International Journal of Emerging Engineering and Technology (IJEET)
ISSN (e): 2958-3764
Volume: 5, Number: 1, Pages: 32 - 39, Year: 2026

Federated Learning and Artificial Intelligence in IoT: Security Challenges, Solutions, and Future Directions

Bilal Gillani, Irshad Ahmed Sumra, and Jahanzaib Sattar

School of Science & Technology, University of Management & Technology, Lahore, 54000, Pakistan
*Corresponding author: Bilal Gillani (Email: syed.gillani.bilal@gmail.com)

Received: 07/02/2026, Revised: 15/05/2026, Accepted: 19/06/2026

***Abstract*— There has been a great proliferation of the Internet of Things in different areas, ranging from healthcare, smart cities, industry, transport to agriculture, among others, and as a result of which an immense volume of distributed data is generated. The problem with traditional centralized machine learning methods is in the need to transfer sensitive data to centralized servers, which leads to additional risks of privacy breaches, communication overhead, and failure of single point. As a way of solving mentioned challenges, there appears the federated learning framework, which allows making collaborative model training without data transfer thus increasing the privacy of the system and saving on communications.**

***Index Terms*— Internet of Things (IoT), Federated Learning (FL), Explainable Artificial Intelligence (XAI), IoT Security, Privacy Preservation, Cyber Attacks, Machine Learning, Intrusion Detection Systems (IDS).**

I. Introduction

The term Internet of Things is used to describe the system of physical objects embedded with sensors and software, which gather and share data in an almost automated fashion [1], [2]. Based on the figures projected in the literature surveyed, the number of connected devices increased from about 9.7 billion in 2020 to reach 29 billion by 2030, with the market value of these devices surpassing USD 33 billion by 2027 [3]. This massive adoption of IoT technologies has rendered them essential for smart cities, smart healthcare, smart industry, smart agriculture, and smart transportation, although at the same time it has amplified two longstanding controversies in distributed computing, namely the privacy-utility and the security-resources controversies in model training at the edge [4].Traditionally, machine learning used by the IoT system relied on centralization, which involved sending the raw data gathered by sensors to a cloud server for analysis [5]. Such practice was characterized by additional communications overhead, had single points of failure, and also exposed users to significant threats to their privacy, including notorious cases such as Cambridge Analytica scandal in 2018 when centrally collected data from the user was being misused [3]. Federated Learning helped solve this problem, enabling devices to perform training locally and share updates of trained models without sending raw data [6], [7]. Nevertheless, FL adds an element of opacity to the process as aggregated global models become hard to comprehend by stakeholders, and in this situation, Explainable AI provides tools such as LIME, SHAP, and Grad-CAM to make these decisions more comprehensible [8], [9]. Meanwhile, heterogeneity and computational limitations that make IoT devices desirable also made them vulnerable to attacks, leading to an independent body of research on the topic of IoT security covering all five layers of the IoT architecture [4], [10], [11].

This paper brings these two threads of investigation together. Section II reviews methodology and scope of each surveyed paper, together with the conceptual framework of Federated Learning, which includes mathematics and taxonomy of the technology. Section III reviews Explainable AI in general and its place within the IoT domain, together with the survey on IoT security, specifically architectural risks, taxonomy of attacks and a Radio-Frequency Identification (RFID) case study. Section IV integrates both frameworks into one picture and demonstrates common ground for FL, XAI and security techniques, together with a comparison of the state-of-the-art enabling technologies (machine learning, deep learning, blockchain, Software-Defined Networking (SDN), edge computing) between these two papers. Section V outlines open issues and future directions and Section VI provides the conclusion.

While this paper draws on the material covered in [3] and [4], its contribution lies in synthesis rather than re-statement. First, it constructs a unified conceptual framework, presented in Section 4, that maps the FL/XAI perspective of Dubey and Kumar [3] directly onto the layered security perspective of Sharma and Bhushan [4], a connection that neither original survey makes explicit. Second, it introduces five comparative tables (Tables 1, 4 and 5 in particular) that place the two surveys' treatment of privacy, attack taxonomies, and enabling technologies such as blockchain, SDN, and edge/fog computing

side by side, allowing differences in scope and emphasis to be assessed directly. Third, it identifies recurring vulnerabilities, such as inference attacks on model updates and adversarial manipulation of intrusion detection systems, that are discussed independently in each source paper but are shown here to be the same underlying problem viewed from two angles. Finally, Section 5 consolidates the future-work recommendations of both surveys into a single five-dimensional research agenda, highlighting overlaps, such as the shared need for resource-efficient algorithms and privacy-preserving explainability, that are not visible when the two papers are read separately.

