

Secure Federated Learning through Blockchain: A Compact Review of Challenges and Advances

مراجعة موجزة للتحديات والتطورات: التعلم الفيدرالي الآمن من خلال تقنية البلوك تشين

Sahar Yousif Mohammed سهر يوسف محمد

Translation Department-Arts College- Anbar University

جامعة الانبار -كلية الاداب- قسم الترجمة

Yousifsahar4@uoanbar .edu.iq

تاريخ تقديم البحث: 2025/7/24

تاريخ قبول البحث: 2025/8/24

Abstract

The exploration of decentralized solutions for training machine learning models via distributed clients; without the need for raw data sharing due ,Concerns regarding data privacy and ownership have prompted.In other hand, federated learning (FL) has appear as a promising method. Although, Federated Learning continues to faced with big challenges related to trust, transparency, and security, particularly in environments characterized by distrust or hostility. The integration of blockchain technology consider as great a complementary system to FL, Blockchain with FL become the provision of immutable record- keeping, decentralized control mechanisms; and governance frameworks based on smart contracts therefore offering the prospect to enhances its functionality. This compact review shows a comprehensive overview of the integration between Fedeareted Learning Technique(FL) and blockchain method. that represent by encompassing architectural frameworks; privacy preservation mechanisms; and data partitioning strategies. The following investigation will explore the ways in which blockchain technology how enhances model integrity also; stimulates participation; and reduces vulnerabilities . these ways represent poisoning and inference attacks. in this paper many subject discusion : real-world applications in healthcare; the Internet of Things (IoT), and emerging environments. Also, case studies are presented to demonstrate the viability of this integration. In fact,It is imperative to acknowledge that we are confronted with persistent issues, known : scalability, interoperability, and regulatory compliance. In this study, we have start a comprehensive initiative to explore the future trends in the development of lightweight; secure; and regulatory-compliant federated learning systems;finally, underpinned by blockchain technology.

Keywords: Block-chain, Decentralized Machine Learning(ML), Federated Learning (FL), Privacy-Preserving AI, Secure Model Aggregation, Horizontal Federated Learning (HFL), Vertical Federated Learning (VFL).

المستخلص

أدى استكشاف حلول لامركزية لتدريب نماذج التعلم الآلي عبر عملاء موزعين، دون الحاجة إلى مشاركة البيانات الخام، إلى إثارة المخاوف بشأن خصوصية البيانات وملكيته. من ناحية أخرى، برز التعلم الفيدرالي (FL) كطريقة واعدة. على الرغم من ذلك، لا يزال التعلم الفيدرالي يواجه تحديات كبيرة تتعلق بالثقة والشفافية والأمان، لا سيما في البيئات التي تتسم بعدم الثقة أو العدائية. يُعد دمج تقنية بلوك تشين نظامًا مكملًا للتعلم الفيدرالي، حيث يوفر حفظ سجلات ثابتًا، وآليات تحكم لامركزية، وأطر حوكمة قائمة على العقود الذكية، مما يتيح إمكانية تحسين وظائفه. يُقدم هذا الاستعراض الموجز نظرة شاملة على التكامل بين تقنية التعلم الفيدرالي (FL) وتقنية بلوك تشين، والتي تشمل الأطر المعمارية الشاملة، وآليات الحفاظ على الخصوصية، واستراتيجيات تقسيم البيانات. سيستكشف البحث التالي الطرق التي تُعزز بها تقنية بلوك تشين سلامة النموذج، وتُحفز المشاركة، وتُقلل من نقاط

الضعف. تُمثل هذه الطرق هجمات التسميم والاستدلال. تناقش هذه الورقة البحثية مواضيع متعددة، منها: التطبيقات العملية في مجال الرعاية الصحية، وإنترنت الأشياء (IoT)، والبيئات الناشئة. كما تُعرض دراسات حالة لإثبات جدوى هذا التكامل. في الواقع، من الضروري الاعتراف بأننا نواجه تحديات مستمرة، تُعرف باسم: قابلية التوسع، والتوافق التشغيلي، والامتثال التنظيمي. في هذه الدراسة، أطلقنا مبادرة شاملة لاستكشاف الاتجاهات المستقبلية في تطوير أنظمة تعلم اتحادية خفيفة الوزن، وأمنة، ومتوافقة مع اللوائح التنظيمية، مدعومة بتقنية البلوك تشين.

الكلمات المفتاحية: تكامل البلوك تشين، التعلم الآلي المركزي، التعلم الفيدرالي (FL)، الذكاء الاصطناعي الذي يحافظ على الخصوصية، تجميع النماذج الآمن، التعلم الفيدرالي الأفقي (HFL)، التعلم الفيدرالي الرأسي (VFL).

