Enhancing Privacy in IoT Systems with Federated Learning: A Multi-Agent Solution for Smart City Networks

1. **Santosh H Lavate, Assistant Professor, AISSMS COE Pune, Savitribai Phule Pune University, Maharashtra, India**
2. **Dr Ravindra K Moje, Assistant Professor, PDEA's COE Pune, Savitribai Phule Pune University, Maharashtra, India**

Abstract: The Internet of Things (IoT) is growing very quickly. This has led to the introduction of smart gadgets that are linked to each other in cities, which has made town's smart cities. These gadgets produce huge amounts of data, which can be used to improve city services and processes. But this information is often very private, which is a big safety issue. The unique ways of gathering and analysing statistics positioned non-public data in a single location, which makes it more likely to be stolen. To cope with these issues, this take a look at indicates a new way to enhance privateness in IoT systems by using Federated learning (FL), a decentralised machine learning technique. In the case of smart metropolis networks, the cautioned solution consists of a multi-agent form that we could many gadgets take care of data regionally, ensuring that raw records never leaves the source. Each IoT device inside the community works on its personal, the usage of its very own information to construct nearby models and solely sending model adjustments to a central computer. Those changes are put together by the server to make a worldwide version better. This protects privacy while keeping the data useful for city-wide analytics. Federated learning helps protect privacy and lowers the risk of data leaks by spreading out the learning process and limiting the moving of raw data. This article also talks about the structure of the multi-agent solution, focussing on the security features, communication protocols, and model synchronisation to make sure the system works well in real-life IoT settings.

Keywords: Internet of Things (IoT), Federated Learning, Privacy Preservation, Multi-Agent Systems, Smart City Networks

I. Introduction

The Internet of Things (IoT) has changed how we interact with our surroundings by linking many different systems and devices, like smart homes and factory machines, together to make huge networks that produce huge amounts of data. These past few years, building "smart cities" has become very important. The goal is to make city management better, public services better, and people's quality of life better by using IoT technologies. IoT devices in smart cities constantly gather information about things like traffic trends, the environment, energy use, public safety, and healthcare tracking. These sources of data have a huge amount of promise to make city operations better and make new services possible. But along with the perks, worries about privacy have become a big problem. Traditional IoT systems gather, handle, and store private data centrally. This puts the data at risk of leaks and unauthorised access because it is easy to get to. For instance, when health, location, and habit-related personal information are stored in one place in databases or the cloud, it can be used for illegal purposes or attacked by hackers. Additionally, the usual method of sending raw data from IoT devices to central servers or cloud platforms frequently goes against the data minimisation principle because it requires sending a lot of private data over the network. This makes the technology less safe and makes people less likely to trust it. So, methods that protect privacy are very important for the smooth implementation and use of IoT in smart towns. Federated Learning (FL) has become a potential way to deal with privacy issues in IoT and other distributed systems. FL is different from standard centralised machine learning methods because it lets machine learning models be trained directly on edge devices, like IoT devices. Because of this these devices can examine from their own data except sending personal facts to a central laptop.

The devices don't share the facts itself; rather, they solely percentage version adjustments, like slopes, after they've been educated domestically. Those changes are put together with the aid of the central pc to enhance the global version. This keeps the privateness of every information supply secure. The danger of statistics breaches is lower with this decentralised approach because personal data stays on the local gadgets. in relation to clever towns, incorporating Federated gaining knowledge of has many benefits, consisting of better privateness, decrease statistics transfer costs, and less difficult expansion. Also, FL lets information processing and analytics take place closer to the source [1]. This makes IoT networks extra efficient by using lowering delay and reducing the amount of labour that centralised computer systems should do. As clever cities get smarter, an increasing number of IoT devices may be added. Federated studying offers a device that may grow with these devices while nevertheless defensive privateness. However there are some issues that want to be constant before Federated gaining knowledge of can be used in clever town IoT systems. One of the biggest troubles is ensuring that IoT gadgets in a diffusion community can communicate to every other and stay in sync. Each device in the network has its own local model, and the central computer only gets changes that are compiled from all the devices.