II. SCOPE AND METHODOLOGY OF THE TWO SURVEYED PAPERS

This survey [3] studies Explainable AI and Federated Learning integration into the Next Generation of IoT. The contributions made by the authors of this survey are: a taxonomy of FL based on data distribution, network topology, and data sharing based on time; a study of different client selection approaches; an overview of the most advanced aggregation methods like FedSGD [7] up to new approaches FedProx, FedMA [12] and FedGKt [13]; the mathematical modeling of operational, data heterogeneity, privacy and resource optimization problems in federated learning for IoT; and convergence of FL and XAI approaches (FED-XAI) in healthcare, manufacturing, smart cities, agriculture, retail and energy areas.

This survey [4] discusses the issue of IoT security from several angles: architecture level issues in perception, network, application, processing, and business layers; component level weaknesses in terms of hardware (sensors and communication channels), and software (operating system and protocols); application level attacks in healthcare, smart home, traffic management, agriculture and industrial IoT applications; attacks taxonomy classified by layers, where RFID is used as a case study; and state of art defensive technologies including machine learning [14], deep learning, blockchain [15], software defined networking, federated learning [16] and edge/fog computing (Table I).

TABLE I: Summarizes the comparative scope of the two papers across major thematic dimensions.

Dimension	Dubey & Kumar [3]	Sharma & Bhushan [4]
Primary focus	Convergence of FL and XAI in IoT	IoT security threats and defenses
IoT architecture treatment	Application-oriented (smart city, healthcare, industry, vehicle, environment, agriculture)	Layer-oriented (perception, network, application, processing, business)
Mathematical formalization	Extensive (FL objective functions, aggregation rules, privacy noise models, resource optimization)	Moderate (reliability/availability equations, energy optimization)
Attack taxonomy	Limited (security treated as a challenge category within FL)	Extensive (layer-wise attacks plus RFID-specific case study)
Case study depth	Sector case studies (diabetic care, manufacturing, smart waste)	Device-level case study (RFID across perception, network, application layers)
Defensive technology review	FL aggregation algorithms, differential privacy, secure aggregation	ML, DL, blockchain, SDN, federated learning, edge/fog computing
Treatment of FL as a security tool	Indirect (privacy by design)	Direct (FL-based intrusion detection, Section 10.5)
Treatment of XAI as a security tool	Central theme (trust, transparency, bias detection)	Not directly addressed

A. Conceptual Background of Federated Learning

In Federated Learning, the models are trained in devices that are distributed, and at the same time, the raw data remains in these local devices [6]. The individual devices train a local model based on their datasets, and then the model parameters are sent to the central server, where the global model is generated by aggregating the model updates from different devices. The mathematical formulation of the whole process has been done by Dubey and Kumar [3], where the objective function for the global model J(w) is defined as the weighted sum of local objective functions,

$$J(w) = \Sigma i \left(w_i \cdot l_i(w)\right) \quad (1)$$

Here w_i is a weight associated with the i-th local loss $l_i(w)$. The objective function of each local device is minimized through the process of stochastic gradient descent, whereas the global model parameters are aggregated using methods like FedAvg [7],

$$w = w + (1/k) \times \Sigma\, (k\ client\ gradient\ updates) \quad (2)$$

B. Classification of Federated Learning

According to the survey [3], there are three ways for FL classification. Firstly, by data distribution: Horizontal Federated Learning (HFL) is used where devices have the same feature spaces but different sample spaces, for example, mobile keyboard prediction; Vertical Federated Learning (VFL) is used where devices have the same sample spaces but different

feature spaces, for example, e-commerce and banks providing services to the same city's citizens; Federated Transfer Learning (FTL) is an extended version of VFL and involves the participation of the devices with both different samples and different features, for example, cross-hospital disease diagnosis. Secondly, by network topology: Centralized Federated Learning (CFL) is based on a central coordinating server, whereas Decentralized Federated Learning (DFL) involves peer-to-peer communication, which is usually embedded with blockchains [17]. Thirdly, by data sharing timing: Synchronous FL synchronizes updates at fixed intervals, whereas Asynchronous FL can allow devices to submit their updates whenever they want [18].

C. Client Selection and Aggregation Algorithms

Since not all the clients are well-suited to be a part of each training cycle, the survey highlights the importance of choosing appropriate clients with consideration of heterogeneity, resource distribution, communication cost, and fairness. Subsequently, the history of development of aggregation algorithms is discussed starting from FedSGD and FedAvg to more specific versions created for solving problems related to heterogeneity, expensive communication, and privacy protection. Table II presents the main aggregation algorithms reviewed in the survey.