Introduction

Over the past four years, concerns about privacy and the potential of deep learning have led to a radical change in the application of machine learning (ML). A new model called federated learning (FL) has emerged as a new model for applying machine learning, offering an alternative to both centralized systems and field analysis its is a decentralized, privacy-preserving approach that stores raw data on devices and trains machine learning locally, reducing the burden of data communications. A federation of learned and shared models is then implemented on a central server to aggregate and share the knowledge built among participants (Zalik & Žalik, 2023, 4). In the federated learning model, the central server and local edge devices share the same model by exchanging updates rather than raw data. This methods protects the privacy of data stored on edge devices. the central server and local edge devices maintain the same model; By sharing model updates instead of raw data in the federated learning.. This methods protects the privacy of data stored on edge devices , that because it does not directly shows the data. This approachs reduce privacy violations caused by the increasing collection of sensitive data. Although FL with a central server has become more commons, It faces performance bottlenecks as new threats grow. The most important reasons are centralized processing; data tampering; and lack of incentives. That make to accelerate the adoption of FL, blockchain-supported FL has attracted very important interest from academia and industry. A large number of innovatives solutions have been created to support diverse use cases. Blockchain-supported federated learning provide theories and techniques to enhancing the performances of federated learning from different dimensiones (Wang & Hu, 2021, 1-18) In the last evolving field of data-driven intelligence; federated learning (FL) has appear as a promising technique; This approach is preserves privacy while enabling collaboration in machine learning. it still faces serious issues, such as model poisoning; trust concerns; and centralized coordination which threaten its effectiveness (Yu, Shen, Wang, Zhang, & Zhao, 2024, 5619). The decentralized and tamper-resistant nature of blockchain architecture; Offers a solution to enhances the security; transparency, and reliability of FL systems. This review give us the basic concepts such as : integration methods, and real-world applications of blockchain-enabled FL without rummage into excessive technical details. The compact review highlights the strengths and limitations of current approaches. Where presents key classifications of Federated Learning architectures and data partitioning strategies. Also identifies common security vulnerabilities; and explores how blockchain technologies: such as smart contracts and consensus protocols. That can helps overcome these vulnerabilities. The review also examines regulatory and ethical considerations theses consideration represent as: data ownership, transparency, and compliance with laws .example of law the General Data Protection Regulation (GDPR), and identifies future directions for building more powerful, scalable, and reliable decentralized AI systems (Awosika, Shukla, & Pranggono, 2024). This paper provides a brief analysis of the integration of clustering learning and blockchain technology. that covering architectural categories, data partitioning, and security challenges. While many previous reviews focus on either clustering or blockchain separately. this paper provides an integrated analysis of both. It also highlights recent techniques, such as secure clustering; swarm learning, and SplitFed that are often overlooked in previous surveys. Our contributions include:

- (1) A categorized overview of FL architectures and partitioning strategies;
- (2) A comparison of traditional FL models and blockchain-based FL models;
- (3) Detailed discussions on privacy, GDPR compliance, and attack mitigation
- (4) real-world applications across healthcare, the Internet of Things (IoT), and autonomous systems; and
- (5) A detailed outline of the current limitations of, and future research directions for, deploying scalable, secure, decentralized AI systems.

Classical ML vs. Federated Learning








Characteristic	Classical ML	Federated Learning
 Data Handling	Raw data to server	Data stays on device
 Privacy Level	Low	High
 Architecture	Centralized	Decentralized/Distributed
 Communication	Data to server	Model between clients
 Scalability	Limited	High
 Security Risks	Data breaches	Model poisoning
 Use Cases	Cloud-based ML	Mobile, IoT, Healthcare

Figure 1 compares traditional machine learning (collecting all data in one place) and federated learning (keeping data on local devices and only sharing model updates). This make federated learning more private and secure (Beltrán et al., 2022, 2983-3013).

2. Background

2.2 Machine Learning Overview

Machine learning (ML) is a technique that enables systems to make accurate predictions or decisions by learning from data. Recent advances of data storage and computational power has accelerated the development of ML techniques via various fields; including bioinformatics. In fields like biology; Where experiments are expensive and complex; Machine Learning give valuable tools for data analysis and prediction (Drainakis et al., 2020, 1-8). Machine Learning systems usually learned from data by three main methods: Supervised, Unsupervised, and Reinforcement learning. Supervised learning involves using labeled datasets to train models to map input features (independent variables) to desired outputs (dependent variables). Unsupervised learning concerns discovering hidden patterns or grouping in data. These patterns without predefined labels. Reinforcement learning operate Through feedbacks mechanisms; Where the system learns via receiving rewards or penalties based on its actions. Traditional centralized Machine Learning methods depending on collecting data from multiples sources and processing it on a central server. While this setups is straightforward; It raises important privacy, security, and communication concerns (Verbraeken et al., 2020,1-33) . when dealing with sensitive data such as : medicals records or financial information. The transmitting raw data over networks increase the risks of breaches and storing. where everything in one place creates a single point of failure. then, centralized approaches struggle in environments ;Where bandwidth is limited or real-time processing is needed, such as in healthcare and IoT applications (Naik & Naik, 2023, p.18–28), Rahman et al., 2021, 124682–124700)

2.2 Federated Learning (FL)

Federated Learning (FL) offers a privacy-preserving alternative to centralized ML. In FL, the data remains on local devices (e.g., smartphones, hospitals), and only model updates not raw data are shared with a central server or among peers. This approach allows collaborative model training across decentralized systems without compromising data privacy. FL is particularly well-suited for sensitive domains that require strict data protection. Federated learning enables powerful models to be built by combining knowledge from multiple sources while protecting data ownership and security. However, Federated Learning proposes unique challenges: such as handling non-identically distributed (non-IID) data; managing device variability; ensuring communication efficiency; and maintaining robust privacy protocols. Federated Averaging (FedAvg) is a core algorithm in federated learning technique. Federated Averaging (FedAvg) that enables multiple clients to collaboratively train a global model without sharing raw data. In spite of simplicity, FedAvg has been shown to generalize well when properly tuned. That making it one of the most widely used techniques in FL (Wen et al., 2023:513–535). Figure 2 shows the basic workflow of federated learning. Each client (like a device or organization) trains a local model on its own data; then sends only the updated model not the raw data to a central server. The server aggregates all updates to create a global model and sends it back to the clients for the next training round. This process repeats, allowing the model to improve collaboratively without sharing private data.