II. Literature Review

A. Overview of IoT privacy challenges

The fast growth of the Internet of Things (IoT) has completely changed how cities work. Now, smart cities use the power of gadgets that are connected to each other to make city life better. However, the huge amounts of data that IoT devices produce often include private and sensitive data about people, like health measurements, location data, and trends of behaviour, which can be very bad for privacy. Wearable tech, outdoor monitors, smart meters, and linked cars are just some of the things that IoT systems use to gather data. When this information is sent or kept centrally, it can be stolen, misused, or accessed by people who aren't supposed to. Statistics leaking is one of the principal safety troubles with IoT. IoT gadgets are often related to cloud-primarily based structures. This process can make personal statistics public without the customers' permission, that's a major breach of privacy. Placing collectively non-public facts from many devices is any other trouble. While this record is joined with different units of records, it could be used to identify particular people [2]. Besides sturdy protections, IoT structures could permit tracking and monitoring besides meaning to. Further, IoT privacy issues go beyond just collecting and storing data. There is the question of who owns and controls the records. There is a lack of openness because customers may not completely apprehend how their data is being used or shared via IoT carrier groups. IoT networks are not centralised, which makes it tougher to apply privacy policies which can be the same across gadgets and platforms. As the range and complexity of IoT devices grow, so do the troubles that include protective data privateness. To lessen risks and defend human being's privateness at the same time as keeping IoT systems practical and beneficial, new thoughts are wished.

B. Current privacy-preserving techniques in IoT

Several privacy-preserving methods have been suggested and put into use to deal with the growing privacy issues in IoT systems. These methods are meant to keep private information safe while still letting IoT systems work well. Data encryption is a popular method that protects data by encoding it before sending it so that people who aren't supposed to can't read it [3]. Encrypting data can make it safer, but it takes computing power and can cause more delay, especially in IoT devices that don't have a lot of resources. Figure 1 shows a number of different IoT privacy-protecting methods for sending data securely and protecting users' privacy.

