Received: 29 July 2026, Accepted: 28 August 2026, Published: 31 August 2026
Digital Object Identifier: https://doi.org/10.63503/j.ijaimd.2026.274

Research Article

Explainable AI for Adaptive Security in Regulated Environments: A Unified Framework Integrating Federated Risk-Based Authentication and Privacy-Preserving On-Premises LLM Deployment

Ashok Kumar[1], Utku Kose[2]

[1]Department of IT, College of Technology, G. B. Pant University of Agriculture and Technology, Pantnagar, Uttarakhand, India

[2]Department of Computer Engineering, Suleyman Demirel University, Isparta, Turkey

[2]Affiliated Researcher, University of North Dakota, USA

ashokkumar.it@gbpuat-tech.ac.in, utku.kose@sdu.edu.tr

*Corresponding author: Naveen, ashokkumar.it@gbpuat-tech.ac.in

ABSTRACT

Healthcare and public-sector organizations face a compounding security imperative: protecting sensitive personal data against adaptive cyber threats while preserving the operational fluency that practitioners require in time-critical workflows. This paper proposes a unified Explainable AI (XAI) security framework that synthesizes federated learning-enhanced Dynamic Risk-Based Authentication (FL-RBA) with privacy-preserving on-premises Large Language Model (LLM) deployment into a cohesive, compliance-native paradigm. The framework introduces SHAP-based rationale generation to address the interpretability gap inherent in transformer-based authentication scoring, producing audit-ready decision trails that satisfy HIPAA and GDPR requirements. Pilot evidence from a live healthcare deployment demonstrates a 95% high-risk interception rate, 2.5% false-positive rate, and sub-1.2-second decision latency. On-premises LLM integration eliminates all external data transmission while delivering clinical intelligence capabilities comparable to cloud-based alternatives. The unified model provides a replicable blueprint for regulated organizations seeking to operationalize AI without compromising data sovereignty or regulatory alignment.

Keywords: *Explainable AI; Federated Learning; Risk-Based Authentication; Healthcare Cybersecurity; On-Premises LLM; Oracle APEX; Privacy-Preserving AI; SHAP; Adaptive MFA; XAI Compliance; Low-Code Security; Transformer Models; Behavioral Biometrics; Data Sovereignty; Differential Privacy; AI Safety; Machine Ethics.*

1. Introduction

The digital transformation of healthcare and public-sector administration has produced profound gains in service delivery, enabling electronic health record access, telemedicine, predictive population health analytics, and interconnected administrative workflows at unprecedented scale. Yet the same networked infrastructure that accelerates institutional decision-making has exponentially expanded the attack surface available to adversaries. Protected Health Information (PHI) commands a premium on illicit markets; individual records reportedly fetch prices exceeding $1,000 each [1]. The IBM Cost of a Data Breach Report 2024 reported an average healthcare breach cost of $9.77 million, the highest of any sector, driven by regulatory penalties, remediation overhead, and clinical disruption [2]. The 2025 Verizon Data Breach Investigations Report

documented 1,710 security incidents in healthcare, with 1,542 confirmed data disclosures representing a 20% year-over-year increase [3].

Authentication represents the most consequential entry point into this threat landscape. Static password systems and uniform multi-factor authentication (MFA) protocols remain standard across much of the sector despite their documented inadequacy. Password credentials are susceptible to credential stuffing, phishing, and brute-force attacks, while blanket MFA imposes friction that practitioners under time pressure frequently circumvent [4]. Neither paradigm accounts for the rich contextual variation characterizing legitimate access: a clinician authenticating from a familiar workstation during a scheduled shift presents a fundamentally different risk profile from an unrecognized device authenticating from a foreign geolocation at 3 a.m.

Simultaneously, regulated-sector organizations face a structural barrier in AI adoption. Institutions handling sensitive population data are under ethical and regulatory pressure to deploy AI-assisted capabilities, natural language query interfaces, automated documentation, and anomaly narration, yet cloud-based AI services require transmitting that data to external servers, creating unacceptable exposure in environments governed by HIPAA, GDPR, and equivalent frameworks.

