Explainable Federated and Blockchain-Enabled Cybersecurity for Smart Healthcare in Smart Cities: A Comprehensive Survey

Sarmad T. Abdul-Samad[1*], Suha Mohammed Hadi[1], Mazin Abed Mohammed[2,3,4]

[1]Informatics Institute for Postgraduate Studies, Information Technology & Communication University

[2]Department of Artificial Intelligence, College of Computer Science and Information Technology, University of Anbar, Anbar 31001, Iraq.

[3]Department of Software Engineering, Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM), Melaka, Malaysia.

[4]Cybersecurity Department, Al-Farabi University, Baghdad 10022, Iraq.

*Corresponding Author: Sarmad.thaer991@rdd.edu.iq

Abstract

The increasing reliance on Cyber-Physical Systems (CPS) and Internet of Medical Things (IoMT) technologies in smart healthcare environments has raised critical concerns regarding data privacy, system security, and the transparency of automated decision-making. Conventional centralized security models are no longer sufficient to meet these demands in distributed and resource-constrained healthcare infrastructures. This paper presents a comprehensive survey of recent cybersecurity frameworks that integrate Explainable Artificial Intelligence (XAI), Federated Learning (FL), and Blockchain technologies to enhance trust and resilience in smart healthcare systems. Research published between 2023 and 2025 is systematically reviewed and classified using a novel taxonomy that groups existing studies into XAI-based, FL-based, hybrid, and fully integrated approaches. The surveyed frameworks are critically analyzed across key security dimensions, including privacy preservation, interpretability, resistance to tampering, robustness against adversarial attacks, scalability, and suitability for real-time deployment. The analysis indicates that while hybrid approaches partially address specific security challenges, they often involve trade-offs between transparency, privacy, and integrity. Moreover, only a limited number of studies attempt a unified integration of all three technologies. These findings highlight a clear research gap and underline the need for lightweight, scalable, and privacy-preserving security frameworks to support trustworthy Healthcare 5.0 systems.

Keywords

Explainable Artificial Intelligence, Federated Learning, Blockchain, Healthcare Cybersecurity, Cyber-Physical Systems, Healthcare 5.0, Smart Cities.

1. Introduction

The rapid evolution of smart cities has accelerated the adoption of advanced digital technologies into healthcare infrastructure, leading to the emergence of interconnected and intelligent medical environments. Central to this transformation are Cyber-Physical Systems (CPS) and the Internet of Medical Things (IoMT), which enable continuous patient monitoring, real-time clinical decision-making, and automated healthcare services through distributed sensing, communication, and control mechanisms [1],[2]. These technologies play a critical role in improving healthcare efficiency, responsiveness, and personalization, particularly within the vision of Healthcare 5.0, where intelligent systems are expected to operate autonomously while maintaining high levels of safety and trust [3].

Despite these advantages, CPS and IoMT-based healthcare systems introduce serious cybersecurity and privacy challenges. Sensitive medical data are continuously generated, transmitted, and processed across heterogeneous and distributed infrastructures, increasing the risk of unauthorized access, data manipulation, and service disruption [1],[4]. Conventional centralized security models are increasingly inadequate for addressing these risks, as they rely on centralized data aggregation and control, creating single points of failure and limiting scalability in large-scale smart city environments [5].

Artificial Intelligence (AI) techniques have been widely adopted to enhance security monitoring and threat detection in smart healthcare systems. Although AI-driven models demonstrate strong detection performance, their decision-making processes are often opaque, which undermines trust and accountability in safety-critical healthcare applications [4],[5]. To address this limitation, Explainable Artificial Intelligence (XAI) has emerged as a promising approach by providing interpretable explanations for model predictions and security alerts, thereby supporting transparency and regulatory compliance [6],[7].

In parallel, Federated Learning (FL) has gained increasing attention as a privacy-preserving learning paradigm for distributed healthcare environments. FL enables collaborative model training without sharing raw patient data, significantly reducing data exposure risks and supporting compliance with strict privacy regulations [1],[8]. However, FL-based systems face several challenges, including vulnerability to poisoning attacks, difficulties in handling non-independent data distributions, communication overhead, and the lack of intrinsic mechanisms for verifying the integrity of distributed model updates [3],[9].

To further enhance trust and data integrity in distributed healthcare systems, Blockchain technology has been explored as a decentralized security layer. Blockchain provides immutable data storage, transparent auditing, and secure access control through cryptographic mechanisms and smart contracts, making it a strong candidate for safeguarding healthcare data and system operations [10],[11]. Nevertheless, blockchain alone lacks intelligent threat detection and explainable decision-making capabilities, limiting its effectiveness when deployed in isolation [12].

Recent research has therefore focused on integrating XAI, FL, and Blockchain technologies to address complementary aspects of healthcare cybersecurity. While several hybrid frameworks combining two of these technologies have been proposed, fully integrated solutions that simultaneously ensure privacy preservation, decision transparency, and data integrity remain limited. This gap highlights the need for a systematic and structured analysis of existing approaches to guide future research toward trustworthy and scalable smart healthcare systems [2],[12]. Motivated by the above challenges, this paper makes the following key contributions:

1. To provide a focused and up-to-date review of cybersecurity frameworks for smart healthcare CPS environments, covering recent studies that employ XAI, FL, and Blockchain technologies.
2. To introduce a structured taxonomy that categorizes existing literature into XAI-based, FL-based, hybrid, and fully integrated security frameworks.
3. To critically analyze existing approaches across key security dimensions, including privacy preservation, interpretability, tamper resistance, robustness to attacks, scalability, and real-time applicability.
4. To highlight the scarcity of fully integrated XAI–FL–Blockchain frameworks and identifies unresolved trade-offs in current hybrid solutions.
5. To outline practical directions toward lightweight, scalable, and privacy-preserving security architectures suitable for Healthcare 5.0 environments.

2. Conceptual Background

Cyber-Physical Systems (CPS) are widely recognized as foundational technology in the modern intelligent environment, such as smart cities. In the context of smart healthcare, these distributed networks deliver highly responsive and autonomous services. CPS architectures integrate sensing, computation, communication, and control into a single intelligent framework to support continuous monitoring, real-time clinical decisions, and emergency response coordination. The integration of

CPS into healthcare is transforming the shape of clinical operations by enabling AI-enabled decision-making and automating critical medical interventions [13]. Due to CPS's ability to support real-time sensing and intelligent control mechanisms, recent studies consider it like a backbone of next-generation medical platforms [14],[15].

This technological evolution is very close to the vision of **Healthcare 5.0**, where in medical environments relay on distributed computing, automation, and AI-empowered decision making. Healthcare 5.0 focuses on real-time intelligence, personalized care and seamless integration of digital systems with physical medical processes. A core component of CPS within the healthcare framework is **Internet of Medical Things (IoMT)**, encompassing a diverse array of edge devices such as wearable sensors, personal health devices (PHDs), smart monitoring implants, and connected diagnostic systems [16,17]. Research emphasizes that IoMT technologies generate a flow of physiological data streams that must be securely and rapidly processed to support real-time healthcare [13]. Other studies have proven that IoMT systems play a critical role in preventive medicine and personalized care through their ability to receive an uninterrupted flow of patient data [18]. However, many cybersecurity challenges have been noticed in IoMT infrastructures. An observation reported by Putta, S. R. in [19] shows that IoMT networks are highly susceptible to many types of attacks, such as unauthorized access, spoofing attacks, malware propagation, and manipulation of sensor outputs, due to the limited computational capabilities of medical devices. Reinforcing this view, it noted that wearable healthcare devices often lack robust security safeguards, which may expose data and operations of the system to significant cyber risks [20].

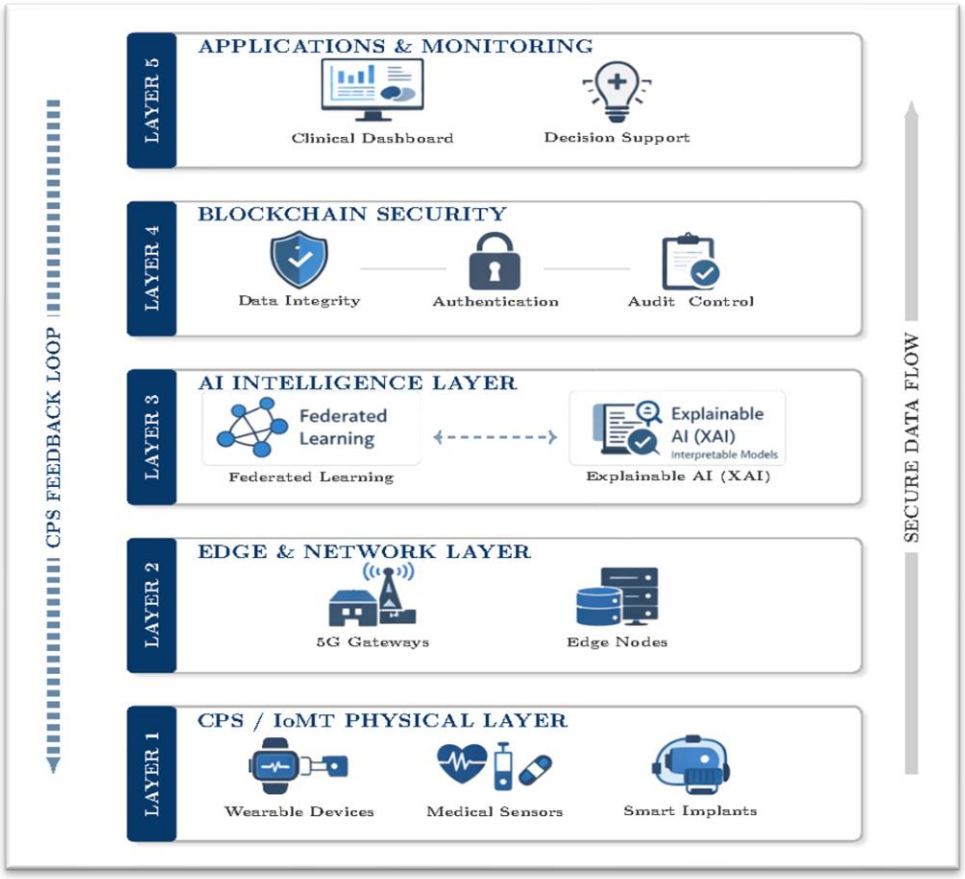

Figure 1. A Multi-Layered Conceptual Architecture for Secure and Explainable Healthcare Cyber-Physical Systems (CPS).

1.1 Security Challenges in CPS and IoMT

As established previously, while CPSs offer significant benefits within healthcare frameworks, they remain highly vulnerable to cyberattacks because of their distributed architecture, heterogeneous medical devices, and continuous connectivity. Previous studies show that CPS in healthcare are exposed to a wide range of attacks, including distributed denial-of-service (DDoS), data poisoning, replay attacks, man-in-the-middle interception, device spoofing, and model inversion techniques targeting AI-driven diagnostic models. Such cyberattacks can result in dangerous outcomes such as manipulation of sensor readings, medical record alternation or deletion, disruption or interruption of communication with critical wearable-sensor devices, or failure in operation of essential hospital equipment [14],[21].

In addition, IoMT-enabled CPS confront compounded security limitations due to their constrained processing capabilities and inadequate protective mechanisms. Various studies like [22] and others highlight that real-time healthcare environments cannot depend on traditional heavyweight cybersecurity techniques, as these approaches affect the performance of the system by increasing the latency and may disturb the continuous need for system availability. As healthcare systems in smart cities continue to develop toward interconnected and autonomous Healthcare 5.0 ecosystems, ensuring cyber resilience for CPS and IoMT remains one of the most persistent and unresolved challenges [23].