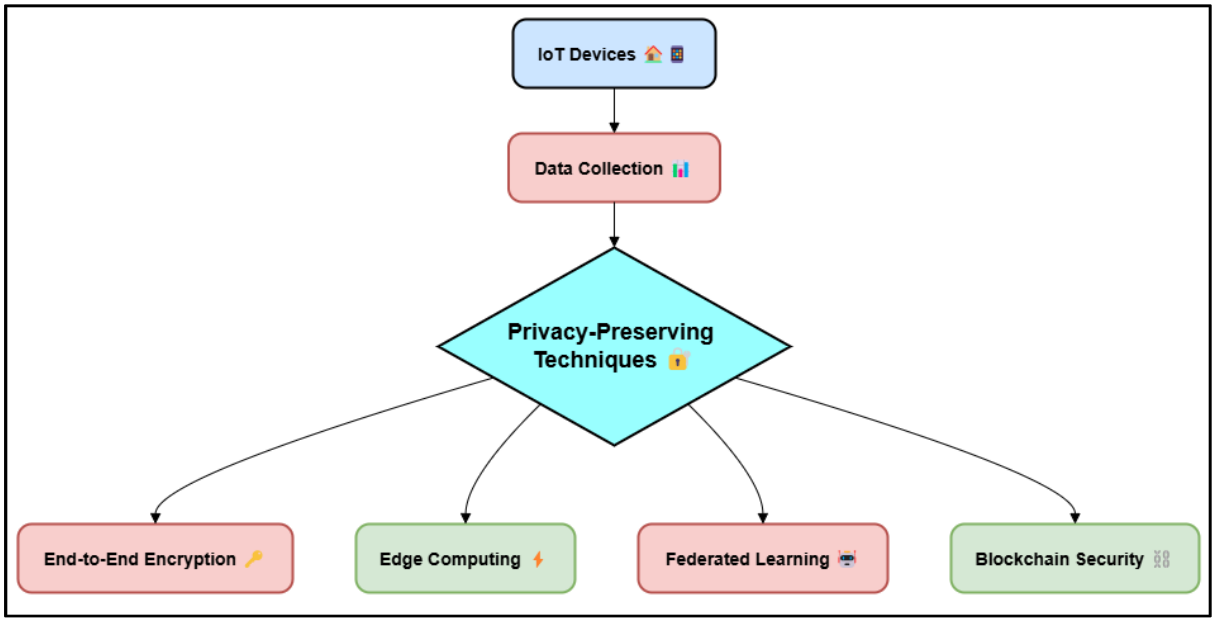

Figure 1: Illustrating Privacy-Preserving IoT Techniques

Anonymisation of data is another way to protect privacy in the IoT. Anonymisation makes sure that the data can't be linked to a specific person by hiding or removing personally identifiable information (PII). But this method might make the data less useful for personalised services because data that has been anonymised can lose its connection to the context. Another potential method is differential privacy, in which noise is added to the data to hide personal information while still letting the data be analysed. This method makes sure that the chance of finding specific people is kept to a minimum while still keeping the general statistical correctness of the data. A lot of people also use access control tools to keep IoT systems safe [4]. These systems control who can see certain data and provide ways to prove that they are who they say they are, so that only authorised users can see or change private data. Blockchain technology is a new way to protect privacy. It

provides a decentralised tracking system that can improve the security of data and make records of data transfers clear. Blockchain helps reduce privacy worries about centralised data keeping by making its design hard to change.

C. Introduction to Federated Learning

Federated Learning (FL) is a totally new way to deal with privacy issues in shared machine learning systems, especially in Internet of Things (IoT) settings. Federated Learning lets models be taught directly on edge devices instead of sending raw data to a central computer like standard centralised machine learning does. The best thing about FL is that the data stays on the devices. This makes it safer and more private for the data to be shared. After FL devices (called "clients") train a model on their own data, they only send model changes to a central computer, like slopes or parameters [5]. These changes are put together by the computer to make a global model. This model is then sent back to the devices to be trained again. Personal data and other private data never leave the device because of this decentralised method. This protects privacy. Federated Learning works really well in IoT settings, where devices often don't have a lot of resources and data protection is very important. FL makes sure that personal information stays private by getting rid of the need to send raw data. This lowers the risk of security leaks. FL can also support a lot of different Internet of Things (IoT) uses where data protection is very important, like smart healthcare, self-driving cars, and environmental tracking [6]. It also has big benefits for computing, since working close to home lowers the load on central computers and helps fix problems with network delay and bandwidth. Table 1 summarizes literature review on applications, benefits, challenges, and future trends in research. Federated Learning has a lot of promise, but it also comes with a lot of problems.

Table 1: Summary of Literature Review

Application	Benefits	Challenges	Future Trends
Smart Traffic Management	Real-time optimization of traffic flow, reduced congestion	Ensuring data accuracy and latency in real-time decisions	Integration of autonomous vehicles with FL for better predictions
Smart Healthcare Monitoring	Personalized health insights, privacy-preserving data sharing	Data heterogeneity across devices, model synchronization	Integration with wearable devices for continuous monitoring
Smart Energy Management [7]	Optimized energy consumption, reduced grid load	Scalability with increasing number of devices, security risks	Adoption of renewable energy data and integration with smart grids
Smart Parking Systems	Efficient parking management, real-time availability updates	Handling high communication costs, dealing with large datasets	Development of predictive models for parking space availability
Environmental Monitoring	Accurate air quality, noise pollution monitoring	Real-time data processing, device failures	Enhanced integration with weather forecasting systems
Smart Home Automation	Energy efficiency, convenience, user-centric control	Security vulnerabilities, device interoperability	Adoption of edge computing for localized decision-making
Autonomous Vehicles in Urban Settings	Safe, efficient navigation, real-time decision making	Data synchronization, ensuring data privacy	Collaborative vehicle-to-vehicle learning for improved safety
Smart Waste Management [8]	Optimized route planning, cost-effective waste collection	Reliability of sensor data, large-scale data processing	Use of drone-based sensors and real-time feedback systems
Public Safety and Surveillance	Improved crime prediction, better emergency response	Privacy concerns, ethical issues in surveillance	Implementation of real-time object detection with federated models
Urban Planning and Development	Data-driven infrastructure planning, better resource allocation	Data accuracy, conflict between different agent goals	Integration of crowdsourced data with federated learning for planning
Supply Chain and Logistics [9]	Efficient resource allocation, reduction of operational costs	Data fragmentation across various agents, real-time analytics	Adoption of blockchain for secure transaction handling
Water Management Systems	Real-time monitoring of water quality, efficient resource distribution	Data privacy in environmental sensors, integration with other systems	IoT-enabled predictive models for water consumption and quality