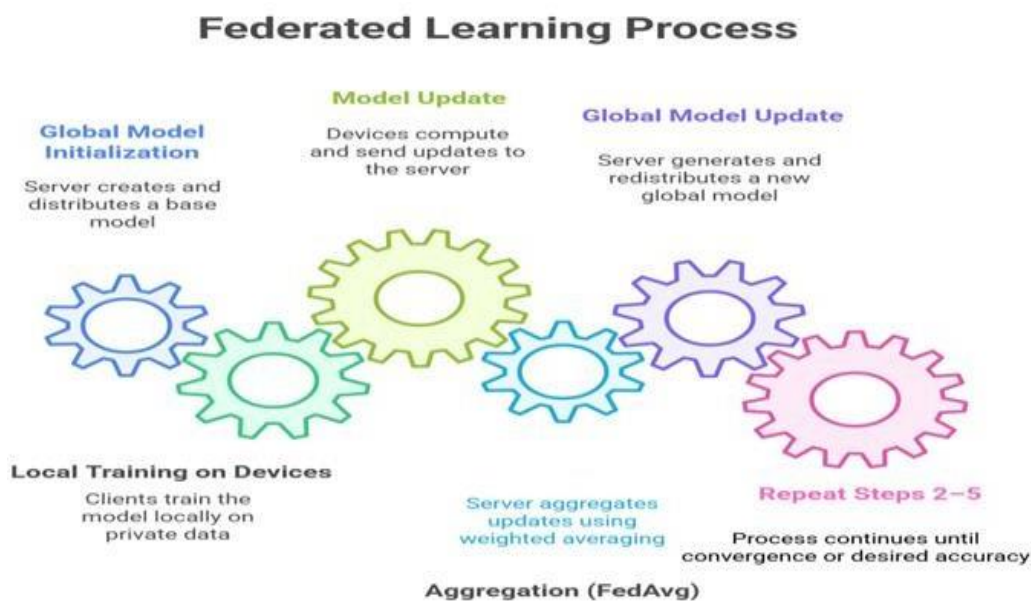


Figure2. Federated Learning Process.

3. Federated Learning Categories and Architecture

Data partitioning is a fundamental concept in both distributed and federated machine learning, as it determines how data is split among participants and how models are collaboratively trained. There are three main types of partitioning: horizontal, vertical, and hybrid each with its own challenges and use cases. In some cases, transfer learning techniques are employed to transfer knowledge across these partitions (Antunes et al., 2022,1-23). Horizontal partitioning (HFL) refers to splitting data by rows. Each participant (e.g., hospital or mobile device) holds data samples with the same features but from different users. Vertical partitioning (VFL) divides data by columns. Here, different entities possess different features about the same individuals. For example, a bank and a telecom company may both hold different types of information on the same customer. Hybrid partitioning combines both horizontal and vertical methods, requiring more complex federated strategies to handle data diversity across both dimensions. Federated learning

can be deployed using two core architectures.

Centralized Federated Learning (CFL) relies on a central server to collect and aggregate model updates from clients.

Decentralized Federated Learning (DFL) eliminates the central server; instead, model updates are shared peer-to-peer among participants.

Choosing between these architectures involves trade-offs in privacy, scalability, fault tolerance, and communication efficiency. FL settings can also be categorized by their scale and trust level:

- ✓ Cross-silo FL involves a small number of reliable, high-resource participants (e.g., hospitals, universities).
- ✓ Cross-device FL supports massive-scale collaboration across many unreliable, resource-limited devices (e.g., smartphones).

Each scenario introduces different challenges related to data heterogeneity, system coordination, and security. Recently, hybrid and hierarchical architectures have been proposed to address these challenges and improve scalability, model accuracy, and resilience (Lo et al., 2022, 111357) - (Shanmugam, Tillu, & Tomar, 2023, p.371–384). Figure 3. Key Design Aspects of Federated Learning (FL) Systems: Four Major Aspects Influence FL Implementation: Architecture Types, Federation Scale, Privacy Mechanisms, And Data Partitioning Strategies. Each Aspect Plays a Critical Role In FL Performance, Scalability, and Security.