1.2 Explainable Artificial Intelligence (XAI) in Healthcare Cyber-Security

Explainable AI (XAI) can be defined as a set of techniques that can provide interpretability and transparency for the judgment or prediction of the AI model. In contrast to the traditional black-box AI models, XAI solutions deal with the deficiency of interpretability that can be extremely problematic in healthcare settings, where cybersecurity decision-making, as well as clinical validation, entails a clear and actionable explanation of the ways model generated a certain output. XAI is applicable in the healthcare cybersecurity setting so that the user can comprehend the way the model works, and why it gave a particular output or classification [24],[25].

Many of them point out the fact that explainability is crucial to the development of security systems that have a direct effect on the clinical workflow [24]. On the same note, the recent studies emphasize the fact that XAI enhances the functionality of the anomaly detection systems by enhancing the legibility of security alerts and their relevance in the context of security operators [25]. Explainable techniques, such as SHAP (SHapley Additive exPlanations), LIME (Local Interpretable Model-Agnostic Explanations), Integrated Gradients, and attention-based visualization have been used on intrusion detection systems in many explainable ways that enhance transiency. As an example, when a cyberattack is detected by an anomaly detection algorithm, the XAI can show the characteristics that are typically influencing the decision, which can include unusual packet rates or suspicious IP activity [26].

In general, XAI can enhance trust and responsibility in AI-based healthcare systems by allowing stakeholders and cybersecurity professionals to reject and interpret the behavior of algorithms. Although this has benefits, the majority of XAI-based security models are based on centralized data processing and lack inherent support of data integrity and adversarial manipulation resistance.

1.3 Federated Learning (FL) for Privacy-Preserving Healthcare Security

As healthcare frameworks generate a massive volume of sensitive data, such as medical records and IoMT sensors measurements, conventional centralized machine learning approaches become inefficient as a result of strict privacy requirements, regulatory limitations, and bandwidth constrains. Federated Learning (FL) has emerged as an alternative technology that allows multiple healthcare devices to train AI models collaboratively without distribution of system or patient raw data. As highlighted in [27], FL works in a decentralized architecture that safeguards the confidentiality of data while supporting large-scale analytics across geographically distributed environments within contexts such as smart cities.

FL plays a critical role in detecting cybersecurity threats in smart healthcare while maintaining strong compliance with privacy regulations such as GDPR and Health Insurance Portability and Accountability Act (HIPAA). Study [28] determines that FL supports cross-institutional IDS to learn from diverse attack patterns without exposing local datasets. Similarly, research highlights that FL decreases the risk of centralized data breaches by not sharing training data and keeping it at the source, which is particularly important for sensitive biomedical information [29].

Despite these advantages, FL presents several technical and security challenges. Studies show that federated models are vulnerable to an attack set including inference attacks, poisoning attacks, and malicious client updates [30],[31]. Resource limitations of IoT and IoMT devices can also create computational problems during local training. Furthermore, studies noted that highly heterogeneous (non-IID) healthcare data can negatively impact global model performance; thus, there is a need for advanced aggregation methods and robust optimization strategies to preserve accuracy and reliability. As smart cities healthcare systems shift toward distributed, interconnected Healthcare 5.0 ecosystems, FL provides what can be considered as a foundation to build privacy-preserving, scalable, and resilient security frameworks across CPS infrastructures. However, while FL effectively protects data locality, it does not inherently provide transparency into model decision-making or robust mechanisms to verify the integrity of model updates, which limits its suitability for trust-critical healthcare applications [32]. With healthcare systems producing a vast amount of sensitive information, including medical records and IoMT sensor measurements, traditional centralized machine learning models cannot be effectively used due to the high demands of privacy, regulatory restrictions, and bandwidth constraints. A new technology, Federated Learning (FL), has appeared to enable several healthcare devices to jointly train AI models without access to system or patient raw data. As emphasized in [27], FL operates in a decentralized

framework which ensures the confidentiality of information and aids in large-scale analytics in geographically distributed settings in scenarios like smart cities.

FL is of paramount importance to identify cybersecurity threats in smart healthcare without being overly aggressive in compliance with privacy standards like GDPR and Health Insurance Portability and Accountability Act (HIPAA). Research [28] concludes that FL assists cross-institutional IDS to learn various patterns of attacks without exposing local datasets. In the same way, studies emphasize that FL reduces the threat of centralized data breaches due to the lack of sharing of training data and retention at the source, especially with sensitive biomedical data [29].

Although there are these benefits, FL is faced with a number of technical and security challenges. Studies show that federated models are vulnerable to an attack set including inference attacks, poisoning attacks, and malicious client updates [30],[31]. During local training, computational issues can also arise due to resource constraints of IoT and IoMT devices. Moreover, researchers observed that a high level of heterogeneous (non-IID) healthcare data has a detrimental effect on the performance of global models; therefore, more sophisticated aggregation tools and effective optimization procedures are required to maintain the accuracy and reliability. With smart cities healthcare systems moving forward to distributed, networked Healthcare 5.0 ecosystems, FL offers what can be regarded as a base to develop privacy-preserving, scalable, and resilient security frameworks across CPS infrastructures. Nevertheless, although FL is an effective defence of data locality, it does not automatically offer visibility of model decision-making or other strong systems to confirm the integrity of model updates, which restricts its application in trust-critical healthcare systems [32].

1.4 Blockchain Technology for Trust and Tamper-Proof Healthcare Security

In distributed healthcare environments, devices are continuously sending critical medical data between each other, trust, integrity, and traceability must be ensured. **Blockchain technology** has arisen as a robust candidate for securing these interconnected systems due to its decentralized architecture, immutable ledger, and cryptographic mechanisms. Blockchain offers tamper-resistant data storage that prevents any unauthorized modification of the data in the environment where even minor alterations to it can cause a real problem [33].

It is also increasing the efficiency of authentication and access control within such ecosystems. The researchers explain the smart contracts and show that they can automatically enforce secure data-

sharing policies, verify device identity, and regulate permissions without requiring centralized control. This model significantly decreases the risk of insider threats and prevents unauthorized data access, which is a major problem in healthcare systems [34]. Additionally, blockchain-based improve the cyber-security of edge, fog, and cloud architectures in healthcare frameworks by providing integrity assurance and decentralized monitoring of device operation [35].

2. Literature survey

The convergence of Cyber-Physical Systems (CPS), the Internet of Medical Things (IoMT), and AI-driven decision support, all of these technologies has catalyzed unprecedented advancements in smart healthcare systems. Consequently, these developments have amplified the critical need for computational models that guarantee transparency, safety, and rigorous privacy preservation. As established before, traditional centralized machine learning has many limitations in healthcare smart environments, such as the risk of data exposure, prone to poisoning attacks, and the problem of interpretability. To circumvent these structural deficiencies, the vanguard of modern cybersecurity research is strategically pivoting toward the following paradigms:

- **Explainable AI (XAI)** for interpretable and transparent systems.
- **Federated Learning (FL)** for decentralized, privacy-preserving model training.
- **Blockchain** for tamper-proof logging and trust access control management.
- **Hybrid systems** combining two or more of these technologies for more powerful cybersecurity systems.

Based on literature, this study categorizes and classifies related works into four groups.” Figure 2” illustrates the taxonomy of the surveyed frameworks, categorizing them into four main groups: XAI-only, FL-only, Hybrid, and Triple-Integrated systems.

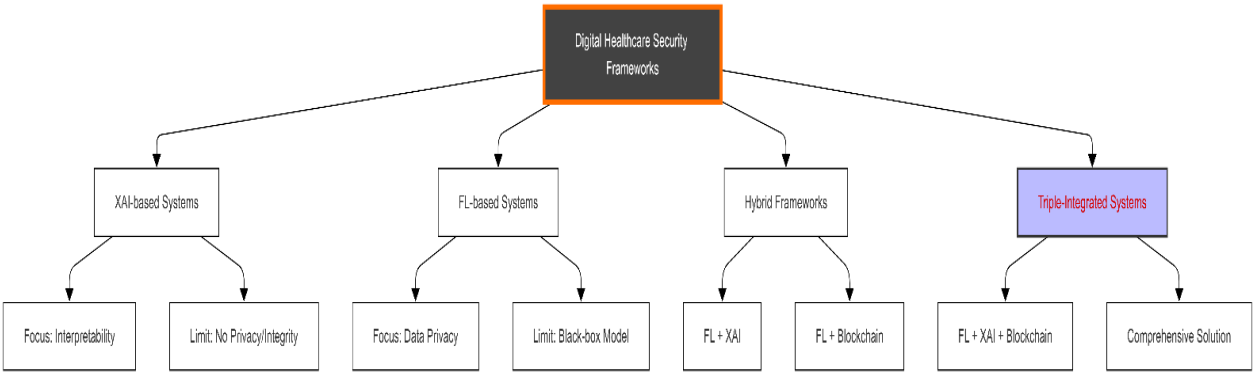

Figure 2: Taxonomy of Research on Explainable, Federated, and Blockchain-based Healthcare Security Frameworks.

The classified groups are defined as follows:

1. XAI-only frameworks,
2. FL-only frameworks,
3. Hybrid frameworks combining two technologies,
4. Frameworks combining all of the three technologies FL, XAI, and Blockchain.

2.1. Categorization of Existing Literature

The reviewed studies are classified into four primary architectural groups derived from our survey:

2.1.1 XAI-based Systems

Explainable Artificial Intelligence (XAI) has arisen as a critical component in modern security systems for CPS and smart environments, which provides transparency and interpretability in security decisions that were previously opaque. Understanding model behavior to aid in decision-making has been the subject of numerous recent studies. "Table 1" summarizes the key XAI-based security frameworks, highlighting their methodologies, results, and interpretability approaches.

Table 1: Explainable AI (XAI)-Based Systems

Author	Year	Focus Area	Method / Framework	result
Shafique Ahmed Memon, Uffe Kock Wiil, Mutiullah Shaikh [36]	2023	**IoMT Security** – Explainable Intrusion Detection (XID)	- **PSO (Particle Swarm Optimization)** for feature selection (reduced features from 32 → 22) - Ten ML classifiers (LR, NB, KNN, SVM, MLP, DT, ET, RF, XGB, SGD) - **LIME** for interpretability	- Accuracy between **97% and 99%** across models - Best: **ET, RF, XGB** (99% accuracy, F1=0.99, Precision=0.99, Recall=0.99) - LIME provided interpretable feature importance for attack/no-attack classification.
Chathuranga Sampath Kalutharage, et al. [11]	2023	**IoT Security** – DDoS Attack Detection with XAI	- **Autoencoder + XAI (SHAP)** for anomaly detection & explanation - Feature selection for DDoS flows (state-exhaustion, application-layer and volumetric features) - Threshold-based classification for attack certainty	- Accuracy: **0.98–1.0** for HULK attacks (No Defense, Evasive, Reqtimeout) vs DT (0.97), RF (0.98), DNN (0.66) - Provided **feature-level explanations** (e.g., flow packets/s, bwd packets/s, packet length features) - More reliable anomaly-to-attack mapping than black-box ML
Marwa Keshk, et al. [37]	2023	IoT Security – XAI for Intrusion Detection	- **LSTM-based IDS** - Novel **SPIP Framework** (SHAP, PFI,	- **NSL-KDD:** Acc=0.811, Prec=0.921, Rec=0.732, F1=0.815 - **UNSW-NB15:** Acc=0.866, Prec=0.811,