TABLE II. Summary of representative Federated Learning aggregation algorithms [3], [7]–[13].

Algorithm	Core Mechanism	Primary Limitation Addressed
FedSGD / FedAvg [7]	Gradient or weighted-parameter averaging	Baseline; struggles with statistical and system heterogeneity
FedMeta [19]	Model-agnostic meta-learning across participant tasks	System heterogeneity; fast adaptation to new tasks
FedPer [20]	Shares base layers, keeps personalized layers local	Statistical heterogeneity; privacy of personalization
FedMD [21]	Transfer learning with knowledge distillation on a public dataset	Allows heterogeneous local model architectures
FedProx [22]	Adds a proximal regularization term to local objectives	System and data heterogeneity; communication overhead
FedAsync [18]	Asynchronous updates with a staleness penalty	Non-IID data; scalability
FedMA [12]	Layer-wise matched averaging	Communication overhead for deep CNN/LSTM models
FedGKt [13]	Asynchronous knowledge transfer between edge and server	Communication overhead; supports large CNNs on edge
MHAT [3]	Knowledge distillation on model outputs rather than parameters	Model heterogeneity; reduces communication rounds by ~82%
FedGA [3]	Combines FedPer with genetic algorithms for weight search	Joint data and model heterogeneity

III. EXPLAINABLE AI AND ITS ROLE IN IOT

XAI deals with the black-box problem of deep learning models which are extensively applied for performing IoT applications like sensor data processing, predictive maintenance, and anomaly detection [8], [9]. The absence of interpretable models makes it difficult to identify the cause of errors, measure bias, and defend machine-driven decisions, which is highly important in critical industries like healthcare and self-driving cars [23]. The survey [3] identifies XAI techniques as intrinsic interpretability, where models like decision trees are intrinsically interpretable, and post-hoc interpretability, where a surrogate model is created in order to interpret a pre-trained black-box model. Post-hoc techniques can be further subdivided into model-specific and model-agnostic techniques. Examples of model-specific techniques include DeepLift and Grad-CAM, whereas examples of model-agnostic techniques include LIME and Partial Dependence Plot (PDP).

The combination of Federated Learning and XAI technologies makes it possible to use the latter not only at the client level to provide an explanation for the contribution of local data to the learning process but also at the server level to understand why specific predictions are produced by the aggregated global model [24]. Such integration is called FED-XAI in the review and is explained by various sector examples such as managing diabetes patients by means of continuous glucose monitoring systems, global quality control of manufacturing processes, smart city traffic and waste management, precision agriculture, retail customization, and energy grid optimization [3], [25], [26].

A. IoT Security: Architecture, Challenges, and Attack Taxonomy

The analysis of security by Sharma & Bhushan [4] is based on an architecture of layered IoT system consisting of perception layer (consisting of sensors, RFID tags, and actuators), network layer (routing and transmission mechanisms including Bluetooth, Zigbee, RPL, and 6LoWPAN), processing layer (consisting of computations, aggregations, and storages), application layer (MQTT, CoAP, XMPP, AMQP and other protocols) and the business layer (decision support and visualization). The security needs of each layer and the attack profiles are different and have been presented in Table III [27], [11].

TABLE III. IoT architectural layers and associated security challenges [4].

Layer	Representative Technologies	Major Security Challenges
Perception	RFID, sensors, Bluetooth, LoRaWAN	Physical tampering, limited compute for encryption, device heterogeneity
Network	6LoWPAN, RPL, CARP, Wi-Fi	Secure transmission, device authentication, scalability, key management
Processing	Cloud computing, DBMS, big data	Secure storage and backup, data integrity and consistency
Application	MQTT, CoAP, XMPP, AMQP	Data privacy, access control, malware, interoperability
Business	HTTP, FTP, dashboards	Regulatory compliance, data ownership, trust management

In addition to the design of such devices, the survey highlights four critical issues that span all IoT solutions: data storage and processing at scale, energy usage (whereby the attacker can purposely deplete the battery), real-time implementation requirements (where delay is itself the attack point, like in traffic management), and the need for both security and privacy together with limited resources.

B. Layer-wise Attack Taxonomy

The survey categorizes the attacks according to the layer being attacked as follows: [4], [28].

At the perception layer: attacks consist of Distributed Denial of Service (DDoS) on sensors, physical attacks on nodes (destruction or capturing), malicious code injection, side-channel attack on power and timing information, physical tampering, spoofing of trusted nodes, and eavesdropping.