III. Federated Learning for IoT Privacy

A. Fundamentals of Federated Learning

Federated Learning (FL) is a decentralised way to do machine learning that lets models be trained on local devices instead of sending raw data to a central computer. In standard machine learning, data is gathered from many sources and put in central systems. The data is then handled and used to teach models. But this centralised method can be very bad for privacy because it requires sending private information that can be stolen, leaked, or used in the wrong way. FL gets around this problem by keeping data on the edge devices that create it, like smartphones, IoT monitors, or wearable tech. Then, each device uses its own data to train a model. It only sends model changes, like slopes or parameters, to a central computer [10]. These changes are sent to the central computer, which puts them all together to make a global model. This model is then sent back to the devices to be improved even more. FL is based on the idea that training models happens locally and that only updated models are sent to each other, not raw data. This design keeps private information on the device at all times, making it much more private while still letting useful trends be learnt from the data. In IoT systems, where there are a lot of gadgets and the data they produce is spread out across the network, FL can be very helpful. FL decentralises the learning process so that IoT devices can work together to make stronger models without sharing private data. This is very important in fields like healthcare, smart houses, and self-driving cars [11]. Each device in a FL setup teaches the model on its own, and the updates are sent together through safe methods to keep bad actors from messing with the system. This decentralised method not only keeps information private, but it also makes centralised computers' jobs easier and fixes problems with data traffic. Federated Learning is a great choice for big IoT networks, especially in critical areas, because it is fast and protects privacy.

- Step 1. Local Model Update (Client Update):

Each client (k) trains its model locally using its local data (D_k). The model update is given by:

$$[\ \theta_k^{\{t+1\}} = \theta_k^t - \eta_k \nabla_{\{\theta_k\} mathcal\{L\}_{k\left(\theta_k^t, D_k\right)]}$$

where:

- (theta_k^t) is the model parameters at time (t),

- (eta_k) is the learning rate,

- $(nabla_{\{theta_k\} mathcal\{L\}_{k\left(theta_k^t, D_k\right)}})$ is the gradient of the local loss function (mathcal{L}_k) with respect to the model parameters, computed on the local dataset (D_k).

- Step 2. Model Aggregation (Server Update):

After clients compute their local updates, the central server aggregates them to update the global model. The aggregation step typically uses a weighted average of the local model updates:

$$[\ \theta^{\{t+1\}} = \sum_{\{k=1\}}^{\{K\} frac\{n_k\}} \{N\} \theta_k^{\{t+1\}]}$$

where:

- (theta^{t+1}) is the updated global model at time (t+1),

- (n_k) is the number of data points at client (k),

- (N) is the total number of data points across all clients,

- (K) is the number of clients.

- Step 3. Iteration of Model Update:

The above steps are repeated for multiple iterations, with each client periodically sending its updated model to the server and receiving the updated global model for further training:

$$[\theta^{\{t+1\}} = \sum_{\{k=1\}}^{\{K\}frac\{n_k\}} \{N\}\theta_k^{\{t+1\}}, quad\ t\ =\ 0, 1, 2, \ldots]$$

where the iteration continues until the convergence criterion is met.

B. Privacy mechanisms in FL

Federated Learning (FL) adds a number of privacy-protecting features to make sure that private information stays private in spread settings. Because FL is based on the idea that data stays on the edge devices and only model changes are sent, it naturally keeps data from getting out. However, more privacy features are needed to keep the info even safer and make the FL process more secure. Differential privacy is one of the main ways that FL protects privacy. This feature adds noise to the model updates that devices share, making it impossible to quickly recreate or identify any one person's data from the collected updates [12].

C. Benefits of FL in smart city IoT applications

In smart cities, federated learning (FL) has many clear benefits, especially for Internet of Things (IoT) systems that care a lot about privacy, speed, and scale. One of the best things about FL is that it can help protect data privacy. In standard centralised machine learning systems, raw data from many IoT devices has to be sent to a central computer, which is very risky for privacy. FL, on the other hand, makes sure that private information stays on the local devices and that only model changes are sent along.