			ICE, PDP) for local & global explanations	Rec=0.988, F1=0.891 - **ToN-IoT:** Acc=0.873, Prec=0.784, Rec=0.880, F1=0.829 - Feature subsets from **SHAP/PFI improved performance** & reduced training/detection time
Marios Siganos, et al. [38]	2023	**IoT & IIoT Security** – Explainable IDS with ML/DL	- AI-powered IDS for IoT/IIoT - **ML/DL models**: NB, SVM, DT, RF, XGBoost, AdaBoost, Logistic Regression, QDA, DNN - **SHAP** for explainability (summary & waterfall plots) - IDS architecture: data capture → preprocessing → detection → explainability → notification	- **IEC 60870-5-104:** RF best (Acc ≈ 0.85, F1 ≈ 0.85); DT/XGBoost close - **CIC-IoT-2022:** XGBoost top (F1 ≈ 0.999), RF/DT/DNN also >0.99 - SHAP visualizations highlighted top 10 features per dataset, improved trust & interpretability
Mohammad Algarni & Shailendra Mishra [7]	2024	Secure and reliable Explainable Artificial Intelligence (XAI) framework for Smart City applications	Machine Learning (Logistic Regression with LIME, Random Forest with SHAP) - Integration of XAI interpretability techniques	LIME + Logistic Regression: ~99.9% accuracy, precision, recall, F1-score - SHAP + Random Forest: ~85% accuracy, precision 82.3%, recall 85.2%, F1-score 82.3% - Models provided both global (SHAP) and local (LIME) interpretability
Abdul Lateef Haroon P.S. & Hareesh K N [39]	2024	Explainable AI (XAI) empowered secured edge-based healthcare systems	Deep Learning (LSTM networks) - Integrated with XAI for interpretability - Edge computing framework with IoT-enabled devices	Proposed LSTM + XAI achieved 99% accuracy, precision 98.9%, recall 98.85%, specificity 99.1%, F1-score 99% - Outperformed GRU, RF, DT, and SVM models
Islam Elgarhy, et al. [40]	2024	**Smart Grid Security** – Electricity Theft (ET) Detection with XAI	- Hybrid **Deep Auto-Encoder (DAE) + One-Class SVM (OCSVM)** - Uses **SHAP explanations** for consumption readings - Cluster-based detector design to reduce generalization	- Proposed detector achieved **ACC up to 97.79%**, **DR = 100%**, and improved robustness vs. FGSM, BIM, CW, ZOO, DeepFool attacks - Outperformed baseline DAE anomaly detector and binary detectors - Achieved high **AUC-ROC (0.978)** and **AUC-PR (0.979)**
Bhawana Sharma, Lokesh Sharma, Chhagan Lal, Satyabrata Roy [26]	2024	IoT Security – Intrusion Detection with Explainable AI	- **Deep Neural Network (DNN)** - **1D-CNN & 2D-CNN** - **Filter-based Feature Selection (Correlation)** - **XAI Methods: LIME (local), SHAP (local + global)**	- **NSL-KDD:** Accuracy = 0.993 (DNN), 0.992 (1D-CNN), 0.994 (2D-CNN) - **UNSW-NB15:** Accuracy = 0.80 (DNN), 0.80 (1D-CNN), 0.81 (2D-CNN) - XAI identified key features (e.g., serror_rate in NSL-KDD, dttl in UNSW-NB15) for trust & interpretability
Osvaldo Arreche, Tanish Guntur, Mustafa Abdallah [41]	2024	Network Security – Intrusion Detection with Explainable AI	- **Seven AI models** (RF, DNN, AdaBoost, MLP, KNN, SVM, LightGBM) - **XAI methods: SHAP & LIME** (global & local explanations) - End-to-end framework: preprocessing → model training → XAI explanations → feature extraction	- RF, DNN, MLP achieved **near-perfect accuracy (≈0.99)** on RoEduNet-SIMARGL2021 - ADA, KNN achieved **0.99 accuracy** on CICIDS-2017 (all features) - On NSL-KDD, **MLP best performer** with high accuracy/F1; others varied - XAI explanations identified **top attack-specific features** & reduced feature sets improved detection

Mohammad Taufik et al. [4]	2025	Adversarial attack detection in CPS with explainability and real-time constraints	Hybrid framework integrating CNN, SHAP feature interpretation, and rule-based reasoning	Accuracy 97.25%, Precision 96.80%, Recall 95.90%, F1-score 96.35%, Inference time +8.5% over baseline CNN (20.1 ms vs 18.5 ms)
Michael Georgiades & Faisal Hussain [5]	2025	Explainable AI (XAI) for **Cross-Layer Intrusion Yousif Hosain Detection in IoMT**	Six-stage methodology (data collection, preprocessing, feature selection, model training, performance comparison, explainability) - K-Means + PCA for attack categorization - SHAP for feature prioritization - PDP & ALE for feature interaction analysis	SHAP-based feature selection improved interpretability and reduced computational overhead - Random Forest, KNN, and DT achieved high precision, recall, and F1-score using optimized features - PDP & ALE visualizations explained cross-layer interactions (biosensor + network features)
Yousif Hosain & Muhammet Çakmak [42]	2025	Explainable Intrusion Detection for **IoMT Security**	Hybrid Random Sampling for class imbalance - Recursive Feature Elimination (RFE) for feature selection - Optimized **XGBoost** classifier - Explainability via **SHAP** & **LIME**	Achieved **99.22% accuracy**, 98.35% precision, 99.91% recall, 99.12% F1-score, and 100% ROC-AUC - Outperformed ML baselines (RF, DT, KNN, SVM, etc.) and prior IDS frameworks
Adel Alabbadi, Fuad Bajaber [43]	2025	**IoT Security / Cross-Domain Intrusion Detection** – XAI + DL	- **X-FuseRLSTM Framework** - Dual-path feature extraction (Deep Encoder + Sparse Transformer) - Feature fusion + PCA reduction - Classification with **Residual LSTM + DNN** - Explainability via **SHAP & LIME**	- Accuracy: **99.40% (TON_IoT)**, **99.72% (NSL-KDD)** - Accuracy: **97.66% (CICIoMT 19-class)**, **98.05% (CICIoMT 6-class)** - Outperformed baselines (CNN, LSTM, PSO-LSTM, Hybrid CNN-LSTM, GRU models) - Robust domain generalization with XAI-based transparency
Yakub Kayode Saheed, Joshua Ebere Chukwuere [44]	2025	CPS-IIoT Security & Privacy-Preserving IDS	- **BiLSTM + Scaled Dot-Product Attention** - **Pearson Correlation + Agglomerative Clustering** for privacy-preserving feature selection - **SHAP (XAI)** for explainability	- **UNSW-NB15:** Accuracy = 99.60%, AUC = 100%, Recall = 97.98%, Precision = 100%, F1 = 98.23%, MCC = 96.54% - **X-IIoTID:** Accuracy = 99.99%, AUC = 100%, Recall = 99.97%, Precision = 99.98%, F1 = 99.87%, MCC = 99.98% - Outperformed **TinyLSTM (+2.7% accuracy)** with only **+11% energy cost**

The experimental data present in "Table 1" indicates that integrating Explainable AI (XAI) into security frameworks significantly enhances the transparency without degrading the computational performance. Notably, frameworks using techniques like LIME and SHAP achieve high classification accuracies, generally ranging from 97% to over 99% across most studies. For instance, using LIME for interpretability does not affect the accuracy and keeping it at roughly 99% and successfully reduces the feature set from 32 to 22 [32]. Similarly, using SHAP to provide feature explanation for DDoS attacks, achieving a near-perfect classification rate [11]. These findings show that XAI successfully bridges the gaps between algorithmic complexity and human understanding, allowing the security experts discern exactly why a specific alert, such as unusual packet rates is triggered. Conversely, a critical architectural limitation is that the majority of XAI-centric frameworks operate on centralized topologies, which inherently introduces severe privacy

risks. As advanced in “Table 1”, these systems require to aggregated a sensitive data in a central server for training, which contradicts the privacy-preserving requirements of the smart city healthcare frameworks infrastructure. Furthermore, while XAI clarifies the decision-making process, it lacks inherent mechanisms to ensure data integrity. The current frameworks do not provide tamper-proofing guarantees for logs or model updates. So, the “Explainable” systems still vulnerable to internal manipulation where the explanation itself can be changed without having a layer of security safeguarding it.

2.1.2 Federated Learning-Based Security Frameworks

Federated Learning (FL) serves as a pivotal technology for privacy-sensitive healthcare systems, as it enables decentralized model training without exposing raw patient data. These FL-centric architectures effectively address critical vulnerabilities related to data confidentiality while facilitating robust, distributed anomaly detection across heterogeneous IoMT and CPS nodes. “Table 2” presents a comprehensive summary of recent FL-driven security studies and their respective methodologies.

Table 2: Federated Learning-Based Security Frameworks

Author	Year	Focus Area	Method / Framework	result
Maryum Butt, et al. [45]	2023	Smart Healthcare – Federated Learning with Privacy Preservation	- **Fog-based Federated Learning (FL)** architecture for **COVID-19 chest X-ray screening** - Uses **CNN** for local hospital training - **FedAvg algorithm** for aggregation - Preprocessing: U-Net lung segmentation + augmentation - Privacy: no raw data sharing, decentralized training	- **FL model outperformed local CNN models** across accuracy, precision, recall, F1 (up to **93–94%**) - Reduced false positives & false negatives compared to local models - Robust generalization across multiple hospital datasets
Jamshed Ali Shaikh, et. Al.[46]	2024	**IoMT Security** – Anomaly-based Intrusion Detection	- **RCLNet**: RF for feature selection + CNN-LSTM hybrid - Self-Adaptive Attention Layer Mechanism (SAALM) - Focal Loss (FL) for handling class imbalance	- **99.78% accuracy**, Precision 99.53%, Recall 99.83%, F1-score 99.57% - Outperformed baselines (KNN, Tree Classifier, PSO-DNN, FusionNet, GBBOA)
JiaMing Wang, Kai Yang, MinJing Li [47]	2024	**IIoT Security** – FL-based NIDS with Secure Aggregation	- **NIDS-FGPA Framework** - Gradient Similarity Model Aggregation (GSA) - **2D-CNN + BiGRU** deep model - **Paillier Homomorphic**	- Accuracy: **94.5% (Edge-IIoTset)**, **99.2% (CIC-IoT2023)** - Outperformed **FedAvg** (89.1% / 93.4%) and **FedNova** (92.6% / 96.4%) - Communication overhead reduced by **22–51%**

			Encryption for secure parameter sharing	- Strong robustness to missing features & incomplete labels
Shihua Sun, et al.[48]	2024	**IoT Network Security** – Federated Learning for Intrusion Detection	- **FedMADE**: Dynamic Aggregation Method - Device clustering with **DBSCAN** - Class Probability Matrices (CPMs) for performance weighting - Robustness to poisoning attacks	- Up to **71.07% improvement** in minority attack classification accuracy - Robust against data/model poisoning - Only **4.7% (5.03s)** latency overhead vs FedAvg - High F1 (>0.99) in binary detection
Mark Devine et al. [49]	2025	**IoT Security** – Intrusion Detection with Federated ML	- **Federated SVM IDS** - Compared with ANN, Random Forest, Isolation Forest - FedAvg aggregation - Focus on **physical metrics** (delay & memory usage) for IoT feasibility	- Centralized SVM: **99.1% accuracy, 97.9% precision, 99.3% recall, 98.1% F1** - Federated SVM (5 nodes): **97.4% accuracy, 97.6% precision, 97.5% recall, 97.3% F1** - Comparable to ANN & RF; Isolation Forest weaker (~73%) - Memory per node reduced (2919 → 819 MB with 5 nodes) - Training delay competitive with ANN, faster than RF
Sanaa A. Sharaf et.al. [10]	2025	**Adversarial Attack Detection** in IoMT using FL	- **AADF-FLEIoTM**: Adversarial Attack Detection Framework with Federated Learning - Min-Max Normalization - **MPA** (Marine Predator Algorithm) for feature selection - **SA-CNN-BiLSTM + Self-Attention** for classification - **RTH Optimizer** for hyperparameter tuning	- Accuracy: **98.24%** - Precision: 98.25%, Recall: 98.24%, F1-score: 98.24%, MCC: 96.49% - Outperformed baselines: ANN (92.25%), RF (92.57%), SVM (93.34%), ResNet-34 (98.0%)
Van Tuan Nguyen & Razvan Beuran [50]	2025	**IoT Intrusion Detection** – Semi-supervised FL	- **FedMSE Framework** - **SAE-CEN** (Shrink Autoencoder + Centroid one-class classifier) - **MSEAvg** aggregation (performance-based weighting) - Semi-supervised FL with anomaly detection	- Detection accuracy improved from **93.98±2.90% → 97.30±0.49%** under high non-IID - Robust to heterogeneous networks and gateway selection ratios (50% gateways sufficient) - Reduced communication & computational costs - Scales well up to 50 gateways with stable accuracy (~98%)
Ragab et al.[51]	2025	Cyberthreat Detection in IoT-assisted Smart Cities	AAIFLF-PPCD Framework (HHO for feature selection + SSAE classifier + WOA optimization under Federated Learning)	Accuracy 99.47%, Precision 97.20%, Recall 96.84%, F1-score 96.92%, AUC 98.28%; Processing time 4.51s (better than existing models)
Muhammed Rafeeq War, Yashwant Singh, Zakir Ahmad Sheikh [52]	2025	**CPS Security** – Federated Learning for Constrained Devices	- **FedSec-CPS Framework** - Horizontal Federated Learning (HFL) & Vertical Federated Learning (VFL) - Base models: **Random Forest (RF), Artificial Neural Network (ANN)**	- Centralized: RF (99% acc., 34s), ANN (97% acc., 2m51s) - HFL: accuracy ranged **74%–100%**, computation time as low as 2.86–5.56s - VFL: accuracy up to **82%**, time 2.28–3.31s