This paper proposes a unified Explainable AI security framework that addresses both challenges through deliberate architectural integration. The framework synthesizes two recent empirical contributions from public health environments, a federated learning-enhanced Dynamic Risk-Based Authentication (FL-RBA) system achieving 95% high-risk interception in a live healthcare portal [5], and an on-premises open-source LLM deployment architecture enabling AI-assisted workflows without any data transmission beyond organizational boundaries [6] into a cohesive compliance-native security paradigm. The unifying contribution is the XAI Rationale Engine: a SHAP-based interpretability layer that bridges the opacity of transformer authentication models with the audit trail requirements of regulated environments.

The framework is grounded in zero-trust principles, in which every authentication event is independently evaluated against a multidimensional contextual risk score computed by locally trained transformer models coordinated through federated learning. All scoring decisions are accompanied by human-readable rationale strings generated via SHAP feature attribution, satisfying HIPAA audit control requirements and GDPR Article 22 obligations. The on-premises LLM layer supports clinical intelligence functions without any PHI leaving the institutional network boundary.

2. Literature Review

A. Risk-Based and Adaptive Authentication

Risk-Based Authentication (RBA) departs from static credential verification by evaluating the contextual risk profile of each authentication attempt in real time. Wiefling et al. conducted one of the first empirical analyses of RBA in production environments, demonstrating that contextual signals, such as IP reputation, device fingerprinting, and behavioral history, can reduce unauthorized access rates by up to 35% while preserving usability for legitimate users [7]. Singh et al. extended this with the RAD-AA framework, demonstrating dynamic threshold adaptation across heterogeneous user populations [8]. Chauhan and Kumar established that machine learning-based adaptive authentication significantly outperforms rule-based systems in detecting shift-based access irregularities in clinical environments [9]. Wairagade et al. confirmed that deep learning models consistently outperform statistical baselines in terms of precision and recall for healthcare

authentication [10]. Despite these advances, existing implementations lack the interpretability mechanisms required for regulatory audit, a gap this paper directly addresses.

B. Federated Learning in Healthcare Security

Federated Learning (FL) enables collaborative model training across distributed data sources without centralizing raw data. McMahan et al. established the FedAvg algorithm as the foundational communication-efficient paradigm for decentralized deep learning [11]. Fereidouni et al. introduced F-RBA, the first federated learning framework specifically designed for risk-based authentication [12]. Mazzocca et al. presented FRAMH, an integrated blockchain-based system for tamper-evident audit trails [13]. Baseri et al. proposed FL-RBA2 to address the non-IID data challenge in federated healthcare deployments [14]. Ali et al. introduced Health-FedNet, demonstrating privacy-preserving, scalable federated analytics across regional health networks [15]. None of these works integrates XAI mechanisms that satisfy HIPAA documentation standards for automated access control, a gap the present framework addresses.

C. Explainable AI: Theory and Healthcare Applications

Explainability has become a defining requirement for responsible AI deployment in regulated domains. Within the author's broader research program on explainable AI, machine ethics, and AI safety, the interpretability of automated decision systems represents a foundational concern [16]. Adadi and Berrada provide a comprehensive taxonomy of XAI methods categorized by model type and explanation target [17]. SHAP (SHapley Additive exPlanations), introduced by Lundberg and Lee, offers theoretically grounded local explanations by computing Shapley values quantifying each feature's marginal contribution to a given prediction [18]. Miller demonstrates that the quality of the explanation determines whether practitioners trust and act on AI recommendations [19]. Saraswat et al. survey XAI in Healthcare 5.0, identifying authentication and anomaly detection as high-priority domains for interpretability investment [20]. Al Ansari et al. demonstrate that decision transparency directly correlates with user acceptance of AI-driven security interventions [21].

D. On-Premises LLM Deployment

The emergence of capable open-source LLMs has created a viable alternative to cloud-based AI for data-sovereignty-constrained organizations. Meta's LLaMA 3, trained on over 15 trillion tokens, demonstrates competitive performance with parameter counts that are feasible for institutional on-premises deployment [22]. Mixtral 8x7B employs a sparse mixture-of-experts architecture to improve efficiency in resource-constrained hardware environments [23]. Google's Gemma 7B provides a lightweight option suitable for edge scenarios [24]. Pati emphasizes combining on-premises inference with differential privacy to bound re-identification risk in healthcare federated contexts [25]. Gubitosa provides a practitioner deployment guide stressing containerization, GPU configuration, and API compatibility layers [26]. Ollama has emerged as the predominant lightweight deployment approach enabling direct substitution into enterprise application stacks without modifying application logic [6].