At the network layer: attacks consist of traffic analysis, sinkhole, blackhole, wormhole, and selective forwarding which are routing attacks, man-in-the-middle interception, and flooding.

At the application layer: attacks consist of malicious code injection, reflection DDoS, exploitation DDoS, reprogramming, cryptanalysis attacks on weak encryption techniques, and side-channel attack because of resource limitation of IoT devices [29].

C. RFID as an Illustrative Case Study

To be able to apply the abstract taxonomy into technologies in practice, the survey [4] uses RFID as its example. Regarding the perception layer, the RFID tag can be damaged, substituted, jammed with sending wrong messages, use of kill commands, and replayed with sending the signals once more. Regarding the network layer, RFID can be cloned through duplication of the information on the tag through using the Physically Unclonable Function, eavesdropped, spoofed, and man-in-the-middle attack in wireless transmission. Lastly, regarding the application layer, RFID software can be manipulated, read illegally through the lack of memory of storing data, and attacked through side channels (Fig. 3).

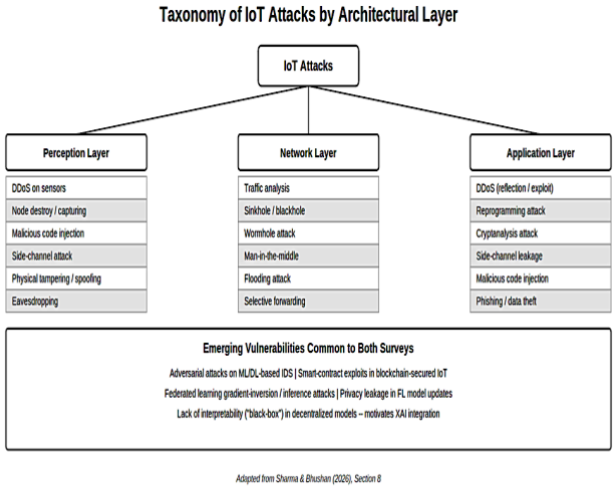

Fig. 1. Taxonomy of IoT attacks organized by perception, network, and application layers, with emerging vulnerabilities shared across both surveyed papers [3], [4].

D. Emerging Vulnerabilities

These two survey articles analyze the same subset of vulnerabilities, which arise because of IoT security technologies. Adversarial attacks on machine learning and deep learning intrusion detection systems, where the attacker manipulates the input data in order not to be detected [14]; vulnerabilities of smart contracts that secure IoT devices on the basis of blockchain technology; the visibility of smart contract code in open-source platforms, such as Ethereum, which could be used in malicious attacks; and federated learning vulnerabilities, where leakage of sensitive data is possible due to the inference attack and decreased accuracy of detection for minority class events due to the locally biased models [30], [31], [16] are considered in Sharma and Bhushan's survey [4]. These vulnerabilities are exactly those ones, which are addressed by the proposed privacy-preserving and explainability techniques in Dubey and Kumar's survey [3], [32], [33].

IV. INTEGRATING THE TWO PERSPECTIVES: A UNIFIED FRAMEWORK

Considering both surveys collectively [3], [4], a comprehensive model emerges where Federated Learning and Explainable AI are not just issues related to AI quality but important parts of the IoT security architecture, and vice versa - the IoT security technologies (such as intrusion detection systems, secure aggregation, and differential privacy) are essential prerequisites to implement FL and XAI safely. The proposed model is illustrated in Fig. 2: the IoT architecture proposed by the security survey serves as input to the Federated Learning layer, and its output becomes understandable with the help of XAI, whereas the whole system operates under permanent threat and protection from the attacks classified in Section III.

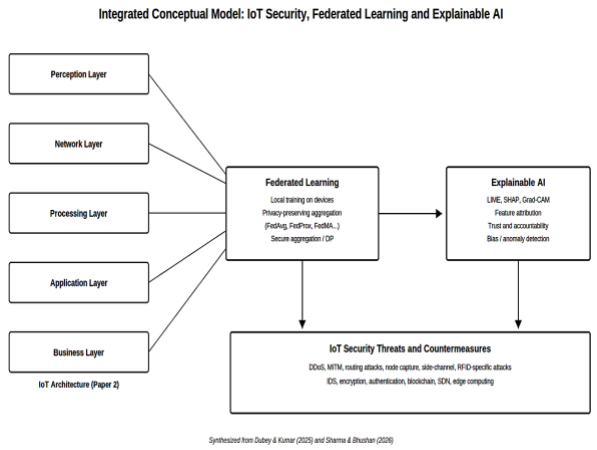

Fig. 2. Integrated conceptual model linking the IoT architectural stack, Federated Learning, Explainable AI, and the security threat-and-countermeasure landscape [3], [4].