			- Implemented with TensorFlow Federated (TFF) - Integrated **data encryption, secure aggregation, differential privacy**	- FL reduced computational overhead while preserving privacy

The analysis of "Table 2" shows that Federated Learning (FL) effectively addresses the inherent problem of data confidentiality in traditional machine learning. By keeping raw data on the local devices, many frameworks are successfully adhered to privacy regulations like HIPAA and GDPR [45],[47]. Operationally, these systems have a robust performance; for example, [46] achieved an accuracy of 99.78% using an anomaly-based detection model. Additionally, FL has proven efficient in resource utilization, with studies showing significant reductions in communication overhead up to 51% in some cases by transmitting only model parameters rather than voluminous datasets such as [52].

However, despite these privacy advantages, FL introduces severe vulnerabilities, the "black box" opacity of its models and a high susceptibility to poisoning attacks. In particular, FL provides protection for data locality but offers little visibility into the decision-making behavior of locally trained models, leaving model reasoning opaque and unsuitable for trust-critical clinical contexts. Moreover, the decentralized update mechanism exposes the global model to manipulation, as compromised participants can deliberately distort training through poisoned or malicious updates. Existing studies indicate that, in the absence of explicit integrity validation, such adversarial contributions can propagate unchecked across the federation. While FL effectively reduces direct data exposure, it lacks native mechanisms to authenticate participating clients or to validate the trustworthiness of their updates, underscoring the need for complementary tamper-resistant technologies, such as blockchain-based verification layers.

2.1.3 Hybrid Frameworks

The integration of dual technologies into a hybrid system such as combining Federated Learning (FL) with Blockchain or FL with Explainable AI (XAI), represents an intermediate evolutionary stage in contemporary security architectures. These synergistic combinations concurrently address multiple systemic limitations by enhancing algorithmic interpretability, facilitating distributed model training, and ensuring robust privacy preservation. A comprehensive compilation of recent hybrid

Table 3: Hybrid Frameworks (Two Technologies Only)

Author	Year	Focus Area	Method / Framework	result
Ayoub Si-Ahmed [53]	2024	Explainable ML-based security & privacy for **IoMT** (intrusion/anomaly detection with privacy preservation)	**ANN-based IDS** + **Federated Learning (FL)** for privacy + **XAI (SHAP)** for explanations; compares optimized FL vs. centralized training; details ANN topologies per dataset (e.g., 7 hidden layers for UNSW-NB15; 5 for ToN-IoT & NSL-KDD)	FL achieves **centralized-comparable** performance. Representative best ranges reported: **ToN-IoT:** Acc ≈ 0.980–0.981, AUC ≈ 0.9976–0.9978; **UNSW-NB15:** Acc ≈ 0.988, AUC ≈ 0.9982–0.9984; **NSL-KDD:** Acc up to **0.9905**, AUC up to **0.9973**; **WUSTL-EHMS:** Acc ≈ 0.938, AUC ≈ 0.899–0.900. Tables also report precision/recall/F1, loss, TP/TN/FP/FN, and communication rounds across client counts and fractions.
Daniel Commey, Sena Hounsinou, Garth V. [54]	2024	**Healthcare Data Privacy & Security** – Blockchain + FL + DP	- **FL-DP-Blockchain Framework** - Federated Learning with **Dynamic Personalization** - **Adaptive Noise Distribution** for DP - Blockchain (Ethereum, Ganache, Web3.py, IPFS) for secure aggregation & storage	- Accuracy: **64.50%** (ε = 8.0, 15 rounds) - Strong privacy guarantees vs inference attacks - Blockchain: stable gas consumption (~22,152 units), latency ~6s - IPFS efficient for storing large model updates (~196 MB each)
Radjaa Bensaid, et al. [55]	2024	Smart Healthcare IoMT Security – Intrusion Detection	- Proposed **SA-FLIDS**: Secure & Authenticated Federated Learning-based IDS - Uses **Blockchain + Self-Sovereign Identity (SSI)** for authentication - **Trimmed Mean** aggregation to resist poisoning - **gRPC + TLS** for secure communication	- **Binary classification:** • Edge-IIoTset: 100% Acc., Prec., Rec., F1 • CICIoT2023: 99.1% Acc., Prec. 100%, Rec. 98%, F1 = 99% - **Multiclass classification:** • Edge-IIoTset: Avg. Acc. 93.48%, best classes (DoS/DDoS, MITM, Normal) = 100%, weakest = Malware (74%) • CICIoT2023: Avg. Acc. 92%, perfect for DDoS/DoS/Mirai, weaker on Spoofing (84%)
Jameel Almalki, et. al [56]	2024	**Healthcare 5.0 Security** – IoMT with FL + Blockchain + IDS	- **Proposed Framework**: Blockchain + Intrusion Detection + Federated Learning - **AT-DLM (Actual Time – Deep Learning Model)** for disease prediction - Secure data commit via blockchain & smart contracts - FL for collaborative model training without data sharing	- Disease prediction accuracy: **93.89%** - Intrusion detection success rate: **97.13%** - Stronger privacy/security than centralized ML - Outperformed CNN, SVM, RF, Logistic Regression baselines
Rabia Abid et al. [2]	2024	Adapting Federated Explainable AI for efficient, privacy-preserving e-healthcare—combining FL with the concept of XAI for secure decision-making.	Optimized Federated Averaging (FMLA) with three comparisons: (i) centralized ML baseline, (ii) Local Differential Model (LDM) in FL, and (iii) FL with differential privacy; typical settings include clients X=5, batch size 16, and multiple epochs. The paper discusses measuring "explainability," but does not operationalize a specific XAI algorithm in the experiments	Centralized baseline: after 20 epochs with BS=16, train acc = 95%, test acc = 90% (slight overfitting); confusion matrix provided. Federated (proposed): training/testing accuracy curves reported for X=5, BS=16 (no single final accuracy number); DP degrades accuracy (privacy ε figure). Computation time (Table 3): FedAvg 37.23±1.75s (5 epochs), 68.42±0.89s (10), 99.58±1.17s (15); FedAvg+DP 195.21±0.59s, 375.28±4.10s, 554.04±0.95s → ~5–6× slower; similar slowdowns across batch-size and client-count settings.

Pietro Ducange et al. [57]	2024	Federated learning of explainable AI (Fed-XAI) in healthcare; case study: predicting Parkinson's disease progression (total_UPDRS).	Two Fed-XAI approaches: (i) interpretable-by-design TSK-FRBS trained in FL; (ii) MLP-NN trained in FL with post-hoc SHAP using a federated SHAP procedure (clients compute Shapley values locally; server averages them). Evaluation compares FL, Local Learning (LL), and Centralized Learning (CL) across four scenarios (IID, NIID-Q, NIID-F, NIID-FQ) with 10 hospitals (cross-silo). Federated feature selection uses Mutual Information to select G=4 features; metrics: RMSE and Pearson r. Key patterns: FL generally > LL; CL best; in NIID-F, MLP-NN has lower RMSE (10.268) while TSK-FRBS has higher r (0.461 vs 0.205).	**Best FL (NIID-F)**: MLP-NN RMSE = 10.268, r = 0.205; TSK-FRBS RMSE = 10.829, r = 0.461. **Best CL (IID)**: MLP-NN RMSE = 8.453, r = 0.485; TSK-FRBS RMSE = 9.309, r = 0.487. FL > LL in most cases but < CL
Ahmed M. et al. [34]	2025	**Healthcare Data Security & Privacy** – Blockchain-based Access Control	- Proposed **ACHealthChain** framework on **Hyperledger Fabric (PBFT + Raft)** - Uses **IPFS** for decentralized storage - Subchains: **EHRChain, DiagnosisChain, PolicyChain, LogChain** - Smart contracts (chaincode) for fine-grained access & revocation - AES + ECC for encryption	- **Improved throughput** by 19.7% and **reduced latency** by 87% vs. baselines - High scalability with 2–26 peers, throughput ~188→132 TPS, latency 0.02→0.19s - Resists **Unauthorized Access, DoS, Eavesdropping, Dolev-Yao, and Sybil attacks** - Superior security & functionality compared to MedRec, MeDShare, MediChain, Healthchain, SPChain, etc.
B. Bhasker, et al. [8]	2025	**IoMT & Healthcare 5.0 Security** – Sustainable Healthcare Systems	- **FBCI-SHS**: Federated Blockchain-IoT framework - FL for decentralized learning - Blockchain for tamper-proof storage & access control - IDS for anomaly detection - IoT medical sensors for real-time health monitoring	- Data privacy & security: **98.73%** - Intrusion detection efficiency: **97.16%** - Disease detection accuracy: **96.42%** - Proactive healthcare management: **98.37%** - Interoperability: **96.74%**
Mohammed Tawfik et al. [58]	2025	**IoMT Cybersecurity** – Federated Few-Shot with XAI	- **FedMedSecure Framework** - Ensemble of CrossTransformer, FEAT, Relation Network, Regularized MAML - Cross-Attention + Meta-Learning - Privacy: Differential Privacy (ε=1.0, $\delta=10^{-5}$) - Communication-efficient aggregation (75% reduction) - Multi-level XAI (SHAP, Attention, Prototypes)	- CICIoMT2024: **99.9% accuracy** supervised, 99.7–99.8% few-shot, 99.8% federated - CIDC2017: **93.3% supervised**, 91.0–99.3% few-shot - Global FL across 8 institutions: **99.8% accuracy** - Few-shot: Adaptation from 91% (5 shots) → 99.3% (50 shots) - Robust XAI showing packet timing & protocol features as key indicators
Kazi Fatema et al. [1]	2025	Explainable Federated Learning-based Intrusion Detection for IoT/CPS Environments	ANN distributed across 4 clients + FedAvg + SHAP explainability on the global model	Training accuracy ≈ 88.4%, test accuracy ≈ 88.2%; Precision ≈ 0.8908, Recall ≈ 0.684, F1 ≈ 0.705 on the test set; SHAP plots for feature importance.