E. AI Ethics, Safety, and Compliance

Rose et al. articulate the NIST Zero Trust Architecture framework, establishing that healthcare security systems must independently verify every access event [27]. Yang et al. identify transparency, fairness, and accountability as non-negotiable properties for federated systems in regulated domains [28]. The intersection of AI safety and healthcare authentication presents particular ethical complexity: over-blocking creates patient safety risks through delayed clinical

intervention, while under-blocking creates privacy risks through unauthorized PHI exposure. The present work addresses this trade-off through configurable role-specific risk thresholds, human-in-the-loop override mechanisms, and XAI rationale delivery, enabling clinicians to contest automated access decisions.

F. Gap Analysis

The reviewed literature reveals three unresolved gaps: (1) no existing federated RBA system integrates XAI mechanisms sufficient for HIPAA audit trail requirements; (2) on-premises LLM deployments remain architecturally disconnected from authentication infrastructure; and (3) the AI safety dimensions of adaptive healthcare authentication have not been systematically addressed within a unified framework. This paper contributes a design that resolves all three simultaneously.

3. Methodology

A. Unified Framework Architecture

The proposed framework extends the five-layer FL-RBA architecture empirically validated in [5] with two additional components: an XAI Rationale Engine and a Privacy-Preserving AI Intelligence Layer. The resulting seven-layer architecture forms a cohesive security ecosystem deployable on Oracle APEX within Oracle Cloud Infrastructure (OCI) or on-premises Oracle Database environments. Table 1 summarizes all seven layers with their constituent technologies and primary compliance functions. Each layer is independently modular, permitting component-level upgrades without disrupting overall system function. Layer 4, the XAI Rationale Engine (marked * in Table 1), represents the primary novel contribution of this paper and is detailed in Section III-C.

Table 1. Seven-Layer Unified XAI Security Framework Architecture

Layer	Component	Key Technology	Compliance Function
1	Client Interface	APEX 24.2; FingerprintJS; JS plugins	Metadata capture; session staging
2	Risk Context Aggregation	PL/SQL; DBMS_CRYPTO; IPStack API	PHI hashing; GDPR Art. 5(1)(c)
3	Federated AI Scoring	TensorFlow Federated; BERT; FedAvg; SMPC	Decentralized training; re-ID risk <1%
4*	XAI Rationale Engine	SHAP DeepExplainer / TreeExplainer; OCI Functions	HIPAA §164.312(b); GDPR Art. 22
5	Auth. Controller	PKG_RBA_CONTROLLER; Twilio OTP; APEX_AUTH	Tiered enforcement; zero-trust

Layer	Component	Key Technology	Compliance Function
6	On-Premises AI	Ollama; LLaMA 3 / Mixtral / Gemma 7B	Zero PHI egress; HIPAA §164.306
7	Audit & Feedback Loop	Oracle TDE; DBMS_SCHEDULER; HITL dashboard	7-yr retention; federated retraining

** Novel contribution of the present work`*

Layer 1 manages user interactions through Oracle APEX 24.2 responsive UI components, augmented by JavaScript plugins that collect client-side metadata during credential submission. Layer 2 compiles device fingerprints, geolocation data, behavioral history, and temporal patterns into a structured JSON payload; sensitive fields are hashed using Oracle's DBMS_CRYPTO, and geolocation is coarsened to the regional level, satisfying GDPR data minimization. Layer 3 coordinates TensorFlow Federated-trained transformer models distributed across regional APEX instances, aggregating local inferences via FedAvg through an ORDS-exposed central aggregator. Layer 5 enforces tiered policy through PKG_RBA_CONTROLLER. Layer 7 captures all authentication events in encrypted Oracle tables and drives weekly federated retraining through a human-in-the-loop annotation interface.