Specific concrete connections also exist. Firstly, FedAvg and its variants [7], which are surveyed in Section 2 (C) [3], are actually employed in Sharma and Bhushan's Section [4] as a way of securing a network: federated intrusion detection systems are trained together on edge devices in such a way that no traffic logs are centralized, i.e., they actually make use of the FL methodology overviewed by Dubey and Kumar [3]. Secondly, the privacy-preserving mechanisms of differential privacy and secure aggregation introduced in Dubey and Kumar [3] and referred to as such in [32], [33], are precisely the approaches suggested by Sharma and Bhushan as means of counteracting data leakage and inference attacks. Thirdly, the use of XAI to uncover bias and anomalies in the behavior of models has much in common with the anomaly-based intrusion detection systems introduced in the security paper [29], hinting at the possibility of using explainability in the security domain.

TABLE IV. Cross-paper linkages between Federated Learning/XAI concepts and IoT security concepts [3], [4].

Concept in Dubey & Kumar [3]	**Corresponding Concept in Sharma & Bhushan [4]**	**Integrated Implication**
Secure aggregation, differential privacy [32], [33]	Countermeasures against data leakage and inference attacks	Privacy mechanisms double as security countermeasures
FedAvg / FedProx aggregation [7], [22]	FL-based intrusion detection (Section 10.5)	Aggregation algorithms can be repurposed for distributed threat detection
XAI bias and anomaly detection	Anomaly-based IDS [29]	Explainability techniques could augment detection of novel attacks
Data heterogeneity / non-IID challenges	Device heterogeneity challenges	Shared root cause: resource-constrained, diverse IoT devices
Communication efficiency (compression, quantization)	Bandwidth management, model compression countermeasures	Same engineering techniques serve both efficiency and security goals
Privacy-preserving explanation methods (future direction)	Adversarial attacks on ML-based IDS [14]	Explaining a security model risks revealing exploitable information

A. State-of-the-Art Enabling Technologies: A Comparative View

However, both studies address the same set of facilitating technologies, albeit from different perspectives, because Dubey and Kumar consider facilitating technologies mainly in terms of tools used to enhance the efficiency and reliability of FL, whereas Sharma and Bhushan view them as security solutions.

B. Machine Learning and Deep Learning

The work by Sharma and Bhushan [4] includes machine learning techniques like Naive Bayes, Decision Trees, Support Vector Machine and k-Nearest Neighbor that are used for applications like intrusion detection, traffic classification, and detecting malicious nodes [14]; on the other hand, deep learning techniques like CNNs, LSTMs, autoencoders, and Generative Adversarial Networks capture more complicated patterns but come with high complexity. The deep learning aspect has been implicitly incorporated in Dubey and Kumar's work in the FL aggregation algorithms in section 2.3 [3] where the majority use CNN or LSTM networks [12],[13].

C. Blockchain

While both surveys recognize the usefulness of blockchain in decentralization, there is a difference in the approach they take towards blockchain technology. The security survey considers the benefits of blockchain in terms of authentication and decentralized intrusion detection through which single points of failure are avoided [15], [17], [31], while the survey on federated learning and explainable artificial intelligence recognizes the use of blockchain in Decentralized Federated Learning, in which there is an exchange of updates between models using blockchain ledgers rather than using a central server [3]. Both surveys highlight the overheads involved in implementing blockchain technology.

D. Software-Defined Networking and Edge/Fog Computing

Sharma and Bhushan define SDN as a solution for the centralization of policy control while leaving IoT devices to do just the data plane tasks, facilitating the implementation of computationally complex security mechanisms. In the case of edge and fog computing, key generation, model retraining, and intrusion detection are delegated to more computationally powerful neighboring devices in order to reduce latency. Thus, the Multi-access Edge Computing (MEC) approach mentioned in the FL/XAI survey when considering FED-XAI implementations is consistent with the ideas described by

Sharma and Bhushan, because both surveys use proximity of edge devices to the source of data as a way to address the trade-off between computational limitations and complex analytics.

E. Federated Learning as a Security Technology

One of the most evident parallels between the two works concerns federated learning itself. Whereas Dubey & Kumar [3] consider federated learning as their main topic of study, Sharma & Bhushan [4] see it as just one among several modern security techniques (Section 10.5), which include such use cases as federated malware detection [16] and federated anomaly detection for IoT traffic [30]. In both articles, FL is considered as an important technique because of its ability to provide decentralization without sharing of the raw data, but both authors recognize the limitations of such a capability as well – leaks of information in the form of model updates, the non-iid problem, and communication costs. Table V represents the comparison of these technologies.