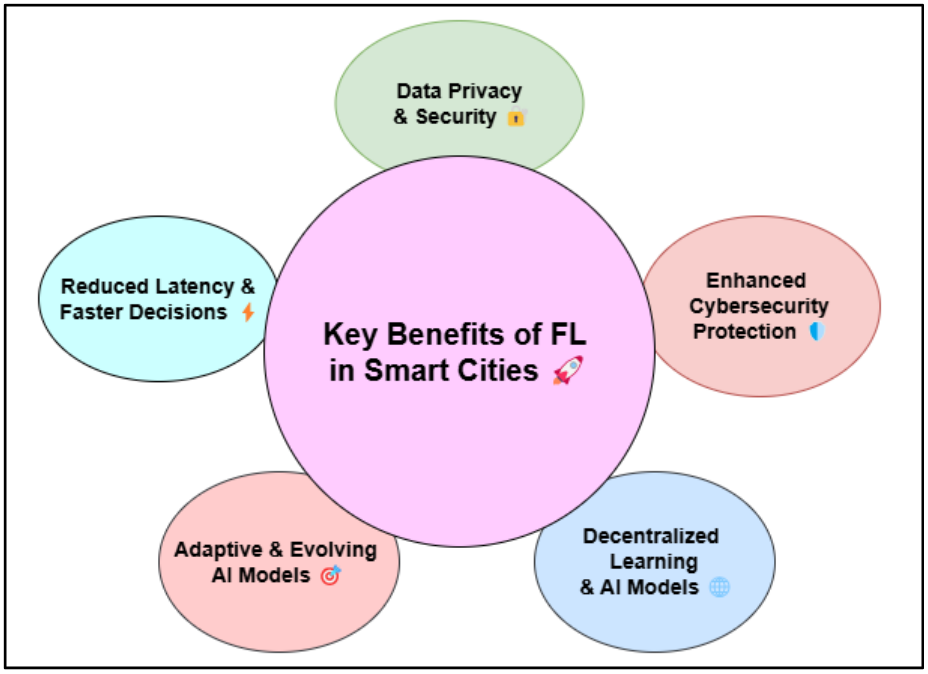

Figure 2: Illustrating Federated Learning Benefits in Smart Cities

This lowers the risk of data leaks and keeps private data, like health or location data, from getting out while the training is going on. The benefits of collaborative learning for improving data protection in smart cities are shown in Figure 2. FL also makes smart city IoT networks more scalable and efficient. There are a lot more IoT gadgets being added to smart towns all the time, and they are making a lot of data. The amount of data and the amount of work that needs to be done on central computers can make centralised processes slow down very fast. FL solves this problem by spreading the computing power across the devices. This lets the models be trained locally without having to constantly send and receive data [13]. This makes the system work better while protecting privacy. This way of working together makes models more accurate and flexible, which makes them more reliable in changing real-world settings. Lastly, FL lets you get personalised services while still protecting your privacy [14]. For instance, FL can make personalised guesses and suggestions in a smart healthcare app without sending private patient data. Based on local health data, each device, like a tracker, can help train the model. This lets for personalised health treatments and insights while keeping data private.

IV. Multi-Agent Systems in Smart Cities

A. Overview of multi-agent systems (MAS)

A couple of independent entities, or "sellers," have interaction with each other in a setting to attain their personal or the group's goals. This is what a Multi-Agent machine (MAS) is. In MAS, every agent can see what is occurring around it, determine what to do, and act based totally on its personal nation or outdoor cues. It is feasible for those bots to work by myself or with others to reply hard obligations. In MAS, sellers may additionally have exceptional degrees of understanding, but they speak to every other and paintings together to get matters accomplished that might be tough for one agent to do on their personal. Autonomy, decentralisation, and engagement are the most important elements of MAS. In MAS, marketers don't depend upon a government. Rather, they follow local rules and talk to different retailers to do their jobs. This decentralised method makes the system greater scalable and bendy, that's essential for makes use of in settings that are always converting and are very complex.