Anas Ali, Mubashar Husain, Peter Hans [59]	2025	**Industrial IoT (IIoT) Security** – Privacy-preserving Intrusion Detection	- **FL-BCID**: Federated Learning + Blockchain - Smart contracts for update validation - Differential privacy (DP) and gradient clipping - Lightweight IDS models for IIoT nodes	- Accuracy: **97.3% (ToN-IoT)**, **96.8% (N-BaIoT)** - Precision: 95.9%, Recall: 96.2% - Reduced communication overhead by **41%** - Faster convergence (21 rounds vs 30+ for standard FL)
Deepak Kumar, Chaman Verma, Zoltán Illés [60]		Privacy-preserving **liver disease prediction** with Explainable AI	- **FL-XAI Framework** - Ensemble ML models (RF, GBC, AdaBoost, LR, DT) - Federated training across decentralized clients - **SHAP** for interpretability - **Isotonic regression** for calibration	- Accuracy: **99%**, F1-score: 98%, Brier score: 0.01, ECE: 0.59 - SHAP identified Direct Bilirubin, SGOT, and Alkaline Phosphatase as most predictive - Comparable or better than centralized models while preserving privacy
Sikander Javed et al. [61]	2025	**Intrusion Detection Systems (IDS)** – Secure, interpretable IDS with FL + Ensemble ML + XAI	- **Fed-Ensemble-XAI Framework** - Federated Learning for privacy-preserving training - Ensemble of SVM kernels (Linear, Polynomial, RBF) - SHAP for interpretability	- **98.6% accuracy**, Precision 98.3%, Recall 98.0%, F1 98.1%, AUC 0.993, FPR 1.2% - Outperformed centralized & federated baselines (Fed-Poly, Fed-RBF, Central-SVM)
Alba Amato & Dario Branco [62]	2025	**Healthcare AI** – Explainable Federated Learning with Semantic Web	- **SemFedXAI Framework** - Ontology-Enhanced Federated Learning - Semantic Aggregation Mechanism - Knowledge Graph-Based Explanation	- Accuracy improved: **FedAvg 73.5% → SemFedXAI 85.3%** - Explanation quality: **0.85 vs 0.70 (FedXAI)** - Explanation comprehensibility: **8.7/10 vs 4.5 (FedXAI)** - Improved trust, semantic consistency, and domain interpretability
Mark Devine et al. [49]	2025	**IoT Security** – Intrusion Detection with Federated ML	- **Federated SVM IDS** - Compared with ANN, Random Forest, Isolation Forest - FedAvg aggregation - Focus on **physical metrics** (delay & memory usage) for IoT feasibility	- Centralized SVM: **99.1% accuracy, 97.9% precision, 99.3% recall, 98.1% F1** - Federated SVM (5 nodes): **97.4% accuracy, 97.6% precision, 97.5% recall, 97.3% F1** - Comparable to ANN & RF; Isolation Forest weaker (~73%) - Memory per node reduced (2919 → 819 MB with 5 nodes) - Training delay competitive with ANN, faster than RF
Gökay Mutlu & Neşe Rihani [63]	2025	**IoT Security** – Intrusion Detection with FL + XAI	- **Hybrid IDS Framework** - Data preprocessing: outlier removal, missing value imputation - **SelectKBest** feature selection (top 10 features) - **SMOTE** for class imbalance - ML models: LightGBM, RF, CatBoost, XGBoost, DNN - **Federated Learning** (3 clients, ensemble aggregation) - XAI: **SHAP** (global & class-wise) + **LIME** (local explanations)	- Centralized models: LightGBM 98.88%, XGBoost 98.81%, CatBoost 98.77%, RF 97.75%, DNN 97.89% (ROC-AUC 0.9914) - Federated Learning model: **93.28% accuracy**, but only **44.12% balanced accuracy** - SHAP: key features = Number, AVG, Tot sum - LIME: clear local feature contributions (e.g., AVG ≤ 60 → Benign)

The review of the hybrid frameworks show that the integration of two technologies has a cooperative effect that recovers the limitations of single-technology frameworks. For instance, integrating Federated Learning with XAI, successfully keeping detection accuracy approximately 98% on ToN-IoT dataset and preserving data privacy comparable to centralized models [53].

Nevertheless, this critical analysis reveals that hybrid security frameworks are inherently constrained by the absence of a third foundational pillar, resulting in structural gaps that prevent them from delivering fully balanced and trustworthy system designs. Architectures that ensure data privacy and integrity by integrating FL with blockchain continue to operate as "black boxes," failing to provide essential interpretability for their automated clinical decisions. Conversely, systems combining FL with XAI successfully offer both privacy and interpretability, but they lack the immutable audit trails necessary to prevent data tampering. Ultimately, the empirical results derived from “Table 3” indicate that intermediate designs fail to simultaneously ensure traceable integrity, data confidentiality, and decision transparency, which in turn exposes unresolved security weaknesses when deployed in highly sensitive IoMT settings.

2.1.4 Triple-Integrated Systems (FL + XAI + Blockchain)

Based on our literature review the most complete category is represented by triple-integrated security frameworks, which integrate Explainable AI, Blockchain, and Federated Learning into a single framework. The end-to-end privacy, transparency, and integrity of healthcare CPS and smart environments are the goals of this framework. The only study completely incorporates all three technologies shown in “Table 4”.

Table 4: Triple-Integrated Farmwork

Author	Year	Focus Area	Method / Framework	result
Tanisha Bhardwaj & K. Sumangali [12]	2025	Privacy-preserving **healthcare data security** with FL + Blockchain + XAI	- **PPFBXAIO Framework**: Privacy-Preserving Federated Blockchain Explainable AI Optimization - SHA-256 for secure updates - **LGOA** (Levy Grasshopper Optimization) for feature selection & hyperparameter tuning - **EDBN (Entropy Deep Belief Network)** for classification - **SHAP, LIME, Grad-CAM** for explainability - Smart contracts for model validation & auditability	- Heart Disease: **93.07% accuracy**, Precision 91.19%, Recall 95.39%, F1-score 93.24% - Breast Cancer: **95.07% accuracy**, Precision 95.44%, Recall 96.54%, F1-score 95.98% - Reduced training loss by **4.93%**, latency reduced by **81 ms**, throughput ↑ **109 TPS** vs baselines (FedAvg, FL-MPC, FL-RAEC, PEFL)

The literature indicates that only one study fully integrates FL, XAI, and blockchain together. This emphasizes a significant chance to contribute, as there is no current solution that achieves distributed learning that preserves privacy, tamper-proof auditing capabilities, and explainable decision-making within a single security framework.

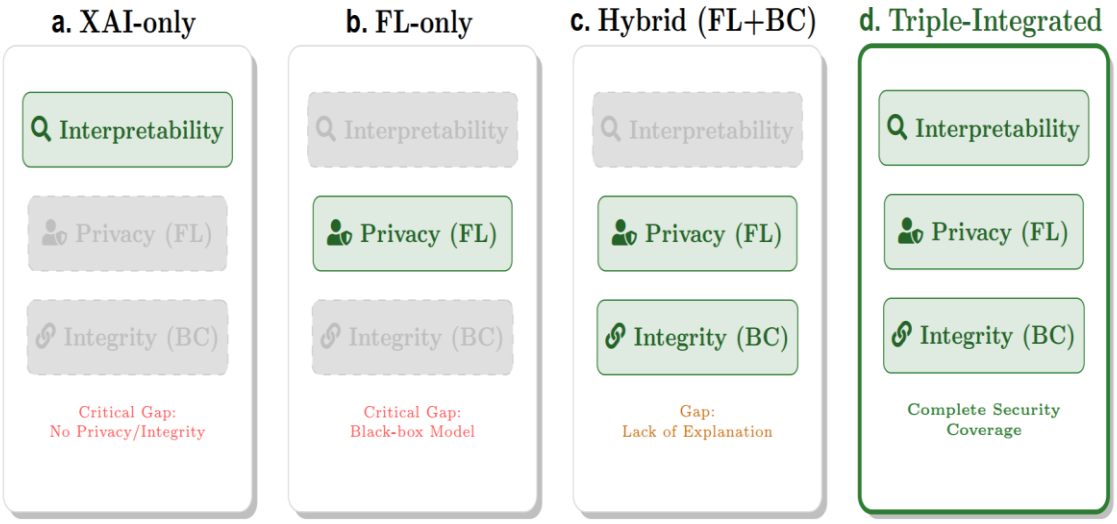

Figure 3. A component-based comparison of the four security framework categories.

Shows the missing component effect **(a)** XAI systems lack privacy and integrity mechanisms; **(b)** FL systems operate as privacy-preserving black boxes susceptible to poisoning; **(c)** Hybrid systems address two aspects but miss the third; **(d)** Triple-Integrated systems provide a comprehensive shield covering Interpretability, Privacy, and Integrity.

2.2 Datasets and Evaluation Environments

Analyzing the evaluation environments is very important before providing a cross-category performance comparison. The surveyed studies utilize a diverse array of datasets, ranging from network traffic logs for Intrusion Detection Systems (IDS) to medical imaging and physiological records for disease prediction. General-purpose network security datasets such as NSL-KDD, UNSW-NB15, and ToN-IoT, are used by a large amount of the literature as a benchmark to evaluate the security performance of healthcare and IoMT frameworks against cyberattacks. Conversely, medical datasets such as the COVID-19 Radiography Database, Heart Disease, and Parkinson's Telemonitoring are used in research on privacy-preserving diagnosis. The use of proxy (e.g., MNIST, SVHN) or synthetic datasets, which might not accurately reflect the complexity of real-

world clinical settings, is a significant drawback in some research. “Table 5” provides a brief summary of the datasets used across the surveyed studies.

Table 5: Summary of Datasets Used in Surveyed Frameworks

Dataset Category	Dataset Name	Description	Data Source (Link)
IoMT & Network Security	**CIC-IoT2023**	Large-scale dataset with 33 attack types (DDoS, Mirai, etc.) from 105 IoT devices.	[64]
	NSL-KDD	Improved version of KDD'99, a benchmark for intrusion detection.	[65]
	UNSW-NB15	Comprehensive network traffic dataset with 9 attack families.	[66]
	ToN-IoT	Telemetry data from IoT/IIoT sensors including heterogeneous data sources.	[67]
	WUSTL-EHMS-2020	Healthcare-specific dataset combining network flows with biometric patient data.	[68]
	Edge-IIoTset	Realistic cybersecurity dataset for IoT & IIoT with 14 attack types.	[69]
	N-BaIoT	Traffic data from 9 commercial IoT devices infected with Mirai/Gafgyt botnets.	[70]
Medical & Diagnosis	**COVID-19 Radiography**	21,165 chest X-ray images for COVID-19 and pneumonia detection.	[71]
	Heart Disease	Clinical records for heart disease prediction (Cleveland dataset).	[72]
	Parkinson’s	Voice recordings and telemonitoring data for disease progression.	[73]
	Liver Disease	Patient records with biochemical attributes for liver disease prediction.	[74]
Proxy & Simulated	**MNIST**	Digit images are used as a proxy for grayscale medical imaging in privacy tests.	[75]
	Smart City Index	Data related to smart city indices used for framework validation.	[76]

3. Cross-Category Comparison of Existing Security Frameworks

The reviewed literature reveals significant heterogeneity in the technical methodologies employed to secure CPS environments. Although each category such as XAI, Federated Learning, Blockchain, Hybrid Frameworks, and Triple-Integrated framework addresses specific limitations of conventional centralized security models, their goals, capacities, and constraints are very different. This section offers a cross-category comparison in order to show the advantages and disadvantages of each paradigm as well as the relationships and trade-offs between privacy, interpretability, scalability, and integrity.

a. Comparison Dimensions

The surveyed studies are evaluated against a set of categories of security solutions that are essential for healthcare Cyber-Physical Systems (CPS) environments to perform a structured comparison. These dimensions reflect both technical and operational requirements:

- **Privacy preservation:** how the system complies with medical privacy regulations and avoids sharing raw data.
- **Model interpretability and transparency:** shows if the framework has the ability to support explainable decision-making and provides understanding into model outputs.
- **Integrity and tamper-resistance:** the ability to prevent manipulation of system updates, logs, and model parameters.
- **Robustness to adversarial or poisoning attacks:** the system's ability to face malicious attempts to corrupt training data or model behavior.
- **Scalability and deployment viability:** suitability for distributed environments and compatibility with large, heterogeneous IoMT and CPS networks.
- **Real-time work:** How well the system satisfies safety-critical healthcare applications' latency and responsiveness requirements.