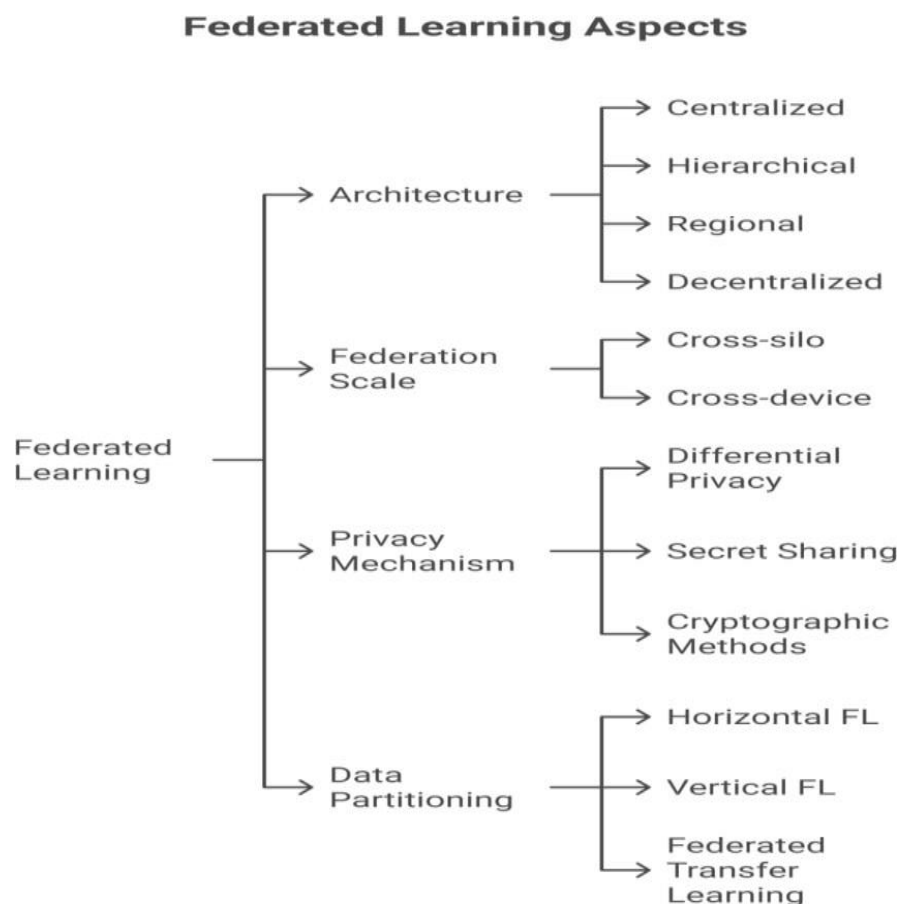


Figure 3. Key Design Aspects of Federated Learning (FL) Systems: Four Major Aspects Influence FL Implementation: Architecture Types, Federation Scale, Privacy Mechanisms, And Data Partitioning Strategies. Each Aspect Plays a Critical Role In FL Performance, Scalability, And Security.

Table 1 below provides a concise overview of key federated learning types, their data distributions challenges, and typical application domains.

Table 1. Design aspects of Federated Learning (FL): Four factors influence FL implementation: architecture, scale, privacy, partitioning. Each impacts performance, scalability, and security.

FL Type	Data Distribution	Key Challenges	Typical Applications
Horizontal (HFL) (Yang et al., 2020, p.49–67)	Same features, different users (samples)	Non-IID data, communication overhead, privacy leakage via updates	Mobile keyboard prediction, hospital collaborations
Vertical (VFL) (Liu et al., 2024, p.3615–) (3634)	Same users, different features	Complex encryption, alignment of shared entities, model synchronization	Financial fraud detection (banks + telecom), healthcare merge
Federated Transfer Learning (FTL) (Saha & Ahmad, 2021,) (p.35–44)	Different users and different features	High complexity, domain adaptation, less overlap in data	Retail & telecom analytics, inter-region education systems

3.Blockchain's Contribution to Overcoming Federated Learning Challenges

In federated learning architectures, Blockchain provides important developments in addressing key weaknesses. These developments represent in terms of privacy, transparency, and incentive mechanisms. Though federated learning is designed to protect data privacy by processing data locally; and sharing gradients only with a central server. this dependence on a central authority can create vulnerabilities. Blockchain technology help this problem via empowering a decentralized approach. That provides greater protection for user privacy that happen by eliminating the single point of failure and control associated with a central server. In addition, the integration of blockchain technology extremely enhance transparency in federated learning (Awosika, Shukla, & Pranggono, 2024, 64551–64560). blockchain's fixed and distributed ledger can solve the issue of lack of transparency that is often found in centralized federated learning systems . this ledger can accurate record all transactions and model upgrades. That making the entire training process verifiable and transparent to all involved entities. This transparency enhances trust among participants regarding the aggregation process and the overall integrity of the global model. With respect to incentives although the submitted text does not outline a broad framework for incentivization, It does introduce a novel data evaluation scheme that leverages miner participation. This refers to a mechanical procedure whereby blockchain contributors. That technique motivates parties to contribute high-quality data or computing resources. blockchain lays the groundwork for powerful incentive structures by providing data quality and a verifiable record,. Sometimes, integrating blockchain into federated learning creates a more decentralized; transparent, and reliable environment. Therefore, enhancing user privacy and enabling verifiable data contribution (Javed et al., 2022, 4394) .In federated learning (FL) Smart contracts are becoming increasingly common to automate the management of rules and rewards. They facilitate transparent; secure; and fair sharing between data owners and contributors. These contracts can distribute rewards such as tokens or NFTs. These contracts can be customised to protect privacy, enhance outcomes;and accelerate service. Smart contracts can also handle important tasks such as : registering users, checking updates, distributing rewards, and managing trust. Also, they protect the system from malicious activities such as: cheating; colluding, or submitting fake data. Smart contract make federated learning more reliable, scalable; and useful in areas such: as AIoT, healthcare, and smart industries Figure 4 provides a high level visualization of the integration of blockchain technology with federated learning and how forms are exchanged and transactions are recorded.

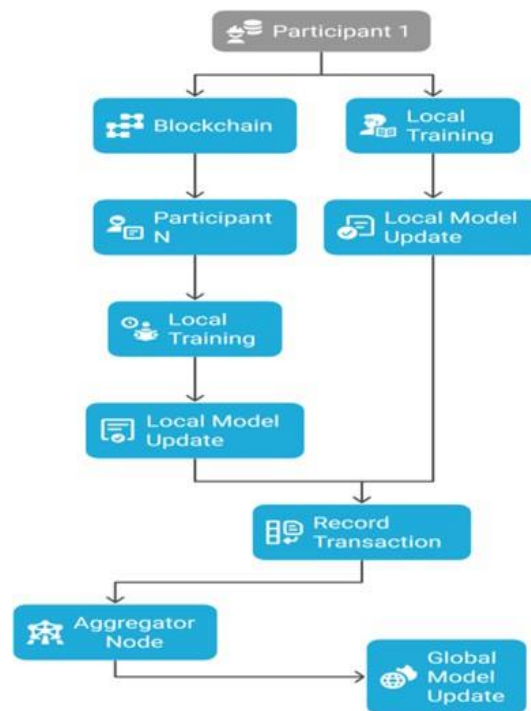


Figure 4.Blockchain-based Federated Learning (FL) paradigm, and Blockchain-based FL.

Table below provides a comparative analysis between: Traditional Federated Learning (FL) and Blockchain-based FL. This analysis covers various key dimensions, including transparency; trust, scalability; and auditability, to better understand the advantages and challenges of integrating blockchain technology with Federated Learning.