TABLE V. Comparative treatment of enabling technologies across the two surveyed papers [3], [4].

Technology	**Role in Dubey & Kumar [3]**	**Role in Sharma & Bhushan [4]**
Machine Learning	Embedded within FL aggregation methods	Direct: classification, anomaly detection, IDS [14]
Deep Learning	CNN/LSTM backbones for local models [12], [13]	CNN, LSTM, autoencoders, GANs for IDS
Blockchain	Substrate for decentralized FL	Decentralized authentication and IDS [17], [31]
SDN	Not directly addressed	Centralized policy and traffic monitoring
Edge/Fog Computing	MEC as deployment platform for FED-XAI	Offloads IDS and key management tasks
Federated Learning	Central subject of the survey	One of several security-enhancing technologies [16]

V. IoT CHALLENGES AND RESEARCH DIRECTIONS

Combining future work from both papers provides an aggregated research agenda around five dimensions.

- Privacy-preserving explainability: Dubey and Kumar [3] point out that XAI methods might unintentionally reveal the sensitive information contained in federated updates in explaining the models' decisions. The risk described by Sharma and Bhushan about performing inference attacks on model updates shared through federation [30], [4] is the same. Future works should design privacy-preserving explanation methods [32].
- Resource-efficient algorithms: Both papers note that the battery, memory, and computing capacity of IoT devices are limited. Dubey and Kumar demand lightweight XAI methods and communication-efficient aggregation [3]; Sharma and Bhushan ask for lightweight cryptography and intrusion detection which takes into account resource constraints [4]. The common limitation suggests that a shared research area will be compressing both AI and security algorithms for the same resource capacity.
- Standardization and benchmarking: As pointed out by Dubey and Kumar [3], there is a lack of standardized datasets and performance metrics for FED-XAI, especially because XAI models such as decision trees need pre-extracted, interpretable features as input, rather than raw imagery/text data. Similarly, Sharma and Bhushan [4] observe a lack of standardization in the communication protocol among vendors and advocate for common frameworks. These gaps reflect the need for common benchmarks considering both the AI quality and security robustness aspects.
- Confronting adversarial and poisoning attacks: The vulnerabilities of machine learning and deep learning algorithms to adversarial attacks presented by Sharma and Bhushan [14], [4], along with the recognition of secure aggregation costs by Dubey and Kumar [33], motivate the development of robust aggregation functions that would be able to filter out poisoned/adversarial inputs while retaining the advantages of FL from the communication viewpoint.
- Decentralized security management at scale: Sharma and Bhushan's description of issues involved in decentralized security management (heterogeneity in devices, lack of centralized control) [4] resembles Dubey and Kumar's description of dynamic participation of clients in FL [3]. Decentralized governance framework, possibly blockchain-powered [17], [31], which takes care of model aggregation and security policy management is described as an area with potential but little research.

The integrated framework proposed in Section 4, along with the cross-paper relationships presented in Table 4, indicates that advances in one area can directly strengthen the others. For example, improved secure-aggregation protocols enhance both privacy protection and resilience against poisoning attacks. Similarly, stronger explainability techniques not only increase user trust but also support more effective anomaly detection. Resource-aware algorithms further contribute by improving AI efficiency while making security solutions more practical for deployment on constrained IoT devices[36,37].

VI. CONCLUSION

This survey brings together insights from two complementary surveys and argues that the future of trustworthy IoT systems depends on viewing Federated Learning (FL), Explainable AI (XAI), and IoT security as interconnected components rather than separate research areas. Dubey and Kumar show that FL helps preserve privacy by ensuring that raw data remains on local devices, while XAI improves transparency by making decentralized AI decisions easier to understand. Together, these technologies support the development of trustworthy AI solutions across domains such as smart cities, healthcare, industry, agriculture, retail, and energy. At the same time, Sharma and Bhushan highlight that the same decentralized and resource-constrained IoT environment faces a wide range of security threats across all architectural layers. These threats

range from physical attacks on perception-layer sensors to application-layer attacks such as DDoS and cryptanalysis. Addressing these challenges requires a comprehensive set of security mechanisms, including authentication, encryption, intrusion detection systems, blockchain, Software-Defined Networking (SDN), and even federated learning itself. Overall, future research should focus on these overlapping areas rather than treating FL, XAI, and security as independent fields of study. Such an integrated approach is likely to produce more resilient, secure, and trustworthy next-generation IoT systems.