B. Federated Risk Scoring Protocol

During each authentication event, the local APEX instance computes a contextual risk inference from the aggregated JSON payload and contributes encrypted gradient updates to the central aggregator. The aggregator applies FedAvg to produce an updated global risk score, returned within 1.2 seconds under standard network conditions. Caching of low-risk user profiles in APEX Collections reduces repeat-login latency to 0.7 seconds. Secure multi-party computation (SMPC) ensures that gradient updates remain confidential during aggregation, preventing the reconstruction of individual sites' training data. Homomorphic encryption bounds re-identification risk below 1% per the NIST differential privacy guidelines [27].

Risk score thresholds are configurable through PKG_RBA_POLICY. Scores below 0.4 result in seamless session establishment via APEX_AUTH. Scores between 0.4 and 0.7 trigger step-up authentication via OTP (Twilio SMS API) or biometric verification. Scores above 0.7 result in session denial with real-time administrative alerting. Emergency override provisions role-specific whitelist entries with time-bounded relaxation windows, and reduce inappropriate denials during critical clinical workflows such as surgical planning and emergency triage.

C. XAI Rationale Engine: SHAP Integration

The XAI Rationale Engine is a Python microservice deployed as an OCI Function that is invoked asynchronously after each federated scoring event. It applies SHAP DeepExplainer to the transformer scoring model, computing per-feature Shapley values quantifying each contextual signal's marginal contribution to the prediction. Features assessed include geolocation deviation from historical baseline, device recognition status, temporal anomaly, behavioral velocity, and IP reputation score.

Shapley values are mapped to human-readable rationale strings through a template engine, for example, "Login from unrecognized geographic region (+0.31 risk contribution)" or "Authentication attempt 4.2 standard deviations outside established access window (+0.28 risk contribution)." Rationale strings are delivered through three channels: (1) logged in encrypted audit tables for HIPAA 45 CFR §164.312(b) audit control compliance; (2) surfaced to security administrators in the APEX Risk Dashboard; and (3) optionally presented to end-users as contextual MFA prompts, satisfying GDPR Article 22 requirements. For latency-sensitive deployments, a lightweight gradient-boosted surrogate model enables the TreeExplainer application with an overhead of 40ms or less, compared to 180ms for DeepExplainer.

D. On-Premises LLM Configuration

The Privacy-Preserving AI Intelligence Layer deploys open-source LLMs via Ollama on organizational server infrastructure and exposes an OpenAI-compatible REST API endpoint. Oracle APEX 24.2's native Generative AI Services framework connects to this endpoint by overriding the commercial API base URL with the local Ollama address, enabling direct substitution without modifying any APEX application logic. LLaMA 3 (8B quantized GGUF) is recommended for general-purpose clinical documentation assistance; Mixtral 8x7B for multilingual and mathematical computation tasks; Gemma 7B for resource-constrained environments requiring only 8GB VRAM. Domain-specific fine-tuning is performed on-premises using LoRA adapters with Hugging Face PEFT libraries. The LLM layer connects to the authentication audit log via PL/SQL procedures, enabling natural-language security intelligence queries with zero PHI egress.

E. Compliance Architecture

Compliance alignment is enforced by design across all seven layers. HIPAA Security Rule requirements are addressed through Transparent Data Encryption (TDE), RBAC via APEX authorization schemes, 7-year audit log retention with automated purging via DBMS_SCHEDULER, and XAI rationale strings satisfying the minimum-necessary documentation standard. GDPR alignment is achieved through data minimization, differential privacy bounds for federated gradient sharing, and delivery of Article 22 rationale. NIST Zero Trust Architecture principles [27] are enforced through per-session independent verification, TLS-encrypted inter-layer communication, and continuous behavioral monitoring.

4. Discussions and Findings

A. Healthcare Pilot Performance

The FL-RBA component was empirically validated in a 45-day production pilot within a public health immunization management portal [5], serving over 2,500 users across three regional OCI instances (Virginia, Illinois, California). Over the pilot period, 32,000 login attempts were processed, issuing 810 MFA challenges and blocking 152 high-risk sessions, with 24 confirmed false positives (2.5% false-positive rate). Ten confirmed intrusion attempts were intercepted, including three credential-stuffing incidents traceable to dark pool credential datasets. Average decision latency was 1.1 seconds, which was reduced to 0.7 seconds with low-risk-profile caching. System uptime was 99.98%.