So, the compassion of XAI-only, FL-only, Blockchain-only, hybrid, and triple-integration approaches will be based on these dimensions.

b. Summary Comparison Across Categories

Different solution categories exhibit varying levels of effectiveness across these dimensions. While single-technology approaches excel in one or two aspects, hybrid and triple-integrated systems offer broader and more balanced coverage. To visualize the trade-offs between these approaches, "Figure 4" provides a comparative analysis across the six key dimensions, highlighting the balanced performance of the triple-integrated framework.

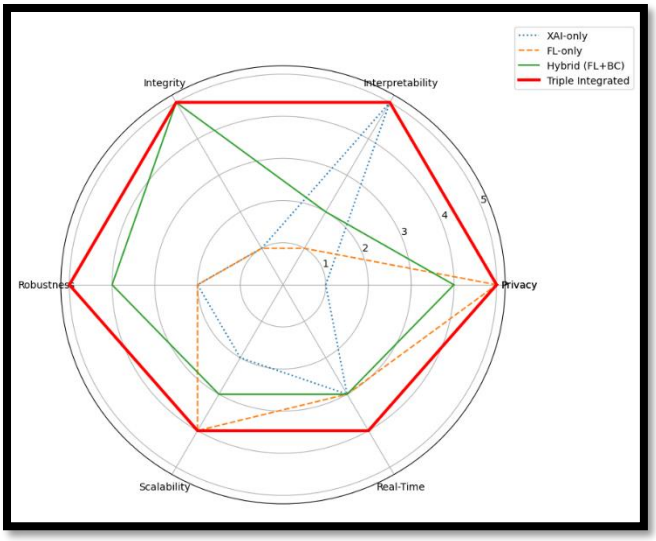

Figure 4: Cross-Category Evaluation of Privacy, Interpretability, and Integrity.

Furthermore, "Table 6" summarizes the detailed performance of each category against the comparison dimensions.

Table 6: Cross-category comparison of security frameworks

Category	Privacy Preservation	Interpretability (XAI)	Integrity / Tamper-Proofing	Robustness to Poisoning	Scalability	Real-Time Suitability
XAI-only Framework	No	High	No	Limited	Model-dependent	Moderate
FL-only Framework	High	None	No	Low without enhancements	Good	Variable
Hybrid (FL + XAI or FL + BC)	Medium–High	High or NO (varies)	High or NO (varies)	Moderate–Good	Good	Variable
Triple (FL + XAI + BC)	High	High	High	Balanced protection	Improving with optimization	Dependent on architecture

c. XAI-Based vs. Federated Learning Approaches

The main goal of explainable AI (XAI) is to make the decision made by the models more transparent. These systems assist the operators of the framework to understand the cause of the security alert and which feature led to this alert. In the healthcare sector, where clinical trust and accountability are paramount, this level of interpretability is vital. [26],[36]. However, because of the majority of current XAI-based solutions depend on centralized data collecting, implementing them in smart healthcare environments where stringent privacy protection is required presents

significant challenges. Furthermore, standalone XAI frameworks do not provide any protections for manipulation of logs or models.

Conversely, Federated Learning (FL) operate on a completely different way than the traditional methods. It eliminates the necessity of aggregating raw data in a central repository; each smart or wearable device can train its model locally. This is very helpful for protecting data privacy. However, FL models usually do not explain how their decisions are made, so they lack transparency. In addition, if there are no methods protect the shared updates, attackers can corrupt the training process by sending fake or corrupted updates to the model [17]. In summary, XAI helps to understand how decisions are made, but it does not protect data privacy, while federated learning protects data privacy, but operators still cannot easily understand how its models make decisions.

d. Federated Learning vs. Blockchain-Based Security

Even though Federated Learning successfully preserves data privacy, it cannot completely protect the model updates themselves. A client that has damaged or the attacked server can send bad updates, reuse old parameters, or interfere with the training process. This leads to many security risks particularly within IoMT systems that there devices are spread out [77],[78].

Blockchain looks at the problem from a different angle. It has no effect with learning, but it makes sure that no change on data can be done without being noticed. Through its shared ledger and agreement rules. It can secure all the data in the framework including records, model updates, and access actions [59],[79]. However, blockchain cannot detect attacks, learn patterns, or even adjust to new threats.

Consequently, when we compare between of these two technologies it can be clear that, FL keeps data private but does not stop tampering, instead Blockchain protects the integrity of data and models but does not have intelligence or explanations.

Given these complementary strengths, combining FL with Blockchain emerges as a robust architectural solution that concurrently provides privacy and integrity [8]. Nevertheless, without XAI, this hybrid system continues to operate as a "black box," and the framework's stakeholders cannot understand or trust its decisions.

e. XAI vs. Blockchain: Two Different Types of Trust

By understanding XAI and Blockchain it can be notable that they offer two different types of trust. Explainable AI (XAI) helps the system operators to understand why the model makes this decision.

While blockchain helps them to trust the integrity (not changed by anyone) of the data exchanged into the framework [25],[80].

A comprehensive review of the literature indicates that most of researches look at these two points separately.

XAI papers mostly do not have secure access control, and blockchain papers often assume that if the data is safe, then the models take a correct decision, which is not always true. This led us to think about healthcare frameworks including both types of trust at the same time, since good systems need safe data and understandable decisions.

f. Hybrid Two-Technology Systems

As shown in "Table 3" some researchers try to combine two technologies together in order to fix the limits of a single framework. The most common mixes are FL with XAI or FL with blockchain. For example:

1. The combination of FL with XAI perceives the system's data privacy and gives clearer explanations for its decision [1].
2. When FL is combined with blockchain the system's data remain private and the training process becomes harder to tamper with [55].

These combinations led the researchers to think in the right direction, but they still do not have privacy, explanations, and strong integrity in one complete solution. Several studies note that even these hybrid approaches are not enough for highly sensitive CPS systems, which can be vulnerable to virus types of attacks.

g. Triple-Integrated Systems (FL + XAI + Blockchain)

Based on literature one recent work tries to bring all of the three technologies together in one framework, which is PPFBXAIO by Bhardwaj & Sumangali [12]. In this system, FL protects the data, XAI explains the model's decisions, and blockchain keeps the records and updates safe from tampering. Few other works suggest similar ideas, such frameworks are still uncommon because they may be hard to build and can make for a heavy performance [81]. For the previous section it can be notable that these triple-layer systems suggest the most complete privacy, explicability and strong integrity. This can make them the fittest for trustworthy digital healthcare environments.

4. Open Research Challenges and Gaps

Even though many recent studies have improved the security of healthcare Cyber-Physical Systems (CPS) using Explainable AI (XAI), Federated Learning (FL), and Blockchain, there are still important unsolved problems. Most recent and current research only focuses on one or two of these technologies, this leaves significant gaps in terms of practical implementation, speed, and real-world testing.

a. High Computational Cost and Energy Usage

Modern security frameworks increasingly rely on computationally intensive techniques, which directly conflict with the limited battery capacity and processing power of IoMT devices.

- Blockchain-based security introduces significant overhead when conventional consensus mechanisms are used. In particular, Proof-of-Work schemes consume excessive energy and incur considerable delays, making them impractical for lightweight medical sensors and wearable healthcare devices.
- Many security solutions rely on complex deep learning models, including BiLSTM, ensemble methods, and attention-based architectures. These models typically demand computing power beyond what standard edge nodes can provide.
- The combined execution of federated learning coordination, blockchain validation processes, and explainability mechanisms further increases system latency. In time-sensitive medical applications, this accumulated delay can surpass acceptable response limits, posing risks to real-time clinical decision-making and patient safety.

b. Lack of Real-World Testing

More of the previously suggested frameworks have not been tested in real digital healthcare environments. Rather, the researchers try to use offline datasets or simulated environments, which do not mirror the noise and the complexity of real-world healthcare systems.

- **Synthetic & Proxy Data:** Several studies use non-medical information as if it were healthcare data, as in some studies. Using the SVHN dataset (digit images) or MNIST as a stand-in for medical imaging is one of the common examples. However, this does not demonstrate that the system is effective for actual patient diagnoses, such as X-rays.
- **Usability Gap:** A lot of XAI explanation models are evaluated for the decisions technically but might be very difficult for non-technical medical staff that work on the framework to understand.

c. Scalability and Latency Issues

Frameworks and systems that work in smart cities must be able to handle thousands of connected devices. However, there is no guarantee that previously existing frameworks will work in the environments because they have primarily been tested on a very small-scale client. The blockchain verification and the time needed for the aggregation and communications of FL updates will be cause an increase in time delay (latency). Also, adding encryption methods like homomorphic encryption to protect the privacy of data slows down the system, this will be creating a trade-off between makes the system fast in result and keeping the data safe.

d. Data Problems and Generalization

In general, the data in smart cities are noisy and unbalanced (non-IID) especially the real-world medical data. This makes some classes or attack types uncommon and many of the current models fail to detect these types accurately. Moreover, a number of previous and current studies evaluated their work on a single dataset, which increases the possibility that the model will not work well if applied in smart cities that include many hospitals or devices. Due to their heavy dependance on high-quality "normal" data, which are not always available, semi-supervised approaches also encounter challenges.

e. Shortage of Triple-Integrated Solutions

There is a clear lack of frameworks that combine all three technologies (FL + XAI + Blockchain) into one solution. While hybrid models (combining two) are common, they often miss one key element: they might offer privacy and integrity but lack explanation, or offer explanation without tamper-proof security. Only one reviewed study attempted to integrate all three, but it faced significant challenges in complexity and was only tested on two datasets. “Table 7” summarizes these limitations and links them to the relevant studies.

Table 7: Summary of Open Research Challenges

Challenge Category	Description	Affected Studies
High Resource Usage	Complex models and blockchain consensus drain battery and processing power of small medical devices.	[8] [44] [49] [58]
Unrealistic Testing	Reliance on simulations, offline data, or non-medical datasets (like MNIST/digits) instead of real clinical trials.	[2] [4] [42] [54]
Scalability & Speed	Systems become slow or fail when adding more devices due to encryption and communication overhead.	[47] [55] [82] [83]
Data Issues	Models fail to generalize to new hospitals or struggle with rare attacks/diseases due to unbalanced data.	[37] [48] [53] [63]
Lack of Integration	Very few studies combine Privacy, Integrity, and Explainability in a single architecture.	[12] and Most Hybrid Studies

5. Future Research Directions

Considering the limitations mentioned in the previous section, the way to develop a secure and intelligent healthcare framework is to balance the trade-off between security, privacy, and performance. In future, researchers should focus on the following practical areas to convert these theoretical concepts into real-world solutions:

1. **Designing and Developing "Lightweight" Frameworks:** The main problem for the IoMT devices is the heavy computational load. Therefore, future works must change from using standard blockchain technologies and deep learning models toward "lightweight" alternatives.

   - **For Blockchain:** Instead of using heavy power techniques like Proof-of-Work, researchers should test to use energy-efficient consensus mechanisms like Proof-of-Stake (PoS) or Directed Acyclic Graph (DAG).
   - **For AI:** To make the deep learning models more exactable and runnable on the wearable medical devices without draining their batteries. The model comparison techniques, such as "pruning" and "quantization," could be used.

2. **Privacy-Preserving Explainability:** The substance here is "*How can we explain a decision without revealing the sensitive data behind it?"* Future work needs to think; of "Privacy-Preserving XAI." This includes creating algorithms that can explain the decisions (using tools like SHAP) while mathematically guaranteeing (via differential privacy) that these explanations cannot be used as a source for exposing the system data.
3. **Scalable Architectures for Smart Cities** there are thousands or millions of devices to be handled in a smart city; the framework cannot store everything on the main blockchain. They should split the network into smaller pieces and using Off-Chain Storage like the InterPlanetary File System (IPFS). In this model, only the small cryptographic hashes are saved on the blockchain, while heavy data and model updates are stored off-chain to ensure speed and low latency.
4. **Building Standardized Triple-Integrated APIs** Since there is a clear shortage of systems that combine FL, XAI, and Blockchain, the research community needs to build open-source, modular A standardized framework will accelerate the adoption of these triple-integrated systems in actual hospitals.
5. **Quantum Computing for Acceleration and Security:** the healthcare networks are very large in scale; traditional computing for mathematics required finding it complex for securing

Blockchain and aggregating Federated Learning updates globally. Future research should think of using **Quantum Machine Learning (QML)** that could significantly speed up complex model training. Furthermore, with the development of quantum computing comes the threat of breaking current encryption; therefore, developing **"Quantum-Resistant"** cryptographic algorithms for the blockchain layer is a critical future requirement to ensure long-term patient data safety.