FUNDING STATEMENT

The author(s) received no specific funding for this study.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest to report regarding the present study.

AUTHOR CONTRIBUTIONS

Conceptualization, A.M.S., S.A.K., and M.R.H.; methodology, A.M.S., S.A.K., and M.R.H; software, A.M.S., S.A.K., and Z.A.L.; validation, A.M.S., S.A.K., and M.R.H.; writing—original draft preparation, A.M.S., S.A.K., and Z.A.L.; writing—review and editing, A.M.S., S.A.K., and Z.A.L.

INSTITUTIONAL REVIEW BOARD STATEMENT

Not applicable.

INFORMED CONSENT STATEMENT

Not applicable.

DATA AVAILABILITY STATEMENT

Data is available on reasonable request.

REFERENCES

[1] P. Sethi and S. Sarangi, “Internet of things: Architectures, protocols, and applications,” J. Electr. Comput. Eng., vol. 2017, pp. 1–25, 2017.

[2] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, “Internet of things: A survey on enabling technologies, protocols, and applications,” IEEE Commun. Surv. Tutor., vol. 17, no. 4, pp. 2347–2376, 2015.

[3] P. Dubey and M. Kumar, “Integrating explainable AI with federated learning for next-generation IoT: A comprehensive review and prospective insights,” Comput. Sci. Rev., vol. 56, p. 100697, 2025.

[4] A. Sharma and K. Bhushan, “A comprehensive survey on IoT security: Challenges, security issues, and countermeasures,” Comput. Sci. Rev., vol. 59, p. 100839, 2026.

[5] M. S. Mahdavinejad, M. Rezvan, M. Barekatain, P. Adibi, P. Barnaghi, and A. P. Sheth, “Machine learning for internet of things data analysis: A survey,” Digit. Commun. Netw., vol. 4, no. 3, pp. 161–175, 2018.

[6] D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, J. Li, and H. V. Poor, “Federated learning for internet of things: A comprehensive survey,” IEEE Commun. Surv. Tutor., vol. 23, no. 3, pp. 1622–1658, 2021.

[7] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, “Communication-efficient learning of deep networks from decentralized data,” in Proc. Artif. Intell. Stat. (AISTATS), 2017, pp. 1273–1282.

[8] I. Kok, F. Y. Okay, O. Muyanli, and S. Ozdemir, “Explainable artificial intelligence (XAI) for internet of things: A survey,” IEEE Internet Things J., 2023.

[9] W. Samek, T. Wiegand, and K.-R. Muller, “Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models,” arXiv:1708.08296, 2017.

[10] R. Roman, P. Najera, and J. Lopez, “Securing the internet of things,” Computer, vol. 44, no. 9, pp. 51–58, 2011.

[11] J. Granjal, E. Monteiro, and J. Sa Silva, “Security for the internet of things: A survey of existing protocols and open research issues,” IEEE Commun. Surv. Tutor., vol. 17, no. 3, pp. 1294–1312, 2015.

[12] H. Wang, M. Yurochkin, Y. Sun, D. Papailiopoulos, and Y. Khazaeni, “Federated learning with matched averaging,” arXiv:2002.06440, 2020.

[13] C. He, M. Annavaram, and S. Avestimehr, “Group knowledge transfer: Federated learning of large CNNs at the edge,” in Proc. Adv. Neural Inf. Process. Syst. (NeurIPS), vol. 33, 2020, pp. 14068–14080.

[14] F. Hussain, R. Hussain, S. A. Hassan, and E. Hossain, “Machine learning in IoT security: Current solutions and future challenges,” IEEE Commun. Surv. Tutor., vol. 22, no. 3, pp. 1686–1721, 2020.

[15] B. K. Mohanta, D. Jena, U. Satapathy, and S. Patnaik, “Survey on IoT security: Challenges and solution using machine learning, artificial intelligence and blockchain technology,” Internet of Things, vol. 11, p. 100227, 2020.

[16] V. Rey, P. M. S. Sanchez, A. H. Celdran, and G. Bovet, “Federated learning for malware detection in IoT devices,” Comput. Netw., vol. 204, p. 108693, 2022.

[17] A. Dorri, S. S. Kanhere, R. Jurdak, and P. Gauravaram, “Blockchain for IoT security and privacy: The case study of a smart home,” in Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops (PerCom Workshops), 2017, pp. 618–623.

[18] C. Xie, S. Koyejo, and I. Gupta, “Asynchronous federated optimization,” arXiv:1903.03934, 2019.