Table 2 presents a comparative evaluation of the proposed FL-RBA framework against three conventional authentication approaches across eight performance dimensions. The proposed framework achieves significant improvements over all baselines across threat interception, false-positive rate, and user satisfaction, while maintaining full PHI data sovereignty.

Table 2. Comparative Performance: FL-RBA + XAI vs. Conventional Methods

Metric	Static Password.	Unified MFA	Rule-Based RBA	Proposed FL-RBA + XAI
High-Risk Interception	~40%	~65%	~72%	95%
False-Positive Rate	~15%	8–12%	~6%	2.5%
Avg. Latency	<0.5 s	5–10 s	1–3 s	1.1 s
Context Awareness	None	Low	Moderate	High
XAI / Audit Trail	None	None	Partial	Full (SHAP)
PHI Sovereignty	High	High	High	High (On-Prem)
Clinician Satisfaction	~70%	~55%	~68%	90%
HIPAA/GDPR Alignment	Partial	Partial	Moderate	Full

The XAI Rationale Engine produced measurable secondary benefits beyond regulatory compliance. Post-pilot surveys indicated that SHAP-generated rationale strings increased security team confidence in blocking decisions by 34%, reducing manual override frequency. Compliance officers confirmed that rationale entries satisfied all documentation requirements under the HIPAA audit control standard (45 CFR § 164.312(b)), thereby eliminating the manual annotation workload previously required for blocked-session records.

B. Cross-Domain Validation

Cross-domain validation was conducted in two sandbox environments. In a financial services loan processing application serving 200 users, the framework achieved a 3% false-positive rate and 98% user satisfaction. On an educational technology platform serving 500 students, 12 anomalous login events were detected, yielding a 2% false-positive rate. In both cases, only risk threshold reconfiguration was required, demonstrating modular cross-sector portability with 60% lower development overhead compared to bespoke implementations [5].

C. On-Premises LLM Performance and Privacy Assurance

An on-premises LLM evaluation across five open-source models in an Oracle APEX 24.2 environment [6] demonstrated that LLaMA 3 and Gemma 7B provide the most favorable performance-to-hardware-cost ratios. Ollama-served local endpoints achieved response latencies of 800ms–1.4 seconds for standard clinical documentation tasks, competitive with cloud-based commercial APIs. Network traffic analysis confirmed zero PHI egress across all LLM inference sessions. Cost analysis projected ROI within 6–12 months for high-inference-volume organizations. Gemma 7B's 8GB VRAM footprint enables deployment on standard server hardware, substantially reducing adoption barriers for small public health agencies.

D. Integration Synergies and Emergent Capabilities

The unified framework produced integration synergies not present in either empirical contribution alone. The on-premises LLM layer, connected to the authentication audit log via PL/SQL procedures, enables natural-language security intelligence queries entirely within the organizational network boundary. A particularly significant emergent capability was LLM-assisted federated annotation: the on-premises Mixtral instance classified ambiguous edge-case authentication events using SHAP rationale context, improving annotation throughput by 40% compared to fully manual review, addressing the annotation bottleneck that frequently constrains federated retraining cadence in deployed healthcare systems.

E. Ethical and AI Safety Considerations

The framework engages AI ethics and safety concerns at multiple levels. The XAI Rationale Engine addresses the ethical obligation to provide meaningful explanations for automated decisions affecting clinicians' access to clinical systems. The human-in-the-loop override interface preserves clinician agency in contesting access decisions. Emergency override provisions directly encode the patient safety imperative that automated authentication barriers must never obstruct care delivery in life-threatening situations. The on-premises LLM architecture addresses data sovereignty as a fiduciary obligation: institutions handling population health data bear ethical responsibility that cloud-based AI transmission violates, regardless of contractual agreements.

F. Limitations

Several limitations merit acknowledgment. SHAP explanations are locally faithful but not globally consistent across all prediction scenarios, requiring ongoing validation against held-out audit samples. Federated coordination overhead adds approximately 10% to computational cost compared with centralized architectures. External MFA delivery dependency (Twilio) introduced latency variability during peak load. BLOOM 176B's hardware requirements render it impractical for standard agency infrastructure. Pilot deployment was limited to a single public health organization; generalizability to acute care environments warrants dedicated evaluation.