6. Conclusion

This survey has examined recent advances in securing smart healthcare environments built on CPS and the IoMT, with particular emphasis on the roles of XAI, FL, and Blockchain technologies. By systematically reviewing and classifying studies published between 2023 and 2025, the paper highlights the ongoing shift from traditional centralized security architectures toward decentralized and intelligence-driven solutions that better address the privacy, transparency, and integrity requirements of Healthcare 5.0 systems. The analysis shows that each of the three technologies contributes to healthcare cybersecurity in a distinct way. XAI improves transparency and trust by making security decisions interpretable, FL enhances privacy by enabling collaborative learning without exposing sensitive data, and Blockchain strengthens integrity through tamper-resistant storage and verifiable audit trails. However, when deployed in isolation, none of these technologies is sufficient to fully protect complex and distributed healthcare infrastructures. Hybrid approaches represent a step forward but often involve trade-offs that leave at least one critical security requirement unaddressed. A key finding of this survey is the limited number of frameworks that achieve full integration of XAI, FL, and Blockchain within a single architecture. This gap indicates a significant opportunity for future research to develop unified security frameworks that balance privacy preservation, decision transparency, and data integrity without imposing excessive computational or communication overhead. Addressing challenges related to scalability, real-time operation, and resource constraints in IoMT devices remains essential for practical deployment. Overall, this survey provides a structured foundation for researchers and practitioners seeking to design trustworthy and scalable cybersecurity solutions for smart healthcare systems. By identifying current limitations and promising research directions, it contributes toward the development of next-generation Healthcare 5.0 infrastructures that are secure, transparent, and resilient.

References

[1] K. Fatema *et al.*, "Federated XAI IDS: An Explainable and Safeguarding Privacy Approach to Detect Intrusion Combining Federated Learning and SHAP †," *Futur. Internet*, vol. 17, no. 6, Jun. 2025, doi: 10.3390/fi17060234.

[2] R. Abid, M. Rizwan, A. Alabdulatif, A. Alnajim, M. Alamro, and M. Azrour, "Adaptation of Federated Explainable Artificial Intelligence for Efficient and Secure E-Healthcare Systems," *Comput. Mater. Contin.*, vol. 78, no. 3, pp. 3413–3429, 2024, doi: 10.32604/cmc.2024.046880.

[3] S. R. Abbas, Z. Abbas, A. Zahir, and S. W. Lee, "Federated Learning in Smart Healthcare: A Comprehensive Review on Privacy, Security, and Predictive Analytics with IoT Integration," Dec. 01, 2024, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/healthcare12242587.

[4] M. Taufik, M. S. Aziz, and A. Fitriana, "Hybrid Explainable AI (XAI) Framework for Detecting Adversarial Attacks in Cyber-Physical Systems," *J. Technol. Informatics Eng.*, vol. 4, no. 1, Apr. 2025, doi: 10.51903/jtie.v4i1.295.

[5] M. Georgiades and F. Hussain, "An Explainable AI Approach for Interpretable Cross-Layer Intrusion Detection in Internet of Medical Things," *Electron.*, vol. 14, no. 16, Aug. 2025, doi: 10.3390/electronics14163218.

[6] N. Kaur and L. Gupta, "Explainable AI Assisted IoMT Security in Future 6G Networks," *Futur. Internet*, vol. 17, no. 5, May 2025, doi: 10.3390/fi17050226.

[7] M. Algarni and S. Mishra, "A Secure and Reliable Framework for Explainable Artificial Intelligence (XAI) in Smart City Applications," *Eng. Technol. Appl. Sci. Res.*, vol. 14, no. 4, pp. 15291–15296, Aug. 2024, doi: 10.48084/etasr.7676.

[8] B. Bhasker *et al.*, "Blockchain framework with IoT device using federated learning for sustainable healthcare systems," *Sci. Rep.*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-06539-z.

[9] Y. Shahsavari, O. A. Dambri, Y. Baseri, A. S. Hafid, and D. Makrakis, "Integration of Federated Learning and Blockchain in Healthcare: A Tutorial," pp. 1–39, 2024, [Online]. Available: http://arxiv.org/abs/2404.10092

[10] S. A. Sharaf and S. Nooh, "Identifying significant features in adversarial attack detection framework using federated learning empowered medical IoT network security," *Sci. Rep.*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-14913-0.

[11] C. S. Kalutharage, X. Liu, C. Chrysoulas, N. Pitropakis, and P. Papadopoulos, "Explainable AI-Based DDOS Attack Identification Method for IoT Networks," *Computers*, vol. 12, no. 2, Feb. 2023, doi: 10.3390/computers12020032.

[12] T. Bhardwaj and K. Sumangali, "An explainable federated blockchain framework with privacy-preserving AI optimization for securing healthcare data," *Sci. Rep.*, vol. 15, no. 1, Dec. 2025, doi:

10.1038/s41598-025-04083-4.

[13] E. Shammar *et al.*, “Threat to trust: A systematic review on Internet of medical things security,” *J. Parallel Distrib. Comput.*, vol. 206, no. 9, p. 105172, 2025, doi: https://doi.org/10.1016/j.jpdc.2025.105172.

[14] M. M. Nair, A. K. Tyagi, and R. Goyal, “Medical Cyber Physical Systems and Its Issues,” *Procedia Comput. Sci.*, vol. 165, no. 2019, pp. 647–655, 2019, doi: 10.1016/j.procs.2020.01.059.

[15] N. Dey, A. S. Ashour, and S. J. Fong, “Medical cyber-physical systems : A survey,” pp. 1–13, 2018.

[16] A. Younesi and T. Fahringer, “HealthCare 5.0: An industry 5.0 perspective for next-generation medical systems with synergistic integration of IoT, AI, and 6G,” *Internet of Things*, vol. 35, no. September 2025, p. 101815, 2025, doi: 10.1016/j.iot.2025.101815.

[17] S. R. Hassan, A. Hassan, A. Maqsood, S. Hijab, A. Younesi, and T. Fahringer, “A survey on intelligent secure and distributed frameworks for Healthcare 5.0,” *Discov. Artif. Intell.*, vol. 5, no. 1, p. 101815, 2025, doi: 10.1007/s44163-025-00572-7.

[18] G. D. Gallo and D. Micucci, “Internet of Medical Things Systems Review: Insights into Non-Functional Factors,” *Sensors*, vol. 25, no. 9, 2025, doi: 10.3390/s25092795.

[19] P. K. Sadhu, V. P. Yanambaka, A. Abdelgawad, and K. Yelamarthi, “Prospect of Internet of Medical Things: A Review on Security Requirements and Solutions,” *Sensors*, vol. 22, no. 15, pp. 1–31, 2022, doi: 10.3390/s22155517.

[20] S. R. Putta, “Security and Privacy of Wearable Internet of Medical Things: Stakeholders Perspective,” *Culminating Proj. Inf. Assur.*, vol. 69, 2018.

[21] R. V Patil, N. P. Ambritta, P. N. Mahalle, and N. Dey, “Medical Cyber–Physical Systems in Society 5.0: Are We Ready?,” *IEEE Trans. Technol. Soc.*, vol. 3, no. 3, pp. 189–198, 2022, doi: 10.1109/TTS.2022.3185396.

[22] A. Alzahrani, M. Alshehri, R. AlGhamdi, and S. K. Sharma, “Improved Wireless Medical Cyber-Physical System (IWMCPS) Based on Machine Learning,” *Healthc.*, vol. 11, no. 3, 2023, doi: 10.3390/healthcare11030384.

[23] A. Gupta and A. Singh, “A Comprehensive Survey on Cyber-Physical Systems Towards Healthcare 4.0,” *SN Comput. Sci.*, vol. 4, no. 2, p. 199, 2023, doi: 10.1007/s42979-023-01669-5.

[24] N. Capuano, G. Fenza, V. Loia, and C. Stanzione, “Explainable Artificial Intelligence in CyberSecurity: A Survey,” *IEEE Access*, vol. 10, pp. 93575–93600, 2022, doi: 10.1109/ACCESS.2022.3204171.

[25] G. Rjoub *et al.*, “A Survey on Explainable Artificial Intelligence for Cybersecurity,” *IEEE Trans. Netw. Serv. Manag.*, vol. 20, no. 4, pp. 5115–5140, 2023, doi: 10.1109/TNSM.2023.3282740.

[26] B. Sharma, L. Sharma, C. Lal, and S. Roy, “Explainable artificial intelligence for intrusion detection in IoT networks: A deep learning based approach,” *Expert Syst. Appl.*, vol. 238, Mar. 2024, doi: 10.1016/j.eswa.2023.121751.

[27] H. Guan, P.-T. Yap, A. Bozoki, and M. Liu, “Federated learning for medical image analysis: A survey,” *Pattern Recognit.*, vol. 151, p. 110424, 2024, doi: https://doi.org/10.1016/j.patcog.2024.110424.

[28] G. Choi, W. C. Cha, S. U. Lee, and S. Y. Shin, “Survey of Medical Applications of Federated Learning,” *Healthc. Inform. Res.*, vol. 30, no. 1, pp. 3–15, 2024, doi: 10.4258/hir.2024.30.1.3.

[29] F. R. da Silva, R. Camacho, and J. M. R. S. Tavares, “Federated Learning in Medical Image Analysis: A Systematic Survey,” *Electron.*, vol. 13, no. 1, pp. 1–22, 2024, doi: 10.3390/electronics13010047.

[30] M. Aggarwal *et al.*, “Federated Learning on Internet of Things: Extensive and Systematic Review,” *Comput. Mater. Contin.*, vol. 79, no. 2, pp. 1795–1834, 2024, doi: 10.32604/cmc.2024.049846.

[31] E. Dritsas and M. Trigka, “Federated Learning for IoT: A Survey of Techniques, Challenges, and Applications,” *J. Sens. Actuator Networks*, vol. 14, no. 1, 2025, doi: 10.3390/jsan14010009.

[32] M. S. Amin, S. Ahmad, and W.-K. Loh, “Federated learning for Healthcare 5.0: a comprehensive survey, taxonomy, challenges, and solutions,” *Soft Comput.*, vol. 29, no. 2, pp. 673–700, 2025, doi: 10.1007/s00500-025-10508-z.

[33] A. J, D. P. Isravel, K. M. Sagayam, B. Bhushan, Y. Sei, and J. Eunice, “Blockchain for healthcare systems: Architecture, security challenges, trends and future directions,” *J. Netw. Comput. Appl.*, vol. 215, no. March, p. 103633, 2023, doi: 10.1016/j.jnca.2023.103633.

[34] A. M. Tawfik, A. Al-Ahwal, A. S. T. Eldien, and H. H. Zayed, *Blockchain-based access control and privacy preservation in healthcare: a comprehensive survey*, vol. 28, no. 8. Springer US, 2025. doi: 10.1007/s10586-025-05308-x.

[35] Z. J. Al-Araji *et al.*, “Healthcare Security in Edge-Fog-Cloud Environment using Blockchain: A Systematic Review,” *Mesopotamian J. CyberSecurity*, vol. 5, no. 2, pp. 606–635, 2025, doi: 10.58496/MJCS/2025/037.

[36] S. A. Memon, U. K. Wiil, and M. Shaikh, “Explainable Intrusion Detection for Internet of Medical Things,” in *International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, IC3K - Proceedings*, Science and Technology Publications, Lda, 2023, pp. 40–51. doi: 10.5220/0012210300003598.