Table 2.Comparison between Federated Learning (FL) and Blockchain-based FL

Aspect	Federated Learning	Blockchain-based FL
Trust (Karandikar et al., 2021, p. 3822)	Relies on a central server	Trustless; relies on blockchain for consensus and auditability
Transparency (Liu et al., 2024, p. 4377)	Limited (central aggregator)	High all updates and transactions are logged on blockchain
Security (Issa et al., 2023, p. 1)	Vulnerable to poisoning attacks, central point of failure	More power decentralized validation and tamper-proof logs
Scalability (Ahmed & Alabi, 2024, p. 102219)	Centralized bottleneck limits scalability	More scalable via P2P structure and smart contracts
Privacy (Zhao et al., 2020, p. 1817)	Strong data privacy (no raw data shared)	Same, plus immutable access records
Auditability (Kalapaaking et al., 2024, p. 102)	Difficult to verify participant behavior	Fully auditable via blockchain ledger
Fault Tolerance (Mounnan et al., 2020, p. 347)	Single point of failure (aggregator)	Decentralized: less risk of complete system failure
Incentivization (Wang et al., 2023, p. 1536)	Not built-in	Can use tokens to reward participation and honesty

4. Security & Privacy Challenges

Blockchain technology can help people control their personal data. This is done by storing it in a decentralized and secure method; Instead of depending on companies to manage their data. users can decide who can see or use it by using smart ; contracts. They can give, remove, or change access at any time. Blockchain also has a clear re+-cord of who has permission .Advanced tools such as encryption help maintain data privacy; Even when it is shared or processed so users know what's happening with their data. This is helpful in fields such as: healthcare or finance where data is sensitive. blockchain usually difficult such as : being slow; hard to use or super expensive. It also needs to be suitable with privacy laws such as: the General Data Protection Regulation (GDPR). Block-chain has a lot of assurance for protecting privacy and giving people more control even with these issues (Ye, Luo, Yang, Choo, & He, 2023). Table 3 shows common attack type in federated learning. Also highlight how smart mechanisms sometimes supported by blockchain, that can help defends against these security and privacy threats.

Table3. Summarizing Attack Types and Mitigation Techniques.

Attack Type	Example Defense Mechanism
Poisoning	Robust aggregation , anomaly detection (Wu et al., 2024, p. 4744)
Backdoor	Pre-aggregation , neuron activation (Hao et al., 2025,p.15)
Membership Inference	Differential privacy (Chen et al., 2020, p. 26)
Free-Rider	Anomaly detection , reward mechanisms (Wang et al., 2023, p. 4377)

5. Blockchain Technology Essentials

In the healthcare area, Federated learning (FL) and block-chain are becoming more integrated to allow for secure, privacy-preserving, and collaborative training of AI models via multiple organisations. This integration between FL and Blockchain manages crucial challenges related to data sharing, privacy, and trust. Especially as healthcare data continues to grow in volume and sensitivity. Both horizontal and vertical Federated Learning benefits from the decentralised and tamper-resistant architecture of blockchain. Horizontal federated learning is a process where institutions such as; hospitals with similar types of data for instance (medical images or patient records), collaborate to train models without sharing raw data. Applications & Case Studies

The integration of federated learning (FL) and blockchain technology has paved the way for innovative applications in diverse fields by enhanced data privacy, security, and collaboratively-based machine learning. This combine is of particular value in the environments that deal with sensitive data, as it provides for decentralized models training without compromise to individual privacy. In the healthcare sector, FL enables the collaborative training of decentralized patient data to enhance disease predicts models, with the blockchain technology ensuring the secure storage and sharing of health records, protecting them from unauthorized access(Li et al., 2021, p. 1) . Similarly, in the Internet of Things (IoT) domain, FL is being applied in smart cities to optimize traffic flow and resource Allocation, with blockchain ensured the security of data transfers between interconnected devices. In the industrial IoT setting, the combine use of FL and blockchain enables predictive maintaining and enhances operating efficiency by analysis of data from distributed sensors without risking data safety. The automotive Industry also benefits from this integration, particular in autonomous vehicleswhereby FL facilitates real-time learning from distributed data sources to enhance safety, while blockchain ensures secure communication between vehicles. Despite its advantages, this integrated Approach still faced challenges such as system heterogeneity and high communication costs, which must be addressing for its wider adoption (Boudjemaa, 2024, p. 15)the following figure shows how federated learning can be applied in various domains, from

data generation to global model optimization, collaboratively and securely((Han & Park, 2022, p. 103888), (Said et al., 2023, p. 198).

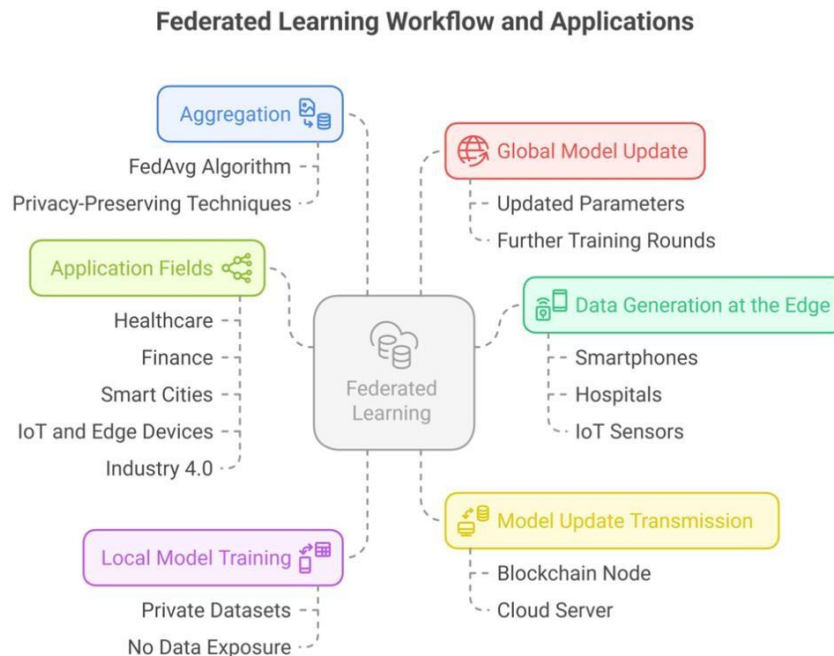


Figure4. Federated learning in different fields to ensure privacy and collaborative learning without sharing raw data.

