future internet

MDPI

Review

Beyond Centralized AI: Blockchain-Enabled Decentralized Learning

Daren Wang [1], Tengfei Ma [1], Juntao Zhu [2] and Haihan Duan [3,*]

[1] Faculty of Engineering, The Chinese University of Hong Kong, Hong Kong 999077, China; darenwang@link.cuhk.edu.hk (D.W.); tengfeima@link.cuhk.edu.hk (T.M.)
[2] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China; jt.zhu1@siat.ac.cn
[3] Artificial Intelligence Research Institute, Shenzhen MSU-BIT University, Shenzhen 518172, China
* Correspondence: duanhaihan@smbu.edu.cn

Abstract

The dominance of centralized artificial intelligence architectures raises significant concerns regarding privacy, data ownership, and control. These limitations have motivated the development of decentralized learning paradigms that aim to remove reliance on a central authority during model training. While federated learning represents an intermediate step by allowing distributed training without raw data exchange, it still depends on a centralized server which could lead to single-point vulnerabilities. Beyond this, a fully decentralized learning in general faces challenges in security vulnerabilities, absence of governance, and lack of incentive alignment. Recent advances in blockchain technology offer a promising foundation for addressing these issues. This paper provides a systematic analysis of blockchain's mechanism-level roles in security, consensus, smart contract, and incentives to support decentralized learning. By reviewing state-of-the-art approaches, this paper suggests that appropriately designed blockchain architectures have the potential to enable practical, secure, and incentive-compatible decentralized learning as technological capabilities continue to evolve.

Keywords: blockchain; decentralized learning; privacy-preserving AI; federated learning; incentive mechanisms; consensus mechanisms; smart contracts

1. Introduction

1.1. Motivation and Background

Centralized artificial intelligence (AI) has increasingly become a source of public concern. A limited number of firms, most notably OpenAI, Anthropic, and Google, control access to the most advanced models, together with the large-scale data and computing resources required to sustain them. This concentration of influence may create vulnerabilities: questions are often raised about privacy protection, copyright compliance, and the ethical implications of concentrated control. These worries are not only theoretical. In September 2025, for example, Anthropic reached a US$1.5 billion settlement after authors alleged that their works had been used without authorization for model training [1].

Public data, once relatively abundant, has already been heavily utilized, leaving limited new material available for training. To develop more advanced models, additional sources of data are required. Yet companies that hold valuable content are increasingly unwilling to make it accessible. In July 2025, for example, Cloudflare announced that it would block AI crawlers by default, effectively allowing its clients to prevent AI firms from scraping their sites [2]. Beyond this, sensitive sectors such as healthcare and finance treat their data

Academic Editor: Wei Yu

Received: 18 January 2026
Revised: 7 February 2026
Accepted: 11 February 2026
Published: 13 February 2026

as highly protected resources and do not release them to AI companies. If these patterns persist, future model training will need to take place under conditions where client data must remain secure and, in many cases, subject to strict safeguards.

1.2. Decentralized Artificial Intelligence

In an ideal scenario, a sustainable decentralized learning ecosystem can function entirely without centralized control. In such a system, numerous data owners or/and computational nodes collaboratively train a shared model while preserving their individual privacy and security. Coordination among parties would emerge through distributed protocols rather than a central authority. This should enable the system to self-organize and operate smoothly even at large scale. The resulting framework would support secure, scalable, and trustworthy model training, ensuring that learning can progress collectively without dependence on any single point of control or failure.

To realize this vision, it is essential to examine the current research on decentralized learning. Federated learning is an important progress introduced by Google in 2016 [3]. In federated learning, clients that own data train models locally using their private data, while only parameter updates are shared with a central server (This paper uses the term client to refer to each node or data owner, as every computational participant is assumed to possess and train on its own local data. Likewise, the term server is used interchangeably with aggregator or coordinator, referring to the entity, whether centralized or decentralized, that performs model aggregation and coordination. We acknowledge that these roles may differ in practical implementations, but this simplification does not affect the conceptual analysis in this work). This preserves privacy and a degree of data ownership. However, federated learning still depends on a centralized aggregator/server to coordinate training and perform model averaging. This introduces the risk of a single point of failure. This paper thus refers to such architectures as **hybrid decentralized learning**, characterized by decentralized training but centralized coordination. This form represents a transitional stage between centralized and fully decentralized paradigms.

Building upon this intermediate model, further decentralization removes the dependence on a single central server. In this paper, **decentralized learning** is defined as the paradigm where both the training process and the server coordination are decentralized. This system eliminates the central authority entirely, allowing all participants to act as peers in both computation and communication. Systems that implement federated learning without a central server can thus be regarded as instances of decentralized learning. Other representative approaches include Gossip Learning [4], where models rather than data are exchanged among peers and gradually converge through random local updates. Another algorithm is Decentralized Parallel Stochastic Gradient Descent (D-PSGD) [5], proposed by Lian et al. (2017), where each client performs local gradient updates and exchanges parameters only with its neighbors. Building on this line of work, Koloskova et al. (2020) [6] proposed a unified theoretical framework that captures various decentralized stochastic gradient methods under dynamic network topologies.

Despite these advances, decentralized learning has not yet achieved large-scale public adoption or enabled the kind of organic ecosystem described above. Several fundamental challenges remain [7,8]. (1) **Security vulnerabilities:** Malicious participants can perform data or model poisoning, inference attacks, or manipulations of aggregation. (2) **Absence of governance:** Without a central coordinator, it is difficult to ensure algorithm rules can be correctly enforced across participants. (3) **Incentive misalignment:** Current frameworks lack economic or reputational mechanisms to encourage honest contributions, leading to potential free-riding behavior.

These limitations suggest that purely algorithmic approaches to decentralization are insufficient. A sustainable decentralized learning ecosystem requires integrated mechanisms for security, governance, and incentive alignment. Such mechanisms are precisely the dimensions in which blockchain technology can offer promising tools.

1.3. Blockchain as the Foundation for Decentralized Learning

Blockchain has emerged as one of the most mature decentralized technologies, with wide adoption in financial applications such as cryptocurrencies, decentralized exchanges, and digital asset management. At its core, a blockchain is a distributed ledger maintained collaboratively by a network of nodes and secured through cryptographic hashing (Node refers to a blockchain participant responsible for maintaining the distributed ledger and participating in consensus. This is distinct from the client discussed in decentralized learning). Several fundamental properties make blockchain a reliable and transparent coordination mechanism [9,10]. (1) Distributed ledger: Records are maintained across multiple nodes with cryptographic linkage, ensuring transparency and tamper resistance. (2) Asymmetric Cryptography: Public-key cryptography ensures identity authentication, data integrity, and confidentiality. (3) Consensus Mechanisms: Algorithms such as PoW and PoS enable collective agreement without centralized control, enhancing robustness and fault tolerance. (4) Smart Contracts: Programmable contracts automate rule enforcement and coordination, enabling trustless interactions. (5) Incentive Mechanisms: Token-based mechanisms motivate honest participation and sustain cooperative behavior.

1.4. Contributions

This paper investigates how blockchain can function as a foundational infrastructure for decentralized learning. We focus on the mechanism-level interplay between blockchain and decentralized learning systems. Table 1 provides a detailed comparison between our work and recent surveys on blockchain-based decentralized learning. The contributions of this paper are as follows.

- We provide a systematic analysis of decentralized learning frameworks, tracing their development from hybrid to fully decentralized architectures, and we review the latest state-of-the-art approaches along with their existing challenges.
- We examine core blockchain mechanisms and evaluate how they contribute to security, governance, and incentive alignment within decentralized learning environments.
- We survey representative blockchain-based decentralized learning systems, analyze their underlying design philosophies, and discuss their advantages, limitations, and potential future research directions.