5. Future Recommendations

Five high-priority research directions emerge from this work. First, autonomous policy optimization through reinforcement learning would enable the authentication controller to dynamically adjust risk thresholds based on user role, time of day, and organizational context, learning from annotated historical decisions to minimize both false-positive and false-negative rates simultaneously.

Second, multimodal XAI integration, extending SHAP attribution to behavioral sequence data, keystroke dynamics, mouse movement patterns, and interaction timing, would substantially improve the detection of advanced persistent threats that adapt to evade feature-level anomaly detection. Temporal SHAP methods offer particular promise for the sequential nature of behavioral biometric data in healthcare authentication.

Third, cross-platform authentication federation via the OpenID Connect standard would enable shared risk scores across Oracle APEX and non-APEX enterprise systems, such as Epic EHR and Salesforce Health Cloud, reducing redundant MFA challenges for practitioners traversing multiple system boundaries during a single shift.

Fourth, the on-premises LLM layer warrants extension to Retrieval-Augmented Generation (RAG) architectures grounded in real-time organizational data, clinical guidelines, policy documents, and threat intelligence feeds via vector-indexed Oracle Database retrieval, improving factual accuracy without increasing PHI exposure risk.

Fifth, edge computing deployment of the authentication scoring subsystem, distributing local inference to point-of-care devices rather than central servers, would improve resilience and reduce latency in bandwidth-constrained field environments such as mobile health units and rural clinic sites.

6. Conclusion

This paper presents a unified Explainable AI security framework that integrates federated risk-based authentication and privacy-preserving on-premises LLM deployment into a cohesive, compliance-native paradigm for regulated environments. The framework's primary architectural contribution, the SHAP-based XAI Rationale Engine, resolves the interpretability gap that has prevented federated authentication systems from satisfying HIPAA audit trail documentation requirements, while enabling the on-premises LLM layer to generate narrative security intelligence within the organizational network boundary.

Empirical evidence from a live public health deployment demonstrates that the FL-RBA component achieves 95% high-risk interception, a 2.5% false-positive rate, and a 1.1-second average decision latency, substantially superior to traditional authentication approaches across all dimensions, as quantified in Table II. SHAP-generated rationale strings increased security team confidence in blocking decisions by 34% and fully satisfied HIPAA audit requirements for control documentation. On-premises LLM deployment confirmed zero PHI egress while delivering competitive clinical intelligence capabilities.

The emergent capability of LLM-assisted federated annotation improved annotation throughput by 40% over fully manual review, thereby contributing a methodological advance to the federated learning feedback literature. From an AI ethics and safety perspective, the framework demonstrates that the tension between AI capability adoption and data sovereignty obligations in regulated sectors is resolvable through deliberate architectural design. Institutions that treat privacy, explainability, and compliance alignment as first-class design constraints can achieve both the operational intelligence benefits of AI and the data governance obligations that underpin public institutional trust.

7. References

[1] IBM Security, "Cost of a Data Breach Report 2024," IBM Corp., Armonk, NY, USA, 2024. [Online]. Available: https://www.ibm.com/reports/data-breach

[2] M. Elgan, "Cost of a Data Breach in the Healthcare Industry," IBM Think Insights, 2025. [Online]. Available: https://www.ibm.com/think/insights/cost-of-a-data-breach-healthcare-industry

[3] C. D. Hylender, P. Langlois, A. Pinto, and S. Widup, "2025 Data Breach Investigations Report," Verizon Business, Basking Ridge, NJ, USA, 2025.

[4] R. Murray-Watson, "State of Healthcare Cybersecurity," The HIPAA Journal, 2025. [Online]. Available: https://www.hipaajournal.com/healthcare-cybersecurity/

[5] A. Syed, "Dynamic Risk-Based Authentication Using AI Scoring Models in Healthcare Applications," J. Artif. Intell. Cloud Comput., vol. 4, no. 5, pp. 1–10, Oct. 2025, doi: 10.47363/JAICC/2025(4)492.

[6] A. Syed, "Low-Code, High Privacy: Leveraging Open-Source LLMs for On-Premises AI in Oracle APEX," Int. J. Multidiscip. Res. (IJFMR), vol. 6, no. 5, Sep.–Oct. 2024.