[19] F. Chen, M. Luo, Z. Dong, Z. Li, and X. He, “Federated meta-learning with fast convergence and efficient communication,” arXiv:1802.07876, 2018.

[20] M. G. Arivazhagan, V. Aggarwal, A. K. Singh, and S. Choudhary, “Federated learning with personalization layers,” arXiv:1912.00818, 2019.

[21] D. Li and J. Wang, “FedMD: Heterogenous federated learning via model distillation,” arXiv:1910.03581, 2019.

[22] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, “Federated learning: Challenges, methods, and future directions,” IEEE Signal Process. Mag., vol. 37, no. 3, pp. 50–60, 2020.

[23] H. Elayan, M. Aloqaily, F. Karray, and M. Guizani, “Internet of behavior (IoB) and explainable AI systems for influencing IoT behavior,” IEEE Netw., 2022.

[24] R. Lopez-Blanco, R. S. Alonso, A. Gonzalez-Arrieta, P. Chamoso, and J. Prieto, “Federated learning of explainable artificial intelligence (FED-XAI): A review,” in Proc. Int. Symp. Distrib. Comput. Artif. Intell. (DCAI), 2023, pp. 318–326.

[25] A. Bucur, F. Manni, A. Bukharev, S. Moorthy, N. I. Mendez, and A. Jain, “Federated learning and explainable AI in healthcare,” in Explainable AI in Healthcare, Boca Raton, FL, USA: Chapman and Hall/CRC, 2024, pp. 279–294.

[26] A. Rahman et al., “Federated learning-based AI approaches in smart healthcare: Concepts, taxonomies, challenges and open issues,” Cluster Comput., vol. 26, no. 4, pp. 2271–2311, 2023.

[27] A. Mosenia and N. K. Jha, “A comprehensive study of security of internet-of-things,” IEEE Trans. Emerg. Topics Comput., vol. 5, no. 4, pp. 586–602, 2016.

[28] N. Neshenko, E. Bou-Harb, J. Crichigno, G. Kaddoum, and N. Ghani, “Demystifying IoT security: An exhaustive survey on IoT vulnerabilities and a first empirical look on internet-scale IoT exploitations,” IEEE Commun. Surv. Tutor., vol. 21, no. 3, pp. 2702–2733, 2019.

[29] A. Khraisat and A. Alazab, “A critical review of intrusion detection systems in the internet of things: Techniques, deployment strategy, validation strategy, attacks, public datasets and challenges,” Cybersecurity, vol. 4, no. 1, pp. 1–27, 2021.

[30] H. N. C. Neto, J. Hribar, I. Dusparic, D. M. F. Mattos, and N. C. Fernandes, “A survey on securing federated learning: Analysis of applications, attacks, challenges, and trends,” IEEE Access, vol. 11, pp. 41928–41953, 2023.

[31] B. K. Mohanta, D. Jena, S. Ramasubbareddy, M. Daneshmand, and A. H. Gandomi, “Addressing security and privacy issues of IoT using blockchain technology,” IEEE Internet Things J., vol. 8, no. 2, pp. 881–888, 2020.

[32] R. C. Geyer, T. Klein, and M. Nabi, “Differentially private federated learning: A client level perspective,” arXiv:1712.07557, 2017.

[33] K. Bonawitz et al., “Practical secure aggregation for privacy-preserving machine learning,” in Proc. ACM SIGSAC Conf. Comput. Commun. Secur. (CCS), 2017, pp. 1175–1191.

[34] S. Nizetic, P. Solic, D. Lopez-de-Ipina Gonzalez-de-Artaza, and L. Patrono, “Internet of things (IoT): Opportunities, issues and challenges

towards a smart and sustainable future," J. Clean. Prod., vol. 274, p. 122877, 2020.

[35] M. Humayun, N. Tariq, M. Alfayad, M. Zakwan, G. Alwakid, and M. Assiri, "Securing the internet of things in artificial intelligence era: A comprehensive survey," IEEE Access, 2024.

[36] I. A. Sumra, I. Ahmad, H. Hasbullah and J. -l. bin Ab Manan, "Behavior of attacker and some new possible attacks in Vehicular Ad hoc Network (VANET)," *2011 3rd International Congress on Ultra-Modern Telecommunications and Control Systems and Workshops (ICUMT)*, Budapest, Hungary, 2011, pp. 1-8.

[37] I. A. Sumra, H. Hasbullah and J. -l. A. Manan, "VANET security research and development ecosystem," *2011 National Postgraduate Conference*, Perak, Malaysia, 2011, pp. 1-4, doi: 10.1109/NatPC.2011.6136344.