The remainder of this paper is organized as follows. Section 2 introduces decentralized learning paradigms, architectures, and challenges. Section 3 discusses blockchain as the infrastructure for decentralized learning. Section 4 reviews state-of-the-art blockchain-enabled decentralized learning systems. Sections 5–7 conclude with open challenges and future opportunities.

Table 1. Comparison of this work with existing surveys on Blockchain and Decentralized Learning.

Paper	Primary Focus	AI Architecture	Role of Blockchain	Reference
Nguyen et al. (2021)	Edge Computing	MEC-based FL	Data Integrity	[11]
Beltrán et al. (2023)	DFL Frameworks	Decentralized FL	Tool for Trustworthiness	[8]
Tang et al. (2024)	Crypto & Privacy	FL (Hybrid)	Security & Privacy	[12]

Table 1. *Cont.*

Paper	Primary Focus	AI Architecture	Role of Blockchain	Reference
Yuan et al. (2024)	Network Topology	Decentralized FL	Network challenges	[7]
Javed et al. (2025)	6G Standards	IoT-based FL	IoT Trust & Standardization	[13]
This Paper	**Mechanism (Security, Governance, and Incentives)**	**Hybrid → Fully Decentralized Learning**	**Foundational Infrastructure for Decentralized Learning**	

Note: FL (Federated Learning), MEC (Mobile Edge Computing), DFL (Decentralized Federated Learning).

2. Decentralized Learning: Paradigms, Architectures, and Challenges

2.1. Paradigms, Architectures and State-of-the-Art Algorithms

Broadly speaking, existing distributed learning approaches can be classified into two main categories based on their network topology and dependence on a central server, as illustrated in Figure 1.

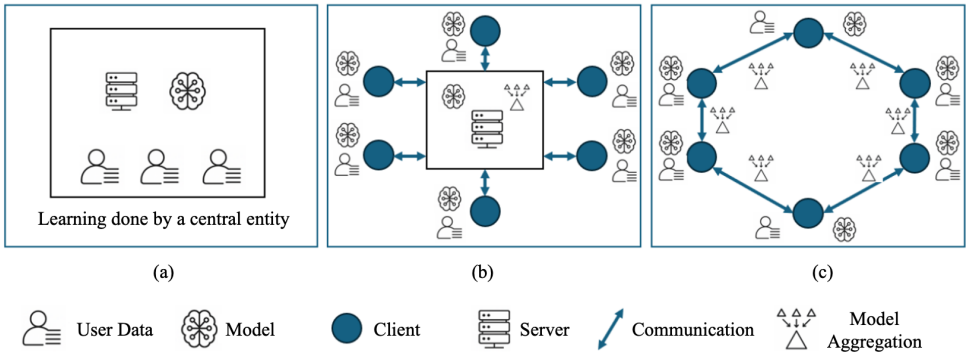

Figure 1. Illustration of (**a**) centralized learning, (**b**) hybrid decentralized learning (federated learning), and (**c**) decentralized learning.

2.1.1. Hybrid Decentralized Learning

In hybrid decentralized learning, or federated learning (FL), model training is decentralized but a centralized server is still required for coordination. Federated learning enables collaborative model training across a large number of edge devices (clients) while keeping raw data localized. Instead of transmitting data to a central server, each client computes updates to a shared global model using its local dataset, and only model parameters or gradients are communicated. The central server then aggregates these updates to refine the global model [3].

Formally, suppose there are K clients where client k holds a local dataset $\mathcal{D}_k$ of size n_k. The local loss of client k is denoted as $\mathcal{F}_k(w)$ and the aggregated loss is $f(w)$, with model parameter w. The global optimization objective is to minimize the weighted sum of all clients' local losses.

$$\min_w f(w) = \sum_{k=1}^{K} \frac{n_k}{n} F_k(w), \quad \text{where } F_k(w) = \frac{1}{n_k} \sum_{i \in D_k} \ell(w; x_i, y_i), \quad n = \sum_{k=1}^{K} n_k \tag{1}$$

Unlike centralized learning, where all data is aggregated in one place to minimize the global loss directly, federated learning performs distributed optimization under communication and privacy constraints. The goal remains to find a single global model weight w that minimizes the overall loss across all clients.

The mainstream optimization algorithm is Federated Averaging (FedAvg) [3]. FedAvg allows each participating client to perform multiple local stochastic gradient descent (SGD) updates before communicating with the central server. At each global round t, the server broadcasts the current global model w_t to a subset of clients S_t. Each client then trains the

model locally for several epochs using its own data and returns the updated parameters w_k^{t+1}. The server aggregates all received models by computing a weighted average:

$$w_{t+1} = \sum_{k \in S_t} \frac{n_k}{m_t} w_k^{t+1}, \quad \text{where } m_t = \sum_{k \in S_t} n_k \tag{2}$$

FedAvg significantly reduces communication frequency by allowing more local computation on each client while maintaining a convergence behavior.

Subsequent research identified two fundamental challenges in federated learning. **(1) Non-IID data distribution:** Client data are often non-independent and non-identically distributed, leading to local model updates that diverge from one another. **(2) System heterogeneity:** Clients differ in computation power, storage capacity, and network conditions. These differences lead to inconsistent training progress and unstable aggregation, which can further degrade model performance.

Together, these challenges cause convergence slowdown and model drift, meaning the global model may fail to align with the true underlying distribution or converge efficiently. To address these challenges, new algorithms have been developed that go beyond simple parameter averaging. Notable state-of-the-art approaches in federated learning include:

- Regularization-based methods (e.g., FedProx [14], FedDyn [15]), which add proximal or dynamic correction terms to stabilize client updates.
- Variance correction methods (e.g., SCAFFOLD [16]), which correct client drift by sharing control variates between the server and clients.
- Normalization decoupling methods (e.g., FedBN [17]), which separate batch normalization parameters locally to accommodate statistical heterogeneity.

These algorithms represent a shift toward principled coordination mechanisms that preserve convergence speed and generalization ability in heterogeneous, real-world environments [14–17]. See Table 2 for a detailed comparison of SOTA federated learning algorithms.

Table 2. Comparison of Representative Federated Learning Algorithms.

Algorithm	Mechanism	Category	Remarks/Limitations	Reference
FedAvg	Weighted averaging of local models after multiple local SGD steps	Baseline	Sensitive to Non-IID data and system heterogeneity; slow convergence	[3]
FedProx	Adds a proximal term to stabilize local updates	Regularization	Balances local–global divergence in Non-IID settings	[14]
FedDyn	Introduces dynamic regularization to align local objectives	Regularization	Accelerates convergence	[15]
SCAFFOLD	Variance reduction via control variates	Variance Correction	More communication overhead	[16]
FedBN	Keeps BatchNorm layers local to handle statistical heterogeneity	Normalization Decoupling	Allow client to inference with non-iid data; supports personalized models	[17]

Note: Other methods are not included but follow similar principles of improving convergence and robustness under heterogeneous conditions.

Federated learning has been adopted in various privacy-sensitive industries, including mobile applications (e.g., Google Gboard), healthcare, and finance [18]. These industries tend to have strict regulatory requirements related to centralized data sharing. To facilitate practical deployment on resource-constrained edge devices, recent advancements have further optimized the framework by introducing edge-assisted decomposition for energy efficiency [19,20] and semantics-guided mechanisms to handle complex data distribu-

tions [21]. However, traditional federated learning assumes the central server is trusted and there is no single point of failure. To achieve decentralized learning, even the centralized servers need to be eliminated or decentralized.