[37] M. Keshk, N. Koroniotis, N. Pham, N. Moustafa, B. Turnbull, and A. Y. Zomaya, “An explainable deep learning-enabled intrusion detection framework in IoT networks,” *Inf. Sci. (Ny).*, vol. 639, Aug. 2023, doi: 10.1016/j.ins.2023.119000.

[38] M. Siganos *et al.*, “Explainable AI-based Intrusion Detection in the Internet of Things,” in *ACM*

International Conference Proceeding Series, Association for Computing Machinery, Aug. 2023. doi: 10.1145/3600160.3605162.

[39] A. Lateef Haroon P.S. and H. K N, "Feasible Implementation of Explainable AI Empowered Secured Edge Based Health Care Systems," *J. Smart Internet Things*, vol. 2024, no. 2, pp. 1–12, Dec. 2024, doi: 10.2478/jsiot-2024-0008.

[40] I. Elgarhy, M. M. Badr, M. Mahmoud, M. Alsabaan, T. Alshawi, and M. Alsaqhan, "XAI-Based Accurate Anomaly Detector That Is Robust Against Black-Box Evasion Attacks for the Smart Grid," *Appl. Sci.*, vol. 14, no. 21, Nov. 2024, doi: 10.3390/app14219897.

[41] O. Arreche, T. Guntur, and M. Abdallah, "XAI-IDS: Toward Proposing an Explainable Artificial Intelligence Framework for Enhancing Network Intrusion Detection Systems," *Appl. Sci.*, vol. 14, no. 10, May 2024, doi: 10.3390/app14104170.

[42] Y. Hosain and M. Çakmak, "XAI-XGBoost: an innovative explainable intrusion detection approach for securing internet of medical things systems," *Sci. Rep.*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-07790-0.

[43] A. Alabbadi and F. Bajaber, "X-FuseRLSTM: A Cross-Domain Explainable Intrusion Detection Framework in IoT Using the Attention-Guided Dual-Path Feature Fusion and Residual LSTM," *Sensors*, vol. 25, no. 12, Jun. 2025, doi: 10.3390/s25123693.

[44] Y. Kayode Saheed and J. Ebere Chukwuere, "CPS-IIoT-P2Attention: Explainable Privacy-Preserving With Scaled Dot-Product Attention in Cyber-Physical System-Industrial IoT Network," *IEEE Access*, vol. 13, pp. 81118–81142, 2025, doi: 10.1109/ACCESS.2025.3566980.

[45] M. Butt *et al.*, "A Fog-Based Privacy-Preserving Federated Learning System for Smart Healthcare Applications," *Electron.*, vol. 12, no. 19, Oct. 2023, doi: 10.3390/electronics12194074.

[46] J. A. Shaikh *et al.*, "RCLNet: an effective anomaly-based intrusion detection for securing the IoMT system," *Front. Digit. Heal.*, vol. 6, 2024, doi: 10.3389/fdgth.2024.1467241.

[47] J. M. Wang, K. Yang, and M. J. Li, "NIDS-FGPA: A federated learning network intrusion detection algorithm based on secure aggregation of gradient similarity models," *PLoS One*, vol. 19, no. 10, Oct. 2024, doi: 10.1371/journal.pone.0308639.

[48] S. Sun, P. Sharma, K. Nwodo, A. Stavrou, and H. Wang, "FedMADE: Robust Federated Learning for Intrusion Detection in IoT Networks Using a Dynamic Aggregation Method," Aug. 2024, [Online]. Available: http://arxiv.org/abs/2408.07152

[49] M. Devine, S. P. Ardakani, M. Al-Khafajiy, and Y. James, "Federated Machine Learning to Enable Intrusion Detection Systems in IoT Networks," *Electron.*, vol. 14, no. 6, Mar. 2025, doi: 10.3390/electronics14061176.

[50] V. T. Nguyen and R. Beuran, "FedMSE: Semi-supervised federated learning approach for IoT

network intrusion detection," Apr. 2025, doi: 10.1016/j.cose.2025.104337.

[51] M. Ragab *et al.*, "Advanced artificial intelligence with federated learning framework for privacy-preserving cyberthreat detection in IoT-assisted sustainable smart cities," *Sci. Rep.*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-88843-2.

[52] M. R. War, Y. Singh, and Z. A. Sheikh, "FedSec-CPS: Federated Learning Based Security Framework for Security of Cyber-Physical Systems," in *Procedia Computer Science*, Elsevier B.V., 2025, pp. 1837–1844. doi: 10.1016/j.procs.2025.04.139.

[53] A. Si-ahmed, M. A. Al-Garadi, and N. Boustia, "Explainable Machine Learning-Based Security and Privacy Protection Framework for Internet of Medical Things Systems," Sep. 2025, [Online]. Available: http://arxiv.org/abs/2403.09752

[54] D. Commey, S. Hounsinou, and G. V. Crosby, "Securing Health Data on the Blockchain: A Differential Privacy and Federated Learning Framework," May 2024, [Online]. Available: http://arxiv.org/abs/2405.11580

[55] R. Bensaid, N. Labraoui, A. A. A. Ari, H. Saidi, J. H. M. Emati, and L. Maglaras, "SA-FLIDS: secure and authenticated federated learning-based intelligent network intrusion detection system for smart healthcare," *PeerJ Comput. Sci.*, vol. 10, pp. 1–34, 2024, doi: 10.7717/peerj-cs.2414.

[56] J. Almalki, S. M. Alshahrani, and N. A. Khan, "A comprehensive secure system enabling healthcare 5.0 using federated learning, intrusion detection and blockchain," *PeerJ Comput. Sci.*, vol. 10, 2024, doi: 10.7717/peerj-cs.1778.

[57] P. Ducange, F. Marcelloni, A. Renda, and F. Ruffini, "Federated Learning of XAI Models in Healthcare: A Case Study on Parkinson's Disease," *Cognit. Comput.*, vol. 16, no. 6, pp. 3051–3076, Nov. 2024, doi: 10.1007/s12559-024-10332-x.

[58] M. Tawfik, A. A. Abu-Ein, H. M. Noaman, A. H. Abdelhaliem, and I. S. Fathi, "FedMedSecure: Federated Few-Shot Learning with Cross-Attention Mechanisms and Explainable AI for Collaborative Healthcare Cybersecurity," Aug. 06, 2025. doi: 10.21203/rs.3.rs-7208692/v1.

[59] A. Ali, M. Husain, and P. Hans, "Federated Learning-Enhanced Blockchain Framework for Privacy-Preserving Intrusion Detection in Industrial IoT," May 2025, [Online]. Available: http://arxiv.org/abs/2505.15376

[60] D. Kumar, C. Verma, and Z. Illés, "Federated learning with explainable AI for liver disease prediction: A privacy-preserving approach," *Intell. Med.*, vol. 12, Jan. 2025, doi: 10.1016/j.ibmed.2025.100285.

[61] S. Javed *et al.*, "Secure and Interpretable Intrusion Detection through Federated and Ensemble Machine Learning with XAI", doi: 10.56979/901/2025.

[62] A. Amato and D. Branco, "SemFedXAI: A Semantic Framework for Explainable Federated

Learning in Healthcare," *Inf.*, vol. 16, no. 6, Jun. 2025, doi: 10.3390/info16060435.

[63] G. Mutlu, N. Rihani, and N. N. Rihani, "Intrusion Detection System with Explainable AI and Federated Learning." [Online]. Available: https://www.researchgate.net/publication/392262172

[64] Canadian Institute for Cybersecurity (CIC), "CICIoT2023 Dataset," 2023. [Online]. Available: https://www.kaggle.com/datasets/akashdogra/cic-iot-2023

[65] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "NSL-KDD Dataset," 2009. [Online]. Available: https://www.unb.ca/cic/datasets/nsl.html

[66] N. Moustafa and J. Slay, "UNSW-NB15 Dataset," 2015. [Online]. Available: https://researchdata.edu.au/the-unsw-nb15-dataset/1957529

[67] A. Alsaedi, N. Moustafa, Z. Tari, A. Mahmood, and A. Anwar, "ToN-IoT Datasets," 2020. [Online]. Available: https://research.unsw.edu.au/projects/toniot-datasets

[68] Washington University in St. Louis, "WUSTL-EHMS-2020 Dataset," 2020. [Online]. Available: https://www.cse.wustl.edu/~jain/ehms/index.html

[69] M. Lemay and J. M. Fernandez, "Edge-IIoTset Cyber Security Dataset," 2022. [Online]. Available: https://ieee-dataport.org/documents/edge-iiotset-new-comprehensive-realistic-cyber-security-dataset-iot-and-iiot-applications

[70] Y. Meidan *et al.*, "N-BaIoT Dataset," 2018. [Online]. Available: https://archive.ics.uci.edu/dataset/442/detection+of+iot+botnet+attacks+n+baiot

[71] T. Rahman, M. E. H. Chowdhury, A. Khandakar, K. R. Islam, and others, "COVID-19 Radiography Database," 2020. [Online]. Available: https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database

[72] UCI Machine Learning Repository, "Heart Disease Dataset," 1988. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/heart+Disease

[73] A. Tsanas, M. A. Little, P. E. McSharry, and L. O. Ramig, "Parkinson's Telemonitoring Dataset," 2009. [Online]. Available: https://archive.ics.uci.edu/dataset/189/parkinsons+telemonitoring

[74] UCI Machine Learning Repository, "ILPD (Indian Liver Patient Dataset)," 2012. [Online]. Available: https://www.kaggle.com/datasets/uciml/indian-liver-patient-records

[75] Y. LeCun, C. Cortes, and C. J. C. Burges, "The MNIST Database of Handwritten Digits," 1998. [Online]. Available: http://yann.lecun.com/exdb/mnist/

[76] IMD World Competitiveness Center, "Smart City Index Datasets," 2024. [Online]. Available: https://www.kaggle.com/datasets/magdamonteiro/smart-cities-index-datasets

[77] E. Bagdasaryan, A. Veit, Y. Hua, D. Estrin, and V. Shmatikov, "How To Backdoor Federated Learning," *Proc. Mach. Learn. Res.*, vol. 108, pp. 2938–2948, 2020.

[78] A. N. Bhagoji, S. Chakraborty, P. Mittal, and S. Calo, “Analyzing federated learning through an adversarial lens,” *36th Int. Conf. Mach. Learn. ICML 2019*, vol. 2019-June, pp. 1012–1021, 2019.

[79] F. Casino, T. K. Dasaklis, and C. Patsakis, “A systematic literature review of blockchain-based applications: Current status, classification and open issues,” *Telemat. Informatics*, vol. 36, no. May 2018, pp. 55–81, 2019, doi: 10.1016/j.tele.2018.11.006.

[80] A. Tandon, A. Dhir, N. Islam, and M. Mäntymäki, “Blockchain in healthcare: A systematic literature review, synthesizing framework and future research agenda,” *Comput. Ind.*, vol. 122, 2020, doi: 10.1016/j.compind.2020.103290.

[81] M. Rahmati and A. Pagano, “Federated Learning-Driven Cybersecurity Framework for IoT Networks with Privacy Preserving and Real-Time Threat Detection Capabilities,” *Informatics*, vol. 12, no. 3, Sep. 2025, doi: 10.3390/informatics12030062.

[82] M. Bai *et al.*, “Adversarial Attack against Intrusion Detectors in Cyber-Physical Systems With Minimal Perturbations,” in *Proceedings - 2024 IEEE International Symposium on Parallel and Distributed Processing with Applications, ISPA 2024*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 816–825. doi: 10.1109/ISPA63168.2024.00109.

[83] A. M. Tawfik, A. Al-Ahwal, A. S. T. Eldien, and H. H. Zayed, “ACHealthChain blockchain framework for access control and privacy preservation in healthcare,” *Sci. Rep.*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-00757-1.