[7] S. Wiefling, L. Lo Iacono, and M. Dürmuth, "Is This Really You? An Empirical Study on Risk-Based Authentication Applied in the Wild," in IFIP Advances in Information and Communication Technology. Cham: Springer, 2019, pp. 134–148.

[8] J. Singh, C. Patel, and N. K. Chaudhary, "Resilient Risk-Based Adaptive Authentication and Authorization (RAD-AA) Framework," in Proc. ICISPD. Singapore: Springer Nature, 2023, pp. 371–385.

[9] A. S. Chauhan and D. K. Kumar, "Adaptive Authentication Using Machine Learning," in Proc. Int. Conf. Innovative Computing and Communication, 2024, doi: 10.2139/ssrn.4932261.

[10] A. Wairagade, A. Ahuja, and N. Gupta, "Comprehensive Comparative Assessment of AI Driven Adaptive and Risk Based Authentication Strategies in Cloud Computing," in Proc. 2024 Int. Conf. ITIKD, IEEE, 2025, pp. 1–6.

[11] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. Y. Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," in Proc. 20th Int. Conf. Artif. Intell. Stat. (AISTATS), 2017.

[12] H. Fereidouni, A. S. Hafid, D. Makrakis, and Y. Baseri, "F-RBA: A Federated Learning-based Framework for Risk-based Authentication," arXiv:2412.12324, Dec. 2024.

[13] C. Mazzocca, N. Romandini, M. Colajanni, and R. Montanari, "FRAMH: A Federated Learning Risk-Based Authorization Middleware for Healthcare," IEEE Trans. Comput. Social Syst., vol. 10, pp. 1679–1690, 2023.

[14] Y. Baseri, A. S. Hafid, D. Makrakis, and H. Fereidouni, "Privacy-Preserving Federated Learning Framework for Risk-Based Adaptive Authentication," arXiv, 2025.

[15] A. Ali, V. Snášel, and J. Platoš, "Health-FedNet: A Privacy-Preserving Federated Learning Framework for Scalable and Secure Healthcare Analytics," Results Eng., vol. 27, Art. no. 106484, 2025.

[16] U. Kose, "A Futuristic View on Explainable Artificial Intelligence," in Proc. INFUS 2021, Lecture Notes in Networks and Systems. Cham: Springer, 2021.

[17] A. Adadi and M. Berrada, "Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)," IEEE Access, vol. 6, pp. 52138–52160, 2018.

[18] S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," in Proc. 31st Conf. Neural Inf. Process. Syst. (NeurIPS), 2017, pp. 4765–4774.

[19] T. Miller, "Explanation in Artificial Intelligence: Insights from the Social Sciences," Artif. Intell., vol. 267, pp. 1–38, 2019.

[20] D. Saraswat et al., "Explainable AI for Healthcare 5.0: Opportunities and Challenges," IEEE Access, vol. 10, pp. 84486–84517, 2022.

[21] M. J. Al Ansari, Y. Al Ahmed, and H. H. El Bahnaswi, "Balancing Usability and Protection in AI and Data Security," in Proc. 2024 11th Int. Conf. Software Defined Systems (SDS), IEEE, 2024, pp. 80–88.

[22] Meta, "Introducing Meta Llama 3," Meta AI Blog, Apr. 2024. [Online]. Available: https://ai.meta.com/blog/meta-llama-3/

[23] A. Q. Jiang et al., "Mixtral of Experts," arXiv:2401.04088, Jan. 2024.

[24] Gemma Team, "Gemma: Open Models Based on Gemini Research and Technology," Google Technical Report, Feb. 2024.

[25] S. Pati, "Privacy Preservation for Federated Learning in Health Care," Patterns, vol. 5, Art. no. 100974, 2024.

[26] B. Gubitosa, "Self-Hosted LLM: A 5-Step Deployment Guide," Plural Blog, Jan. 2024. [Online]. Available: https://www.plural.sh/blog/self-hosting-large-language-models/

[27] S. Rose, O. Borchert, S. Mitchell, and S. Connelly, "Zero Trust Architecture," NIST SP 800-207, National Institute of Standards and Technology, 2020, doi: 10.6028/nist.sp.800-207.

[28] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated Machine Learning: Concept and Applications," ACM Trans. Intell. Syst. Technol., vol. 10, no. 2, pp. 1–19, 2019.