6. Regulatory & Ethical Considerations

8.2GDPR, CCPA, and Blockchain Immutability Conflicts

Blockchain technology has emerged with transparency, security; and stability due raising serious concerns about compliance with privacy regulations. These concern represent by some rules such as the General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA). blockchain systems are designed to be tamper-proof and maintain permanent records. Though privacy protection laws grant individuals basic rights over their personal data. These rights contain the ability to amend or delete data, also known as the right to be forgotten. The immutable nature of blockchain technology presents significant legal challenges because it is difficult to change. If not impossible to implement these rights in systems that forever and indelibly store data. The decentralized nature of blockchain technology makes identifying data controllers or processors difficult, which adds another layer of legal complexity when determining liability or implementing rights (Tiwari, 2024, p. 822). A range of technical and structural solutions has been proposed to overcome these challenges. Although there is no perfect solution. One of the most advanced options is editable blockchains, which use cryptographic techniques, such as variable hashing or dynamic, attribute-based encryption, that allow for the selective modification of specific data blocks. While these methods are innovative, they can compromise the integrity and trustless nature of blockchain systems. Another widely explored process is using pseudonyms, where only pseudonymous data is stored on the blockchain, and any personally identifiable information is securely kept off the blockchain. This helps to make anonymised data, but it transfers the responsibility of security to another location. Permissioned blockchains offer another path by restricting access and clearly defining roles within the system, making compliance easier, but that effect at the expense of decentralization (Chen & Lobo, 2024, p. 871). Although these methods but still ongoing challenges. Legal uncertainties are still an issue, especially when it comes to defining roles and responsibilities in decentralized ecosystems. When it comes to

technology, using blockchain systems often means making choices that balance how well those systems meet certain standards. Regardless, these choices sometimes make the systems less decentralized or less secure. Researchers are still actively exploring how to balance innovation and regulation. Until clear legal frameworks evolving and more adaptable blockchain architectures are developed, this tension between data permanence and privacy rights is likely to continue (Ye et al., 2023, p. 1669).

Table4. Key Approaches to Blockchain–GDPR/CCPA Conflicts

Approach	How It Works	Limitations/Notes
Redactable Blockchains	Use cryptographic methods (e.g., chameleon hashes) to allow selective data editing	May weaken blockchain security; adds complexity (Lapwattanaworakul et al., 2023,p2470-2480.)
Pseudonymization	Store pseudonymized data on-chain; keep identifiers off-chain	Requires secure, reliable off-chain storage (Polge et al., 2021, p. 229)
Permissioned Blockchains	Access is limited to verified participants; roles clearly defined	Reduces decentralization; better accountability (Asgarinia et al., 2023, p. 351)

8.2 Ethical Issues: Data Ownership, Fairness, and Decentralization

Current discussions refer to how to manage data in a world that increasingly uses digital technology. They are focused on ethical issues related to data ownership, fairness, and decentralisation. Data ownership is considered as one of the main concerns of data not only in terms of legality but also regarding control, acknowledgement and honest sharing of benefits. Ownership data is represented as super important in areas like healthcare and agriculture, Where unclear data ownership can lead to exploitation or misuse of sensitive informations. The goal are shift the power dynamics in the data economy to prioritize the rights of individuals and communities.

Equality is another challenge due to algorithmic bias unequal access to data, and opaque decision-making processes. equality requires inclusive data processing practices, continuous bias detection, and transparency in AI systems. Ethical design principals and regulations such as the General Data Protection Regulation (GDPR); play a key role in supporting these efforts. Ultimately decentralization through systems such as blockchain and federated learning provides a way to empower users and reduce centralized control. these systems can enhance privacy and distribute decision-making power, which sounds ideal. They also cause complications, such as reduced accountability, fragmented oversight and the need to balance technical performance with ethical safeguards. So, although decentralization has potential, it is not a panacea (Mustafa et al., 2025, p. 37).

Table 5. Ethical Issues in Data-Driven Systems

Issue	Key Concerns & Potential Solutions
	Calls for user control, benefit-sharing, and clear governance over data use
Fairness	Tackling bias, preventing discrimination, ensuring transparency, and complying with regulations
Decentralization	Balancing privacy and distributed control with accountability and system efficiency

Conclusion and Future Work

This compact review presented how federated learning can be integrated with blockchain technology, and how this technology has contributed to practical, secure solutions for AI systems. It also included several privacy-preserving solutions, including data decentralization. It also explained how we can address the challenges faced by federated learning in terms of transparency, trust, and incentives while preserving data privacy through blockchain's decentralized ledger. While this combined technology has solved many challenges, such as healthcare analytics, Internet of Things environments, and trusted autonomous systems, other challenges or consequences remain, such as high communication costs, scalability limitations, and complexities that still prevent the deployment of these technologies. The results of this study emphasize the importance of developing lightweight blockchain architectures, mechanisms for maintaining opinion privacy, and advanced trust assessment techniques to improve the feasibility of blockchain-enhanced federated learning. Future studies should prioritize real-world benchmarking, establishing large-scale testbeds, and validating hybrid models that combine privacy protection and efficiency. Addressing these gaps will help transform the reviewed frameworks from theoretical models into deployable, regulatory-compliant, scalable, and industrial-ready systems.

References

1. A. Ahmed and O. O. Alabi, "Secure and scalable blockchain-based federated learning for cryptocurrency fraud detection: A systematic review," *IEEE Access*, vol. 12, pp. 102219–102241, 2024. doi: 10.1109/ACCESS.2024.3412963.
2. R. S. Antunes, C. André da Costa, A. Küderle, I. A. Yari, and B. Eskofier, "Federated learning for healthcare: Systematic review and architecture proposal," *ACM Trans. Intell. Syst. Technol. (TIST)*, vol. 13, no. 4, pp. 1–23, 2022. doi: 10.1145/3510576.
3. H. Asgarinia, A. Chomczyk Penedo, B. Esteves, and D. Lewis, "Who should I trust with my data? Ethical and legal challenges for innovation in new decentralized data management technologies," *Information*, vol. 14, no. 7, p. 351, 2023. doi: 10.3390/info14070351.
4. T. Awosika, R. M. Shukla, and B. Pranggono, "Transparency and privacy: The role of explainable AI and federated learning in financial fraud detection," *IEEE Access*, vol. 12, pp. 64551–64560, 2024. doi: 10.1109/ACCESS.2024.3390878.
5. E. Beltrán, M. Pérez, P. S'anchez, S. Bernal, G. Bovet, M. Pérez, G. P'erez, and A. Celdrán, "Decentralized federated learning: Fundamentals, state of the art, frameworks, trends, and challenges," *IEEE Commun. Surveys Tuts.*, vol. 25, pp. 2983–3013, 2022. doi: 10.1109/COMST.2023.3315746.
6. Y. Boudjemaa, "Ensuring regulatory compliance in cloud-based big data systems: A framework for global operations adhering to GDPR and CCPA," *Stud. Knowl. Discov. Intell. Syst. Distrib. Anal.*, vol. 14, no. 9, pp. 15–27, 2024.
7. J. Chen, W. H. Wang, and X. Shi, "Differential privacy protection against membership inference attack on machine learning for genomic data," in *BIOCOMPUTING 2021: Proc. Pac. Symp.*, 2020, pp. 26–37.
8. S. Chen and B. C. Lobo, "Regulatory and implementation considerations for artificial intelligence," *Otolaryngol. Clin. North Am.*, vol. 57, no. 5, pp. 871–886, 2024. doi: 10.1016/j.otc.2024.05.004.
9. G. Drainakis, K. V. Katsaros, P. Pantazopoulos, V. Sourlas, and A. Amditis, "Federated vs. centralized machine learning under privacy-elastic users: A comparative analysis," in *Proc. 2020 IEEE 19th Int. Symp. Netw. Comput. Appl. (NCA)*, Nov. 2020, pp. 1–8. doi: 10.1109/NCA51143.2020.9306687.
10. S. Han and S. Park, "A gap between blockchain and general data protection regulation: A systematic review," *IEEE Access*, vol. 10, pp. 103888–103905, 2022. doi: 10.1109/ACCESS.2022.3215004.
11. L. Hao, K. Hao, B. Wei, and X. S. Tang, "Multi-target federated backdoor attack based on feature aggregation," *arXiv preprint arXiv:2502.16545*, 2025.

12. W. Issa, N. Moustafa, B. Turnbull, N. Sohrabi, and Z. Tari, "Blockchain-based federated learning for securing internet of things: A comprehensive survey," *ACM Comput. Surv.*, vol. 55, no. 9, pp. 1–43, 2023. doi: 10.1145/3569905.
- A. R. Javed et al., "Integration of blockchain technology and federated learning in vehicular (IoT) networks: A comprehensive survey," *Sensors*, vol. 22, no. 12, p. 4394, 2022. doi: 10.3390/s22124394.
13. P. Kalapaaking et al., "Auditable and verifiable federated learning based on blockchain-enabled decentralization," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 36, no. 1, pp. 102–115, 2024. doi: 10.1109/TNNLS.2022.3231209.
14. N. Karandikar, A. Chakravorty, and C. Rong, "Blockchain based transaction system with fungible and non-fungible tokens for a community-based energy infrastructure," *Sensors*, vol. 21, no. 11, p. 3822, 2021. doi: 10.3390/s21113822.
15. J. Lapwattanaworakul, C. Srisa-An, and T. Aribarg, "Blockchain-based auxiliary systems for pseudonymization and consent management," *TEM J.*, vol. 12, no. 4, pp. [pages missing], 2023.
16. D. Li, Z. Luo, and B. Cao, "Blockchain-based federated learning methodologies in smart environments," *Cluster Comput.*, pp. 1–15, 2021. doi: 10.1007/s10586-021-03424-y.
17. J. Liu et al., "Enhancing trust and privacy in distributed networks: A comprehensive survey on blockchain-based federated learning," *Knowl. Inf. Syst.*, vol. 66, no. 8, pp. 4377–4403, 2024. doi: 10.1007/s10115-024-01917-7.
18. Y. Liu et al., "Vertical federated learning: Concepts, advances, and challenges," *IEEE Trans. Knowl. Data Eng.*, vol. 36, no. 7, pp. 3615–3634, 2024. doi: 10.1109/TKDE.2022.3212710.
19. S. K. Lo, Q. Lu, L. Zhu, H. Y. Paik, X. Xu, and C. Wang, "Architectural patterns for the design of federated learning systems," *J. Syst. Softw.*, vol. 191, p. 111357, 2022. doi: 10.1016/j.jss.2022.111357.
20. S. Y. Mohammed, M. Aljanabi, and M. M. Mijwil, "Detecting denial-of-service (DoS) attacks with edge machine learning," in *Proc. 3rd Finance, Accounting, and Law in the Digital Age Conf.*, Salamanca, Spain, May 2024, pp. 119–127. doi: 10.1007/978-3-031-67511-9_8.
21. K. Mohanta et al., "Addressing security and privacy issues of IoT using blockchain technology," *IEEE Internet Things J.*, vol. 8, no. 2, pp. 881–888, 2020. doi: 10.1109/JIOT.2020.3000423.
22. O. Mounnan et al., "Privacy-aware and authentication based on blockchain with fault tolerance for IoT enabled fog computing," in *Proc. 2020 5th Int. Conf. Fog Mobile Edge Comput. (FMEC)*, Apr. 2020, pp. 347–352. doi: 10.1109/FMEC49853.2020.9144759.
23. G. Mustafa, W. Rafiq, N. Jhamat, Z. Arshad, and F. A. Rana, "Blockchain-based governance models in e-government: A comprehensive framework for legal, technical, ethical and security considerations," *Int. J. Law Manag.*, vol. 67, no. 1, pp. 37–55, 2025.
24. Naik and N. Naik, "The changing landscape of machine learning: A comparative analysis of centralized machine learning, distributed machine learning and federated machine learning," in *Proc. UK Workshop Comput. Intell.*, Cham: Springer Nature Switzerland, Sep. 2023, pp. 18–28.
25. J. Polge, J. Robert, and Y. Le Traon, "Permissioned blockchain frameworks in the industry: A comparison," *ICT Express*, vol. 7, no. 2, pp. 229–233, 2021. doi: 10.1016/j.icte.2020.06.002.
26. K. J. Rahman et al., "Challenges, applications and design aspects of federated learning: A survey," *IEEE Access*, vol. 9, pp. 124682–124700, 2021.
27. S. S. Rao, S. L. Fernandes, C. Singh, and R. R. Gatti, Eds., *AIoT and Big Data Analytics for Smart Healthcare Applications*, vol. 5. Sharjah, UAE: Bentham Science Publishers, 2023.
28. S. Saha and T. Ahmad, "Federated transfer learning: Concept and applications," *Intelligenza Artificiale*, vol. 15, no. 1, pp. 35–44, 2021. doi: 10.3233/IA-210107.
29. L. Shanmugam, R. Tillu, and M. Tomar, "Federated learning architecture: Design, implementation, and challenges in distributed AI systems," *J. Knowl. Learn. Sci. Technol.*, vol. 2, no. 2, pp. 371–384, 2023. ISSN: 2959-6386.