B. Position of MAS In Smart City IoT Networks

Multi-Agent systems (MAS) are very important in clever metropolis IoT networks due to the fact they make the networks more efficient, scalable, and bendy. With such a lot of IoT devices in clever cities, every one creates a massive amount of statistics that needs to be looked after through and used. With MAS, this fact is managed in a decentralised way, with every tool or display performing as its very own impartial agent and making decisions based totally on its surroundings. Those bots speak to each different, sharing records and making plans moves to attain shared dreams like making sure anybody is secure, enhancing traffic glide, or lowering electricity use. One of the satisfactory matters approximately MAS in clever metropolis IoT networks is that they can assist humans make decisions from one-of-a-kind places.

V. Implementation

A. System design and architecture

Federated Learning (FL) and Multi-Agent Systems (MAS) are combined in the suggested solution to make an IoT network system in smart towns that is decentralised and protects privacy. The system design is made up of several important parts that work together to make sure that the model is trained quickly, that data is handled safely, and that IoT devices (bots) can talk to each other reliably. The edge devices, which work on their own in the MAS, are at the heart of the system. IoT devices like smart monitors, personal tech, traffic cameras, and other things that collect data and are spread out across the city are among these. Each gadget is in charge of gathering data from its own area, processing it, and then using that data to build a machine learning model. The raw data is never sent to a central computer, which is very important for protecting user anonymity. The chief supervisor is in charge of making the whole system work. It takes model changes (like gradients or parameters) sent by edge devices and updates the world model by putting them all together.

B. Simulation environment and tools

To check and confirm the Federated Learning and Multi-Agent System approach, we need a virtual system that acts like an actual smart city where IoT devices would be used. This setting lets you test success indicators like protecting privacy, being able to grow, and being reliable. We use Python as our main computer language for the modelling and tools like TensorFlow Federated and Mesa to make Federated Learning work and simulate Multi-Agent Systems, respectively. TensorFlow Federated gives you tools to make FL models where the model's processing is spread out. This way, changes can be shared without sending raw data. In addition, it lets you simulate shared learning across multiple devices, which lets you test how well the system protects privacy in a controlled setting. Mesa is a Python tool for agent-based modelling that is used to model how IoT devices (agents) behave in a network.

C. Key features of the IoT network

Federated Learning and Multi-Agent Systems run the IoT network in a smart city. It has a number of important features that make processes quick, flexible, and private. Decentralisation, local model training, data protection, joint learning, and making decisions in real time are some of these features.

- Decentralization: The IoT community is constructed on a decentralised layout, because of this that every IoT tool (referred to as an "agent") works on its own. Those devices accumulate and method facts on-web site, so they do not want a central laptop to store or manner the facts. This decentralisation makes the system extra scalable and makes sure it can cope with the numerous devices that are not unusual in a smart town.

- Local version education: We don't ship uncooked statistics to a central laptop; instead, every device makes use of the records it receives to teach its very own gadget mastering model locally. Only version changes, like slopes or parameters, are sent to the central pc. This makes privacy a lot better with the aid of restricting the amount of information that is visible.
- Privacy renovation: The IoT network is built around protecting privacy. Federated getting to know and privacy-defensive strategies inclusive of differential privateness, safe aggregation, and encryption are utilized by the system to ensure that personal statistics is in no way made public. If attackers get into the interactions among gadgets and the central laptop, they won't be capable of get private statistics from model modifications because they are encrypted.
- Collaborative Learning: The gadgets within the network percentage adjustments to their models with each different to construct a global model. by means of using facts from distinctive assets, like climate monitors, healthcare gadgets, and clever meters, except breaching privateness, this joint learning makes positive that the model is usually getting higher.