2.1.2. Decentralized Learning

In a Decentralized Learning paradigm, both the training process and the coordination mechanism are distributed across participating nodes, eliminating the need for a central server.

Decentralized federated learning (DFL) is a specific form of federated learning. Instead of sending model updates to a central aggregator, DFL relies on peer-to-peer communication to disseminate information across clients. This communication is typically implemented through gossip-based algorithms [22], which have become the state-of-the-art approach for parameter synchronization in fully decentralized settings.

Gossip Learning, first introduced by Hegedűs et al. (2019) [23], represents a fully decentralized alternative to traditional federated learning. In gossip learning, client k maintains a local dataset $\mathcal{D}_k$, a local model w_k, and periodically exchanges it with randomly selected peers, allowing models to perform random walks across the network. Upon receiving a remote model w_r with training age t_r, the node performs an age-weighted merge defined as

$$w_k \leftarrow (1-a)\, w_k + a\, w_r, \quad a = \frac{t_r}{t_k + t_r} \tag{3}$$

where t_k and t_r denote the cumulative number of samples (or updates) each model has been trained on. The age parameter thus reflects the maturity of a model, giving higher influence to models trained on more data. Repeated merging operations of this form enable the system to converge toward a globally consistent solution, even without any central coordination.

Another SOTA method is the Decentralized Parallel Stochastic Gradient Descent (D-PSGD) algorithm proposed by Lian et al. [5]. It aims to address the fundamental bottleneck of communication congestion in centralized training, enabling large-scale distributed optimization without a central parameter server. In this setting, the system is organized as a network topology, where each client is connected to only a few neighboring clients for parameter exchange. The update rule at round $b+1$ can be summarized as:

$$x_i^{(b+1)} = \left(\sum_{j \in \mathcal{N}(i)} W_{ij} x_j^{(b)} \right) - \gamma \nabla F_i \left(x_i^{(b)} \right) \tag{4}$$

where $\mathcal{F}_i(\cdot)$ denotes the local loss function based on node i's data, and W_{ij} represents the communication weight between node i and node j, determined by the underlying network connectivity. This simple rule combines information diffusion (through neighbor averaging) and local optimization (via gradient descent) in a single step, allowing all nodes to gradually reach consensus without centralized coordination. As a result, D-PSGD significantly reduces communication cost, making it a highly efficient and scalable approach for decentralized learning.

Building upon D-PSGD, subsequent state-of-the-art decentralized learning algorithms aim to further enhance convergence efficiency. In particular, GT-DSGD introduces the idea of gradient tracking, where each node not only updates its own parameters but also maintains and iteratively refines an estimate of the global gradient [24]. In addition, DeTAG implements this insight with accelerated gossip and gradient tracking, achieving near-optimal convergence rates under practical network constraints [25].

Current state-of-the-art decentralized learning algorithms are primarily represented by Decentralized Federated Learning [22], Gossip Learning [4], and D-PSGD [5] along with

their variants. These methods focus mainly on the theoretical foundations of achieving efficient model training without a central server, particularly under complex client network topologies. Essentially, they rely on peer-to-peer parameter exchange among clients, which gradually leads to a global consensus model. A conceptual comparison highlighting the key differences between these paradigms is presented in Table 3.

Table 3. Conceptual comparison between Federated Learning, Gossip Learning, and D-PSGD.

Aspect	**Federated Learning [3]**	**Gossip Learning [4]**	**D-PSGD [5]**
Control	Centralized server	Fully decentralized	Fully decentralized
Communication topology	Star-shaped (client → server)	Random peer-to-peer (client ↔ client)	Fixed structured network (e.g., ring, mesh, or graph-based)
Synchronization	Typically synchronous rounds	Asynchronous updates	Synchronous parallel updates across all nodes
Aggregation	Global averaging (FedAvg)	Local weighted averaging	Neighbor-weighted averaging followed by local gradient descent
Fault tolerance	Relies on central server	High tolerance, resilient to node failure	Moderate; depends on network connectivity

2.2. Challenges in Decentralized Learning

The theoretical foundations of current hybrid and decentralized learning have not been widely adopted in real-world systems. Several critical challenges remain.

2.2.1. Security Vulnerabilities

Security is considered as the most significant challenge. Most decentralized and federated learning frameworks assume that all clients and servers are honest participants. However, in large-scale practical deployments, this assumption rarely holds. In decentralized networks, where numerous participants communicate with one another, various types of adversarial attacks can emerge.

In federated learning, where a central server still exists, a compromised or malicious server may perform model inversion, backdoor aggregation, or intentionally manipulate model parameters.

In fully decentralized learning, although the absence of a central server eliminates the risk of single-point failure, vulnerabilities can also arise from malicious clients. Representative attacks include the following [8,9]:

- Byzantine attacks: Malicious clients inject corrupted gradients or adversarial updates (e.g., random noise, targeted attacks, label flipping, or model poisoning) to degrade global model performance.
- Sybil attacks: Adversaries create a large number of fake clients to bias or dominate the global aggregation process.
- Evasion attacks: Corrupted clients manipulate local training data or model updates to cause the global model to misclassify or ignore specific adversarial inputs during inference.

Most existing defenses focus on robust aggregation. State-of-the-art Byzantine-resilient aggregation algorithms, such as Krum [26], Mean-Median (MeanMed) [27,28], and Bulyan [29], aim to detect and down-weight (or exclude) parameter updates that deviate significantly from the consensus [30]. More recent approaches, including RFLPA [31] and FoundationFL [32], enhance robustness by refining existing aggregation rules rather than designing entirely new ones. While effective in certain settings, these methods still struggle against small-magnitude, directionally consistent poisoning attacks and often introduce significant computational complexity.

Although decentralized learning keeps raw data local, shared model parameters can still leak sensitive information through inference attacks, such as membership or property inference. To mitigate such risks, differential privacy (DP) techniques are commonly applied by adding noise to model updates or aggregated results [33]. However, while DP enhances privacy guarantees, it inevitably introduces a trade-off between privacy, model utility, and convergence speed.

2.2.2. Absence of Governance and Trust Management

In decentralized learning systems, the absence of a central coordinator removes a unified mechanism for accountability among participants. Each participant must decide independently whether to trust others and their shared models. Without a trusted entity to validate model updates, enforce rules, or resolve conflicts, coordination and knowledge exchange become difficult.

Recent research [34] suggests that regulation and compliance could help address governance. However, enforcing such regulations in a complex network environment is costly and difficult to manage. More importantly, introducing regulatory oversight inevitably increases centralization, which conflicts with the fundamental objective of decentralized learning.

2.2.3. Lack of Incentive Alignment

Current decentralized learning frameworks often lack robust economic or reputational incentives to sustain honest participation. As a result, rational clients may engage in free-riding behavior, benefiting from the global model without contributing meaningful local updates, which undermines collaboration and slows convergence. Recent studies address this issue primarily through contribution-based, game-theoretic, and reputation-driven mechanisms to motivate active and reliable participation while preserving privacy [35]. However, most of these methods remain theoretical frameworks without deployment in practical systems. These approaches still face critical challenges in fairly evaluating data value and ensuring efficiency and security in large-scale, dynamic environments.

Overall, these challenges should be addressed for decentralized learning to function effectively in practice. As a decentralized technology, blockchain could offer potential solutions.