30. T. Ye, M. Luo, Y. Yang, K. K. R. Choo, and D. He, "A survey on redactable blockchain: Challenges and opportunities," *IEEE Trans. Netw. Sci. Eng.*, vol. 10, no. 3, pp. 1669–1683, 2023. doi: 10.1109/TNSE.2023.3235690.
31. Y. Tiwari, "Data privacy challenges and regulatory responses in cross-border cryptocurrency transactions: A comparative analysis," *Educ. Adm. Theory Pract.*, vol. 30, no. 1, pp. 822–832, 2024.
32. J. Verbraeken, M. Wolting, J. Katzy, J. Kloppenburg, T. Verbelen, and J. S. Rellermeyer, "A survey on distributed machine learning," *ACM Comput. Surv.*, vol. 53, no. 2, pp. 1–33, 2020.
33. Z. Wang and Q. Hu, "Blockchain-based federated learning: A comprehensive survey," *arXiv preprint arXiv:2110.02182*, pp. 1–18, 2021.
34. Z. Wang, Q. Hu, R. Li, M. Xu, and Z. Xiong, "Incentive mechanism design for joint resource allocation in blockchain-based federated learning," *IEEE Transactions on Parallel and Distributed Systems*, 34(5), 1536–1547, 2023.
35. J. Wen, Z. Zhang, Y. Lan, Z. Cui, J. Cai, and W. Zhang, "A survey on federated learning: challenges and applications," *Int. J. Mach. Learn. Cybern.*, vol. 14, no. 2, pp. 513–535, 2023. doi: 10.1007/s13042-022-01554-0.
36. Z. Wu, H. Li, Y. Qian, Y. Hua, and H. Gan, "Poison-resilient anomaly detection: Mitigating poisoning attacks in semi-supervised encrypted traffic anomaly detection," *IEEE Trans. Netw. Sci. Eng.*, vol. 11, no. 5, pp. 4744–4757, 2024. doi: 10.1109/TNSE.2024.3382194.
37. S. Xing, Z. Ning, J. Zhou, X. Liao, J. Xu, and W. Zou, "N-fedavg: Novel federated average algorithm based on fedavg," in *Proc. 14th Int. Conf. Commun. Softw. Netw. (ICCSN)*, Jun. 2022, pp. 187–196. doi: 10.1109/ICCSN54991.2022.9877128.
38. Q. Yang, Y. Liu, Y. Cheng, Y. Kang, T. Chen, and H. Yu, "Horizontal federated learning," in *Federated Learning*, Cham, Switzerland: Springer, 2020, pp. 49–67. doi: 10.1007/978-3-030-42266-2_3.
39. Yu, S. Shen, S. Wang, K. Zhang, and H. Zhao, "Communication-efficient hybrid federated learning for e-health with horizontal and vertical data partitioning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 36, pp. 5614–5628, 2024. doi: 10.1109/TNNLS.2024.3383748.
40. K. Zalik and M. Žalik, "A review of federated learning in agriculture," *Sensors (Basel, Switzerland)*, vol. 23, 2023. [Online]. Available: <https://doi.org/10.3390/s23239566>.
41. Y. Zhao et al., "Privacy-preserving blockchain-based federated learning for IoT devices," *IEEE Internet Things J.*, vol. 8, no. 3, pp. 1817–1829, 2020. doi: 10.1109/JIOT.2020.3013860.