VI. Conclusion

Federated Learning (FL) and Multi-Agent Systems (MAS) working together could be a good way to protect privacy in Internet of Things (IoT) systems used in smart city networks. Ensuring privateness and security is blanketed could be very challenging because IoT devices in smart cities create a number of private information. Federated learning protects privacy by using spreading out the learning method and ensuring that raw data by no means leaves the threshold devices. This lets smart town apps make smart decisions based on data. This look at delivered a new multi-agent framework that blends FL with the decentralised form of MAS. Each IoT tool works as its very own agent, able to acquire records, procedure it, and train nearby fashions. Version adjustments are amassed by using the central pc, which improves the sector model without affecting human being's privacy. This design cuts down on the amount of data that needs to be sent, which makes it harder for hackers to get into systems and steal data. Adding methods that protect privacy, like differential privacy, secure aggregation, and encryption, makes the system even better at keeping private data safe. The suggested answer also needs to be able to grow and be reliable in the big and changing settings that are typical of smart towns. As the number of IoT devices increases, the system can still change to keep working without affecting the safety or speed of data. The multi-agent Federated Learning model gives a safe, adaptable, and expandable framework for many smart city uses, like managing traffic, keeping an eye on healthcare, and making the best use of energy.

References

1. Zheng, G.; Kong, L.; Brintrup, A. Federated Machine Learning for Privacy Preserving, Collective Supply Chain Risk Prediction. Int. J. Prod. Res. 2023, 1–18.
2. Liu, Y.; Yu, W.; Ai, Z.; Xu, G.; Zhao, L.; Tian, Z. A Blockchain-Empowered Federated Learning in Healthcare-Based Cyber Physical Systems. IEEE Trans. Netw. Sci. Eng. 2022.
3. Wu, J.M.T.; Teng, Q.; Huda, S.; Chen, Y.-C.; Chen, C.-M. A Privacy Frequent Itemsets Mining Framework for Collaboration in IoT Using Federated Learning. ACM Trans. Sens. Netw. 2023, 19, 27.
4. Lu, Y.; Huang, X.; Dai, Y.; Maharjan, S.; Zhang, Y. Blockchain and Federated Learning for Privacy-Preserved Data Sharing in Industrial IoT. IEEE Trans. Ind. Inform. 2020, 16, 4177–4186.
5. Qammar, A.; Karim, A.; Ning, H.; Ding, J. Securing Federated Learning with Blockchain: A Systematic Literature Review. Artif. Intell. Rev. 2023, 56, 3951–3985.
6. Chen, Y.; Luo, F.; Li, T.; Xiang, T.; Liu, Z.; Li, J. A training-integrity privacy-preserving federated learning scheme with trusted execution environment. Inf. Sci. 2020, 522, 69–79.
7. Li, Y.; Chen, C.; Liu, N.; Huang, H.; Zheng, Z.; Yan, Q. A blockchain-based decentralized federated learning framework with committee consensus. IEEE Netw. 2020, 35, 234–241.
8. Imteaj, A.; Thakker, U.; Wang, S.; Li, J.; Amini, M.H. A survey on federated learning for resource-constrained IoT devices. IEEE Internet Things J. 2021, 9, 1–24.
9. Nguyen, D.C.; Ding, M.; Pathirana, P.N.; Seneviratne, A.; Li, J.; Poor, H.V. Federated learning for internet of things: A comprehensive survey. IEEE Commun. Surv. Tutor. 2021, 23, 1622–1658.
10. Zhang, T.; Gao, L.; He, C.; Zhang, M.; Krishnamachari, B.; Avestimehr, A.S. Federated learning for the internet of things: Applications, challenges, and opportunities. IEEE Internet Things Mag. 2022, 5, 24–29.
11. Issa, W.; Moustafa, N.; Turnbull, B.; Sohrabi, N.; Tari, Z. Blockchain-based federated learning for securing internet of things: A comprehensive survey. ACM Comput. Surv. 2023, 55, 1–43.

12. Gugueoth, V.; Safavat, S.; Shetty, S. Security of Internet of Things (IoT) using federated learning and deep learning—Recent advancements, issues and prospects. ICT Express 2023, 9, 941–960.
13. Tariq, A.; Serhani, M.A.; Sallabi, F.; Qayyum, T.; Barka, E.S.; Shuaib, K.A. Trustworthy federated learning: A survey. arXiv 2023, arXiv:2305.11537
14. Yaacoub, J.P.A.; Noura, H.N.; Salman, O. Security of federated learning with IoT systems: Issues, limitations, challenges, and solutions. Internet Things Cyber-Phys. Syst. 2023, 3, 155–179.