3. Blockchain as the Infrastructure for Decentralized Learning

Decentralized learning fundamentally struggles with three unresolved challenges: security vulnerabilities, absence of governance, and lack of incentive alignment. Importantly, these limitations do not stem from the learning algorithms themselves, but from the lack of a reliable trust and coordination infrastructure that enables participants to cooperate without central authority.

Blockchain is a distributed system that allows multiple parties to maintain a shared and tamper-resistant record without relying on a central authority. It could provide the missing foundation for decentralized learning through the following properties. See Table 4 that summarizes how foundational blockchain mechanisms directly address core limitations of decentralized learning.

3.1. Distributed Ledger

Blockchain maintains a shared, append-only ledger replicated across all nodes in the network [36–38]. Each recorded event is timestamped and cryptographically chained to previous entries, making the historical record immutable. In the context of decentralized learning, this means the full evolution of the global model including model update and aggregation, becomes transparent and auditable. Any participant can verify which contributions influenced the model and evaluate their legitimacy. To support this at scale, advanced

architectures have been proposed to integrate distributed databases, thereby enhancing auditability and data reliability [39]. The distributed ledger eliminates the reliance on implicit trust between clients and addresses security vulnerabilities from malicious participants.

Table 4. Blockchain capabilities in addressing decentralized learning challenges.

Challenge in Decentralized Learning	Blockchain Property	Resulting Capability
Security vulnerabilities	Distributed ledger + cryptographic authentication	Verifiable and tamper-resistant model provenance
Absence of governance	Consensus protocols + smart contracts	Transparent, rule-enforced coordination without central authority
Lack of incentive alignment	Token and reputation mechanisms	Stable long-term collaboration based on economic rationality

3.2. Asymmetric Cryptography

Blockchain uses asymmetric cryptography, where each participant holds a public–private key pair. This serves as the basic identity and authentication mechanism: participants use their private keys to sign their actions, and others can verify these signatures using the corresponding public keys [40,41]. Each update can be traced back to a specific contributor, preventing impersonation or forged submissions.

At the same time, cryptographic techniques allow the system to record proofs of contribution without revealing any underlying data, helping preserve the privacy goals of decentralized or federated learning. Overall, asymmetric cryptography provides a secure and accountable identity layer, which is essential for open, permissionless collaboration.

3.3. Consensus Mechanisms

The global state of the ledger is maintained through consensus protocols such as Proof-of-Work, Proof-of-Stake [42–44]. Instead of trusting a central aggregator, participants collectively validate and agree upon which updates are accepted. This ensures that the global model state cannot be manipulated by any single party. Consensus therefore provides the governance capability that decentralized learning lacks: the system can coordinate legitimately even when participants do not trust each other.

3.4. Smart Contracts

Smart contracts extend blockchain beyond static record-keeping by enabling automated execution of predefined conditions [45,46]. These contracts can specify admissibility criteria for model updates, define verification procedures for accuracy improvement, manage access control to shared models, or enforce penalties for malicious behavior. Importantly, smart contract execution is deterministic and requires no trusted enforcement authority. Governance therefore becomes protocol-driven rather than institution-driven, enabling reliable coordination without central oversight.

3.5. Incentive Mechanisms

Blockchain ecosystems natively support token-based or reputation-based incentive structures that reward constructive participation and penalize harmful behavior [42,47,48]. In decentralized learning, these mechanisms ensure that clients contributing legit model updates receive quantifiable compensation, while free-riding or poisoning attacks become economically unfavorable. This resolves the incentive misalignment problem that commonly undermines collaborative learning, allowing the system to sustain participation over time. Recent frameworks have further instantiated this through relay-based sustainable learning models [49] and win-win incentive designs for content creators [50].

Blockchain does not modify the optimization dynamics of learning algorithms. It provides the institutional infrastructure needed for decentralized learning.

4. Blockchain-Enabled Decentralized Learning: Current State-of-the-Art Systems

Recent studies have examined the integration of blockchain technology into decentralized learning and have proposed a range of conceptual frameworks. In this paper, we categorize the state-of-the-art algorithms into two broad groups: those that primarily address security and governance, and those that focus on incentive mechanisms.

4.1. Blockchain-Enabled Security and Governance

One notable approach to blockchain-enabled decentralized learning is Swarm Learning (SL), proposed by Warnat-Herresthal et al. in 2021 [51]. SL can be regarded as an architectural extension of federated learning, where the centralized server is replaced by a blockchain-based leader selection and coordination mechanism. Multiple clients train models locally on their private datasets and periodically share model parameters. A smart contract on the blockchain dynamically elects the first client that signals readiness as a temporary leader, who then aggregates parameters and distributes the updated model to the rest of the network. The blockchain functions as a permissioned distributed ledger, collaboratively maintained by all participating clients, recording leader elections, merge rules, and model states.

This design leverages blockchain's immutability, transparency, and consensus properties to eliminate the need for a trusted central coordinator and ensure a verifiable training history. In essence, blockchain achieves governance and trust without a central authority. The original work targeted medical applications, using multi-institutional transcriptomic and imaging datasets, and achieved performance comparable to centralized models when the number of clients ranged from 3 to 32.

Nevertheless, the framework retains several limitations. It assumes partially honest clients: the leader selection follows a first-come-first-serve policy, allowing a malicious client to falsely claim readiness and be chosen as the leader, which introduces a potential single point of failure. The system also lacks mechanisms to verify model updates, making it vulnerable to Byzantine or free-rider behaviors in the absence of incentives. While an immutable ledger enables traceability within small, regulated clinic networks (as demonstrated in the 32-node experiments), scalability remains a challenge for larger, heterogeneous environments. Overall, Swarm Learning represents a state-of-the-art approach for small-scale decentralized networks, providing a strong foundation for future research on secure and verifiable collaborative learning.

Blockchain-enabled Incentivized and Secure Federated Learning (BIT-FL) is a blockchain-based decentralized federated learning framework that integrates incentive compatibility, differential privacy (DP), and Byzantine-resilient consensus into a unified architecture [52]. In each training round, clients submit encrypted bids representing their computational costs to a smart contract. Using an exponential mechanism under differential privacy, the contract selects a subset of participants as trainers or validators, concealing individual cost information. Trainers locally update the global model using private datasets and inject Gaussian noise into their gradients to ensure privacy preservation. Validators then evaluate these local models on a public validation dataset to determine their contribution to the global objective.

The final decision is achieved through validator voting, and only models that improve validation performance are accepted for global aggregation and rewarded according to their reported cost. Trainers whose models fail to contribute positively are removed, while validators whose votes deviate from the majority are expelled. BIT-FL achieves robustness against

Byzantine and inference attacks under the assumption that a majority (50%) of trainers and validators behave honestly. The validation mechanism effectively deters free-riders and enforces transparent rules through smart contracts and consensus voting mechanism.

However, one unresolved issue is that, since the public validation dataset is accessible to all participants, malicious trainers may overfit their models to this dataset to secure higher validation scores and rewards. Therefore, despite its strong conceptual design, BIT-FL remains primarily a theoretical framework.

Blockchain-Enabled Secure and Privacy-Preserving Decentralized Learning System (SPDL) proposes a flexible decentralized framework that can be extended to other decentralized learning algorithms not limited to federated learning [53]. In SPDL, each client initializes a unique cryptographic identity and performs local training on private data. A verifiable random function randomly selects a temporary leader for each training round. Before communication, clients add Gaussian noise to their local gradients to guarantee differential privacy (DP). The noisy gradients are aggregated using a Byzantine-resilient rule such as Krum, which filters out anomalous updates and preserves convergence under the presence of malicious clients.

The aggregation results are validated through an on-chain consensus protocol, where all clients collaboratively verify and append the new model update. Upon receiving a proposed model, each client first verifies the leader's digital signature using its registered public key to ensure that the message is authentic and tamper-free. Then, it independently recomputes the aggregated gradient and checks whether the received parameters deviate beyond a preset threshold. This distance-based validation involves only lightweight vector arithmetic and is negligible compared to local gradient computation. Furthermore, SPDL employs a reputation mechanism that penalizes clients submitting abnormal gradients and excludes them from future leadership. Through the seamless integration of DP, Byzantine robustness, cryptographic verification, and blockchain consensus, SPDL achieves secure, verifiable, and privacy-preserving decentralized learning.

However, its security is relatively weaker compared to BIT-FL since the system requires at least two-thirds of the nodes to behave honestly for consensus to hold. Another issue is that no incentive mechanism is incorporated to discourage malicious or non-cooperative behaviors.

There are several other notable blockchain-based federated learning frameworks, including Fantastyc, BlockFLA, PBFL [54–56], etc. In essence, these systems are variants of decentralized federated learning that leverage blockchain to achieve secure coordination and trust management among distributed participants. Blockchain provides an immutable and transparent infrastructure that enables flexible rule enforcement, auditability, and integration of various privacy-preserving mechanisms such as differential privacy, homomorphic encryption, and secure aggregation. Different voting and consensus schemes have also been proposed to enhance Byzantine fault tolerance and ensure fair model aggregation.

Overall, while these frameworks demonstrate significant progress in decentralizing learning, most of them still lack a comprehensive incentive design. The absence of robust economic mechanisms limits their ability to sustain long-term participant engagement.

4.2. Blockchain-Enabled Incentive Mechanisms

Research on incentive mechanisms for decentralized learning remains relatively nascent, with only a limited number of studies exploring how blockchain-based incentive architectures can be effectively integrated into decentralized learning systems. Within this emerging frontier, existing frameworks primarily focus on two directions:

4.2.1. Incentive Based on Clients' Contribution

Incentive framework design in blockchain-enabled decentralized learning primarily relies on blockchain-based consensus mechanisms to distribute rewards according to each participant's effective contribution to global model improvement.

FedToken [57] adopts a Shapley-value-based formulation to precisely measure each client's marginal contribution to the global model. Formally, each client's contribution is defined as

$$u_n = \text{Average}[V(\text{Model with client } n) - V(\text{Model without client } n)], \tag{5}$$

where $V(\cdot)$ denotes the global model's performance metric (e.g., accuracy or loss). This follows the cooperative game-theoretic principle that each participant's fair reward corresponds to its average marginal improvement across all possible participation orders. Although computationally intensive, this approach ensures high fairness and accurate contribution quantification.

To improve scalability, Wu and Seneviratne propose a gradient-alignment-based approximation in their BFSIF framework [58]. The algorithm evaluates each client's update through directional similarity with the global optimization trajectory:

$$S_i = g_i \cdot g_{\text{global}} \cdot \frac{n_i}{N} \tag{6}$$

where g_i represents the local gradient from client i, g_{global} is the aggregated global gradient, n_i is the number of local samples, and N is the total sample size. This alignment-based score S_i captures both the consistency and data-weighted significance of each contribution while significantly reducing computational overhead.

Together, these frameworks illustrate two major design paradigms in using block chain to align incentives: Shapley-based fairness with higher computational cost versus alignment-based efficiency with approximate fairness. Both leverage blockchain as a trustless settlement and verification layer, ensuring transparent, auditable, and contribution-aware reward allocation that sustains long-term participation.

4.2.2. Blockchain-Enabled Incentive Platform for Decentralized Learning

Recent studies also explored integrating blockchain consensus mechanisms to construct large-scale decentralized learning platforms. FedChain and POFL [59,60] exemplifies this concept by envisioning a blockchain-based ecosystem where model publishers can openly post learning tasks, such as semantic analysis or biomedical image recognition, along with target accuracy and token incentives. Clients possessing relevant data may then participate as miners, using their local computational resources and data to train models and earn rewards once their contributions are verified. In such frameworks, the blockchain functions as an essential technological layer that ensures transparency, immutability, decentralized coordination, and verifiable incentive distribution among untrusted participants. However, the proposed platform remains theoretical, as practical implementation would require further advances in scalability, secure model verification, and incentive mechanism design.

From a theoretical perspective, blockchain can serve as a fundamental infrastructure for decentralized learning. It enhances system security, enables management and trust without a central authority, and supports the design of incentive mechanisms. The inherent properties of blockchain directly address core challenges in decentralized learning, as outlined in Table 5.

- Distributed Ledger ensures that all clients share access to a consistent and verifiable record.
- Asymmetric Cryptography enables identity authentication and role-based access for each participant.

- Consensus Mechanisms support decentralized decision-making, such as selecting the optimal global model and excluding malicious nodes.
- Smart Contracts enforce system rules transparently. Their flexibility allows different decentralized learning systems to adopt customized architectural designs. Moreover, blockchain can be seamlessly integrated with other privacy-preserving technologies, such as differential privacy.
- Incentive Mechanisms offer transparent and programmable structures that encourage clients to participate and contribute to the learning process over the long term.

Despite these theoretical advantages, most existing studies remain at the framework design stage. Significant challenges and open research directions need to be discussed.

Table 5. Comparison of Representative Blockchain-Enabled Decentralized Learning Frameworks.

Framework	Paradigm/ Design Focus	Use of Blockchain	Challenges Addressed	Limitations/ Comments	Reference
Swarm Learning (SL)	DFL; Blockchain-based leader selection	Distributed ledger; Smart contracts for leader election and model update recording	Governance and trust management without a central server	Vulnerable to malicious leader selection; lacks Byzantine robustness and incentives; scalability limited to small, semi-trusted networks	[51]
BIT-FL	DFL; Blockchain-enabled validator voting and incentive compatibility	Smart contract for role assignment and rewards; Consensus voting	Security vulnerability, free-rider behavior, inference attack resilience	Public validation dataset may be overfitted; assumes $\geq$50% honest participants	[52]
SPDL	Secure and privacy-preserving decentralized learning with cryptographic validation	Distributed ledger; digital signature verification; reputation tracking	Byzantine robustness; governance and decentralized verification	Requires $\geq$2/3 honest nodes; lacks incentive mechanism for sustained participation	[53]
FedToken	Contribution-aware FL reward distribution via cooperative game theory	Blockchain as transparent settlement and audit layer	Fair reward allocation; long-term participation sustainability	Shapley-value computation is expensive; mainly suitable for small to medium-scale networks	[57]
BFSIF	Gradient-alignment contribution scoring for scalable incentive assignment	Blockchain records contribution scores and payment transactions	Reduces computational overhead while maintaining fairness approximation	Approximate fairness; sensitive to noisy gradients; relies on robust aggregation	[58]
FedChain/ POFL	Blockchain-based open collaborative learning marketplace	Blockchain as coordination infrastructure for task posting, model training, and token distribution	Large-scale decentralized participation; transparent task–reward matching	Mostly conceptual; requires secure model verification and high scalability; not yet practically deployable	[59,60]

5. Challenges and Open Problems

An examination of the current state-of-the-art landscape in blockchain-enabled decentralized learning algorithms reveals several unresolved challenges and problems.

5.1. System Design Vulnerabilities

Blockchain-enabled decentralized learning systems rely heavily on the underlying protocol and system design, including rules for training participation, leader election, model aggregation, and result verification. Our analysis indicates that several state-of-the-art frameworks still expose design-level vulnerabilities.

For instance, in Swarm Learning, the leader is selected solely based on the clients' self-reported readiness, where the first participant claiming readiness becomes the leader. Such a mechanism is susceptible to manipulated or falsified readiness signals, enabling malicious nodes to seize leadership [51].

Critically, current incentive designs often rely on simplified honest-majority assumptions (e.g., BIT-FL requires >50% honest validators [52], while SPDL requires >2/3 [53]), which are fragile against colluding rational actors. Beyond purely destructive attacks, systems are vulnerable to strategic manipulation where rational actors exploit protocol rules for profit. A prime example is the vulnerability in BIT-FL, where the reliance on a publicly accessible validation dataset incentivizes trainers to intentionally overfit their local models to this specific validation set. This allows adversaries to maximize their reward scores ("gaming the system") while degrading the global model's generalization capability. Furthermore, free-riding remains a significant economic threat; while frameworks like FedToken attempt to mitigate this via temporal discounting [57], detecting sophisticated "lazy" updates that statistically mimic valid gradients remains an open challenge.

In scenarios where such strategic manipulations become pervasive or the honest-majority assumption is irrevocably violated, algorithmic defenses alone may fail. A potential governance-oriented solution to such systemic failures lies in adopting a fork mechanism as studied in blockchain governance research [61]. A hard fork is a non-backward-compatible protocol change that permanently splits the blockchain into two distinct networks. In practice, it often serves as a governance mechanism that enables the community to transition to a new and improved chain when consensus on the existing rules can no longer be maintained. In decentralized learning systems, this implies the possibility of reinitializing the training ecosystem with revised governance rules when the existing protocol becomes unstable or compromised.

5.2. Smart Contract Risk

Smart contracts play a foundational role in blockchain-enabled decentralized learning systems. Nevertheless, any vulnerability or programming error in the smart contract code may compromise the system's integrity. This makes the entire system open up to attacks, and undermines governance [9,10]. On the mitigating side, smart contract source code is publicly accessible, which enables community-wide review and audit. Open auditing also increases the chance of detecting flaws before deployment and thereby strengthens trust in the protocol.

5.3. Computation Burden and Communication Overhead

Decentralized learning systems inherently incur higher computation burdens and communication overheads compared to centralized learning, largely due to the complexity of peer-to-peer networking. The integration of a blockchain layer further increases this burden, as all participants are required to maintain and synchronize a distributed ledger.

Quantitative analysis reveals significant bottlenecks in real-world deployment. Regarding latency, empirical results from BIT-FL indicate that adding layers of standard consensus mechanisms, such as Practical Byzantine Fault Tolerance (PBFT), can impose prohibitive delays exceeding 12,000 s. While optimized sharded consensus mechanisms can reduce this to approximately 52 s, the latency overhead remains tangible compared to non-blockchain approaches [52]. On-chain aggregation is also economically constrained. Wu et al. demonstrated that smart contract gas costs grow exponentially with model size, reaching USD 4.7 billion in gas for a 100,000-parameter model, rendering full on-chain updates economically impractical [58]. Regarding scalability, while frameworks like Swarm Learning have proven effective, their empirical validation has been primarily limited to smaller networks (≈32 nodes) [51]. Achieving scalability for thousands of participants (as simulated in BIT-FL with 4,000 validators) has to adopt complex sharding mechanisms that introduce their own trade-offs in terms of system complexity and cross-shard latency [52].

Current studies mitigate this overhead by restricting on-chain storage to lightweight information (such as verification proofs or participant identification) while keeping heavy data (e.g., full model parameters) off-chain and local [7,8]. Nonetheless, blockchain-enabled decentralized learning continues to demand increased computation and communication resources. This tradeoff appears to be unavoidable if decentralized security and governance guarantees are to be preserved.

6. Future Research Opportunities

Recent research mainly studies blockchain's role in the training phase of decentralized learning. As the infrastructure for verification and incentive mechanisms, blockchain may also support other components in the decentralized AI ecosystem. In particular, future research could explore blockchain's integration to decentralized data collection and model inference.

6.1. Decentralized Data Collection and Incentivization

Most existing blockchain-enabled decentralized learning frameworks implicitly assume that each participant holds local private data and contributes to model training. In practice, data may originate from distinct entities that act as data providers rather than model trainers. This leads to a need for mechanisms that validate data quality, ensure provenance, and incentivize high-value data contribution. Blockchain can facilitate such an ecosystem by offering verifiable records of data contribution events and enabling transparent incentive mechanisms [47]. Through cryptographic verification techniques, data quality assessment can be performed without exposing raw data, thereby preserving privacy while promoting trustworthy decentralized data exchanges.

6.2. Verifiable and Incentivized Model Inference

Existing approaches generally assume that the final trained model is locally available to all participants. However, when model size increases or storage and computation become unevenly distributed, model inference may require querying model hosts across the network. This setting introduces the risk that a host may return an altered or degraded model to reduce computation cost. Future research may explore blockchain-based verifiable inference protocols leveraging Zero-Knowledge Proofs (ZKP) or Trusted Execution Environments (TEE), where the correctness of inference outputs or model parameters can be cryptographically checked without full model disclosure. Additionally, incentive structures can be incorporated to encourage honest model hosting, ensuring that model users receive the accurate model version agreed upon by the network. Beyond inference, blockchain-based frameworks can be extended to broader model applications such as decentralized Multi-Agent Systems (MAS), where they ensure secure coordination and scalable collaboration among autonomous entities [62].

7. Conclusions

This paper examined blockchain as a foundational infrastructure for decentralized learning. Both hybrid decentralized architecture and fully decentralized systems face persistent challenges, including security vulnerabilities, absence of governance, and lack of incentive alignment. These limitations indicate that algorithmic decentralization alone is inadequate to sustain reliable learning.

Blockchain provides structural capabilities that align with the core requirements of decentralized learning. Distributed ledger offers transparent and secure record keeping. Asymmetric cryptography ensures identity authentication. Consensus mechanisms support decentralized coordination. Smart contracts enable verifiable rule enforcement. Incentive structures promote honest contribution and sustained participation.

Despite these advantages, several open problems remain. System design vulnerabilities may arise in leader selection, aggregation, and reward distribution. Smart contract failures may compromise system-level security. Communication and computation overhead can also be significant with blockchain integration.

Future research may extend blockchain involvement beyond model training to decentralized data acquisition and verifiable inference. Potential directions include provenance-aware data contribution mechanisms, privacy-preserving data quality verification, and incentive-aligned model hosting. Advancing these areas could support a more transparent, secure, and collectively governed AI ecosystem in which learning proceeds without reliance on centralized authorities.

Author Contributions: Conceptualization and methodology, D.W. and T.M.; validation, J.Z. and H.D.; formal analysis, investigation, resources, and data curation, D.W. and T.M.; writing—original draft preparation, D.W. and T.M.; writing—review and editing, D.W. and T.M.; visualization, D.W. and T.M.; supervision, J.Z. and H.D. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Key Field Projects of Ordinary Universities in Guangdong Province (No. 2025ZDZX3050).

Data Availability Statement: No new data were created or analyzed in this study.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. The Guardian. AI Startup Anthropic Agrees to Pay $1.5bn to Settle Book Piracy Lawsuit. *The Guardian*, 5 September 2025. Available online: https://www.theguardian.com/technology/2025/sep/05/anthropic-settlement-ai-book-lawsuit (accessed on 5 September 2025).
2. Cloudflare, Inc. Cloudflare Just Changed How AI Crawlers Scrape the Internet-at-Large; Permission-Based Approach Makes Way for a New Business Model. Press Release, San Francisco, CA, 1 July 2025. Available online: https://www.cloudflare.com/en-au/press/press-releases/2025/cloudflare-just-changed-how-ai-crawlers-scrape-the-internet-at-large/ (accessed on 1 July 2025).
3. McMahan, B.; Moore, E.; Ramage, D.; Hampson, S.; y Arcas, B.A. Communication-efficient learning of deep networks from decentralized data. In Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS), Fort Lauderdale, FL, USA, 20–22 April 2017; pp. 1273–1282.
4. Ormándi, R.; Hegedűs, I.; Jelasity, M. Gossip learning with linear models on fully distributed data. *Concurr. Comput. Pract. Exp.* **2013**, *25*, 556–571. [CrossRef]
5. Lian, X.; Zhang, C.; Zhang, H.; Hsieh, C.J.; Zhang, W.; Liu, J. Can decentralized algorithms outperform centralized algorithms? A case study for decentralized parallel stochastic gradient descent. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 5336–5346.
6. Koloskova, A.; Loizou, N.; Boreiri, S.; Jaggi, M.; Stich, S. A unified theory of decentralized SGD with changing topology and local updates. In Proceedings of the 37th International Conference on Machine Learning (ICML), Virtual, 13–18 July 2020; pp. 5381–5393.
7. Yuan, L.; Wang, Z.; Sun, L.; Yu, P.S.; Brinton, C.G. Decentralized federated learning: A survey and perspective. *IEEE Internet Things J.* **2024**, *11*, 34617–34638. [CrossRef]
8. Beltrán, E.T.M.; Pérez, M.Q.; Sánchez, P.M.S.; Bernal, S.L.; Bovet, G.; Pérez, M.G.; Pérez, G.M.; Celdrán, A.H. Decentralized federated learning: Fundamentals, state of the art, frameworks, trends, and challenges. *IEEE Commun. Surv. Tutor.* **2023**, *25*, 2983–3013. [CrossRef]
9. Singh, A.; Parizi, R.M.; Zhang, Q.; Choo, K.K.R.; Dehghantanha, A. Blockchain smart contracts formalization: Approaches and challenges to address vulnerabilities. *Comput. Secur.* **2020**, *88*, 101654. [CrossRef]
10. Jiao, T.; Xu, Z.; Qi, M.; Wen, S.; Xiang, Y.; Nan, G. A survey of ethereum smart contract security: Attacks and detection. *Distrib. Ledger Technol. Res. Pract.* **2024**, *3*, 1–28. [CrossRef]
11. Nguyen, D.C.; Ding, M.; Pham, Q.V.; Pathirana, P.N.; Le, L.B.; Seneviratne, A.; Li, J.; Niyato, D.; Poor, H.V. Federated learning meets blockchain in edge computing: Opportunities and challenges. *IEEE Internet Things J.* **2021**, *8*, 12806–12825. [CrossRef]
12. Tang, Y.; Zhang, Y.; Niu, T.; Li, Z.; Zhang, Z.; Chen, H.; Zhang, L. A Survey on Blockchain-Based Federated Learning: Categorization, Application and Analysis. *Comput. Model. Eng. Sci. (CMES)* **2024**, *139*, 2451–2477. [CrossRef]

13. Javed, F.; Zeydan, E.; Mangues-Bafalluy, J.; Dev, K.; Blanco, L. Blockchain for Federated Learning in the Internet of Things: Trustworthy Adaptation, Standards, and the Road Ahead. *arXiv* **2025**, arXiv:2503.23823. [CrossRef]
14. Li, T.; Sahu, A.K.; Zaheer, M.; Sanjabi, M.; Talwalkar, A.; Smith, V. Federated optimization in heterogeneous networks. *Proc. Mach. Learn. Syst.* **2020**, *2*, 429–450.
15. Durmus, A.E.; Yue, Z.; Ramon, M.; Matthew, M.; Paul, W.; Venkatesh, S. Federated learning based on dynamic regularization. In Proceedings of the International Conference on Learning Representations, Virtual, 3–7 May 2021.
16. Karimireddy, S.P.; Kale, S.; Mohri, M.; Reddi, S.; Stich, S.; Suresh, A.T. Scaffold: Stochastic controlled averaging for federated learning. In Proceedings of the 37th International Conference on Machine Learning (ICML), Virtual, 13–18 July 2020; pp. 5132–5143.
17. Li, X.; Jiang, M.; Zhang, X.; Kamp, M.; Dou, Q. Fedbn: Federated learning on non-iid features via local batch normalization. *arXiv* **2021**, arXiv:2102.07623.
18. Hard, A.; Rao, K.; Mathews, R.; Ramaswamy, S.; Beaufays, F.; Augenstein, S.; Eichner, H.; Kiddon, C.; Ramage, D. Federated learning for mobile keyboard prediction. *arXiv* **2018**, arXiv:1811.03604.
19. Shi, Y.; Duan, H.; Chi, Y.; Gai, K.; Cai, W. Edge-assisted federated learning: An empirical study from software decomposition perspective. In *Proceedings of the 20th International Conference on Algorithms and Architectures for Parallel Processing (ICA3PP), New York, USA, 2–4 October 2020*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 200–214.
20. Shi, Y.; Duan, H.; Yang, L.; Cai, W. An energy-efficient and privacy-aware decomposition framework for edge-assisted federated learning. *ACM Trans. Sens. Netw.* **2022**, *18*, 1–24. [CrossRef]
21. Zhang, J.; Xu, Y.; Li, S.; Liang, F.; Duan, H.; Dong, Y.; Leung, V.; Hu, X. FedSM: Robust Semantics-Guided Feature Mixup for Bias Reduction in Federated Learning with Long-Tail Data. *arXiv* **2025**, arXiv:2510.27240. [CrossRef]
22. Shi, Y.; Shen, L.; Wei, K.; Sun, Y.; Yuan, B.; Wang, X.; Tao, D. Improving the model consistency of decentralized federated learning. In *Proceedings of the 40th International Conference on Machine Learning*; PMLR: Honolulu, HI, USA, 2023; pp. 31269–31291.
23. Hegedűs, I.; Danner, G.; Jelasity, M. Gossip learning as a decentralized alternative to federated learning. In *Proceedings of the IFIP International Conference on Distributed Applications and Interoperable Systems*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 74–90.
24. Xin, R.; Khan, U.A.; Kar, S. An improved convergence analysis for decentralized online stochastic non-convex optimization. *IEEE Trans. Signal Process.* **2021**, *69*, 1842–1858. [CrossRef]
25. Lu, Y.; De Sa, C. Decentralized learning: Theoretical optimality and practical improvements. *J. Mach. Learn. Res.* **2023**, *24*, 1–62.
26. Blanchard, P.; El Mhamdi, E.M.; Guerraoui, R.; Stainer, J. Machine learning with adversaries: Byzantine tolerant gradient descent. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 119–129.
27. Xie, C.; Koyejo, O.; Gupta, I. Generalized byzantine-tolerant sgd. *arXiv* **2018**, arXiv:1802.10116. [CrossRef]
28. Yin, D.; Chen, Y.; Kannan, R.; Bartlett, P. Byzantine-robust distributed learning: Towards optimal statistical rates. In *Proceedings of the 35th International Conference on Machine Learning*; PMLR: Stockholm, Sweden, 2018; pp. 5650–5659.
29. Guerraoui, R.; Rouault, S. The hidden vulnerability of distributed learning in byzantium. In *Proceedings of the 35th International Conference on Machine Learning*; PMLR: Stockholm, Sweden, 2018; pp. 3521–3530.
30. Baruch, G.; Baruch, M.; Goldberg, Y. A little is enough: Circumventing defenses for distributed learning. *Adv. Neural Inf. Process. Syst.* **2019**, *32*, 8632–8642.
31. Mai, P.; Yan, R.; Pang, Y. Rflpa: A robust federated learning framework against poisoning attacks with secure aggregation. *Adv. Neural Inf. Process. Syst.* **2024**, *37*, 104329–104356.
32. Fang, M.; Nabavirazavi, S.; Liu, Z.; Sun, W.; Iyengar, S.S.; Yang, H. Do we really need to design new byzantine-robust aggregation rules? *arXiv* **2025**, arXiv:2501.17381. [CrossRef]
33. Naseri, M.; Hayes, J.; De Cristofaro, E. Local and central differential privacy for robustness and privacy in federated learning. *arXiv* **2020**, arXiv:2009.03561.
34. Monschein, D.; Pérez, J.A.P.; Piotrowski, T.; Nochta, Z.; Waldhorst, O.P.; Zirpins, C. Towards a peer-to-peer federated machine learning environment for continuous authentication. In *Proceedings of the 2021 IEEE Symposium on Computers and Communications (ISCC)*; IEEE: Piscataway, NJ, USA 2021; pp. 1–6.
35. Zhan, Y.; Zhang, J.; Hong, Z.; Wu, L.; Li, P.; Guo, S. A survey of incentive mechanism design for federated learning. *IEEE Trans. Emerg. Top. Comput.* **2021**, *10*, 1035–1044. [CrossRef]
36. Nakamoto, S. Bitcoin: A Peer-to-Peer Electronic Cash System. 2008. Available online: https://bitcoin.org/bitcoin.pdf (accessed on 10 February 2026).
37. Yli-Huumo, J.; Ko, D.; Choi, S.; Park, S.; Smolander, K. Where is current research on blockchain technology?—A systematic review. *PLoS ONE* **2016**, *11*, e0163477. [CrossRef]
38. Zheng, Z.; Xie, S.; Dai, H.N.; Chen, X.; Wang, H. Blockchain challenges and opportunities: A survey. *Int. J. Web Grid Serv.* **2018**, *14*, 352–375. [CrossRef]

39. Muzammal, M.; Qu, Q.; Nasrulin, B. Renovating blockchain with distributed databases: An open source system. *Future Gener. Comput. Syst.* **2019**, *90*, 105–117. [CrossRef]
40. Katz, J.; Lindell, Y. *Introduction to Modern Cryptography: Principles and Protocols*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2007.
41. Bonneau, J.; Miller, A.; Clark, J.; Narayanan, A.; Kroll, J.A.; Felten, E.W. Sok: Research perspectives and challenges for bitcoin and cryptocurrencies. In *Proceedings of the 2015 IEEE Symposium on Security and Privacy*; IEEE: Piscataway, NJ, USA 2015; pp. 104–121.
42. Wood, G. Ethereum: A secure decentralised generalised transaction ledger. *Ethereum Proj. Yellow Pap.* **2014**, *151*, 1–32.
43. Garay, J.; Kiayias, A.; Leonardos, N. The bitcoin backbone protocol: Analysis and applications. *J. ACM* **2024**, *71*, 1–49. [CrossRef]
44. Nguyen, C.T.; Hoang, D.T.; Nguyen, D.N.; Niyato, D.; Nguyen, H.T.; Dutkiewicz, E. Proof-of-stake consensus mechanisms for future blockchain networks: Fundamentals, applications and opportunities. *IEEE Access* **2019**, *7*, 85727–85745. [CrossRef]
45. Buterin, V. A Next-Generation Smart Contract and Decentralized Application Platform. *White Paper*. 2014. Available online: https://ethereum.org/whitepaper/ (accessed on 10 February 2026).
46. Christidis, K.; Devetsikiotis, M. Blockchains and smart contracts for the internet of things. *IEEE Access* **2016**, *4*, 2292–2303. [CrossRef]
47. Huang, J.; Lei, K.; Du, M.; Zhao, H.; Liu, H.; Liu, J.; Qi, Z. Survey on blockchain incentive mechanism. In *Proceedings of the International Conference of Pioneering Computer Scientists, Engineers and Educators*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 386–395.
48. Karim, M.M.; Qu, Q.; Cai, Y.; Liu, T.; Meng, X. Bitcoin reimagined: A comprehensive study of ordinals and inscriptions protocols for Web3 asset innovation. *Blockchain Res. Appl.* **2025**, 100379. [CrossRef]
49. Duan, H.; Ma, T.; Qin, Y.; Zeng, R.; Cai, W.; Leung, V.C.M.; Hu, X. DeRelayL: Sustainable Decentralized Relay Learning. *IEEE Trans. Mob. Comput.* **2025**, *24*, 8913–8929. [CrossRef]
50. Duan, H.; Saddik, A.E.; Cai, W. Incentive Mechanism Design Toward a Win–Win Situation for Generative Art Trainers and Artists. *IEEE Trans. Comput. Soc. Syst.* **2024**, *11*, 7528–7540. [CrossRef]
51. Warnat-Herresthal, S.; Schultze, H.; Shastry, K.L.; Manamohan, S.; Mukherjee, S.; Garg, V.; Sarveswara, R.; Händler, K.; Pickkers, P.; Aziz, N.A.; et al. Swarm learning for decentralized and confidential clinical machine learning. *Nature* **2021**, *594*, 265–270. [CrossRef]
52. Ying, C.; Xia, F.; Wei, D.S.; Yu, X.; Xu, Y.; Zhang, W.; Jiang, X.; Jin, H.; Luo, Y.; Zhang, T.; et al. BIT-FL: Blockchain-enabled incentivized and secure federated learning framework. *IEEE Trans. Mob. Comput.* **2024**, *24*, 1212–1229. [CrossRef]
53. Xu, M.; Zou, Z.; Cheng, Y.; Hu, Q.; Yu, D.; Cheng, X. SPDL: A blockchain-enabled secure and privacy-preserving decentralized learning system. *IEEE Trans. Comput.* **2022**, *72*, 548–558. [CrossRef]
54. Boitier, W.; Del Pozzo, A.; García-Pérez, Á.; Gazut, S.; Jobic, P.; Lemaire, A.; Mahe, E.; Mayoue, A.; Perion, M.; Rezende, T.F.; et al. Fantastyc: Blockchain-based federated learning made secure and practical. In Proceedings of the 2024 43rd International Symposium on Reliable Distributed Systems (SRDS), Nara, Japan, 25–27 September 2024; pp. 260–270.
55. Desai, H.B.; Ozdayi, M.S.; Kantarcioglu, M. Blockfla: Accountable federated learning via hybrid blockchain architecture. In Proceedings of the Eleventh ACM Conference on Data and Application Security and Privacy, Virtual, 26–28 April 2021; pp. 101–112.
56. Miao, Y.; Liu, Z.; Li, H.; Choo, K.K.R.; Deng, R.H. Privacy-preserving Byzantine-robust federated learning via blockchain systems. *IEEE Trans. Inf. Forensics Secur.* **2022**, *17*, 2848–2861. [CrossRef]
57. Pandey, S.R.; Nguyen, L.D.; Popovski, P. Fedtoken: Tokenized incentives for data contribution in federated learning. *arXiv* **2022**, arXiv:2209.09775. [CrossRef]
58. Wu, B.; Seneviratne, O. Blockchain-based Framework for Scalable and Incentivized Federated Learning. In Proceedings of the Companion Proceedings of the ACM on Web Conference 2025, