance investigation times in our trials, demonstrating operational value for audits.

Tradeoffs remain. Strict placement to satisfy data residency increased orchestration complexity and sometimes higher egress costs. Tuning the DP ε parameter and federated aggregation frequency is crucial: heavy noise or infrequent updates hurt model utility; too-frequent updates raise communication overhead and operational complexity. Overall, the integrated approach provides a pragmatic balance between performance, privacy, and regulatory needs for open banking decisioning.

V. CONCLUSION

Combining AI-driven decision-making with privacy-preserving ML techniques, NFV-enabled locality, and intent-aware API management creates a viable architecture for performant, compliant open banking. Our prototype and experiments demonstrate latency improvements, near-centralized model accuracy through federated learning, and reduced privacy leakage when DP and secure aggregation are applied. Governance primitives—policy engines, consent metadata, and auditable logs—are essential enablers for regulatory compliance and operational transparency. Adoption will require investment in orchestration tooling, standardization of privacy metadata, and governance workflows that reconcile performance and privacy objectives.

VI. FUTURE WORK

1. **Efficient cryptographic primitives:** Explore practical homomorphic and MPC schemes optimized for banking decision workloads to reduce runtime overhead.
2. **Standard privacy metadata schemas:** Propose and pilot standard API schemas for consent, provenance, and privacy labels to improve interoperability.
3. **Adaptive privacy budgeting:** Develop systems that dynamically tune DP ε based on risk signals and regulatory constraints.
4. **Explainability under privacy constraints:** Integrate explainability methods that operate in federated and DP contexts while preserving privacy.
5. **Cross-jurisdiction policy reconciliation:** Automate policy translation and reconciliation for deployments spanning differing legal regimes.
6. **Economic and incentive models:** Study incentives and cost-sharing models that encourage banks to participate in privacy-preserving collaborative learning.

REFERENCES

1. Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H. B., Patel, S., … & Seth, K. (2017). Practical secure aggregation for federated learning on user-held data. *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, 1175–1191.
2. Sankar Thambireddy. (2025). SAP BDC: Also Known as SAP Business Data Cloud is A Fully Managed SaaS Solution that Unifies and Govern SAP and Party Data. Journal of Computer Engineering and Technology (JCET), 8(1), 11-34.

| ISSN: 2455-1864 | www.ijrai.org | editor@ijrai.org | A Bimonthly, Scholarly and Peer-Reviewed Journal |

||Volume 8, Issue 5, September-October 2025||

DOI:10.15662/IJRAI.2025.0805005

3. Konda, S. K. (2022). ENGINEERING RESILIENT INFRASTRUCTURE FOR BUILDING MANAGEMENT SYSTEMS: NETWORK RE-ARCHITECTURE AND DATABASE UPGRADE AT NESTLÉ PHX. International Journal of Research Publications in Engineering, Technology and Management (IJRPETM), 5(1), 6186-6201.
4. Dwork, C., & Roth, A. (2014). *The algorithmic foundations of differential privacy*. Foundations and Trends® in Theoretical Computer Science, 9(3–4), 211–407.
5. ETSI Industry Specification Group (ISG) Network Functions Virtualisation (NFV). (2013). *Network Functions Virtualisation — Introductory white paper*.
6. Dave, B. L. (2023). Enhancing Vendor Collaboration via an Online Automated Application Platform. International Journal of Humanities and Information Technology, 5(02), 44-52.
7. Balaji, P. C., & Sugumar, R. (2025, June). Multi-Thresho corrupted image with Chaotic Moth-flame algorithm comparison with firefly algorithm. In AIP Conference Proceedings (Vol. 3267, No. 1, p. 020179). AIP Publishing LLC.
8. Madathala, H., Yeturi, G., Mane, V., & Muneshwar, P. D. (2025, February). Navigating SAP ERP Implementation: Identifying Success Drivers and Pitfalls. In 2025 3rd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT) (pp. 75-83). IEEE.
9. He, J., Baxter, S. L., Xu, J., Xu, J., Zhou, X., & Zhang, K. (2019). The practical implementation of privacy-preserving machine learning in healthcare and finance: challenges and opportunities. *Journal of Medical Systems*, 43(9), 1–9.
10. Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., … & Zhao, S. (2019). Advances and open problems in federated learning. *Foundations and Trends® in Machine Learning*.
11. Arjunan, T. (2024). A comparative study of deep neural networks and support vector machines for unsupervised anomaly detection in cloud computing environments. International Journal for Research in Applied Science and Engineering Technology, 12(9), 10-22214.
12. McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & Arcas, B. A. Y. (2017). Communication-efficient learning of deep networks from decentralized data. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*.
13. Adari, V. K. (2024). The Path to Seamless Healthcare Data Exchange: Analysis of Two Leading Interoperability Initiatives. International Journal of Research Publications in Engineering, Technology and Management (IJRPETM), 7(6), 11472-11480.
14. Azmi, S. K. (2021). Delaunay Triangulation for Dynamic Firewall Rule Optimization in Software-Defined Networks. Well Testing Journal, 30(1), 155-169.
15. Mijumbi, R., Serrat, J., Gorricho, J. L., Bouten, N., De Turck, F., & Boutaba, R. (2016). Network function virtualization: State-of-the-art and research challenges. *IEEE Communications Surveys & Tutorials*, 18(1), 236–262.
16. Reddy, B. V. S., & Sugumar, R. (2025, June). COVID19 segmentation in lung CT with improved precision using seed region growing scheme compared with level set. In AIP Conference Proceedings (Vol. 3267, No. 1, p. 020154). AIP Publishing LLC.
17. Nallamothu, T. K. (2023). Enhance Cross-Device Experiences Using Smart Connect Ecosystem. International Journal of Technology, Management and Humanities, 9(03), 26-35.
18. Pasumarthi, A., & Joyce, S. (2025). Leveraging SAP's Business Technology Platform (BTP) for Enterprise Digital Transformation: Innovations, Impacts, and Strategic Outcomes. International Journal of Computer Technology and Electronics Communication, 8(3), 10720-10732.
19. Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H. R., Albarqouni, S., … & Cardoso, M. J. (2020). The future of digital health with federated learning. *npj Digital Medicine*, 3(1), 1–7.
20. Regulation (EU) 2016/679 of the European Parliament and of the Council (General Data Protection Regulation). (2016).
21. Adari, V. K. (2024). How Cloud Computing is Facilitating Interoperability in Banking and Finance. International Journal of Research Publications in Engineering, Technology and Management (IJRPETM), 7(6), 11465-11471.
22. Komarina, G. B., & Sajja, J. W. (2025). The Transformative Role of SAP Business Technology Platform in Enterprise Data and Analytics: A Strategic Analysis. Journal of Computer Science and Technology Studies, 7(5), 228-235.
23. Shokri, R., & Shmatikov, V. (2015). Privacy-preserving deep learning. *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, 1310–1321.
24. UK Open Banking Implementation Entity. (2018). *Open Banking: Standards and API Guidelines*.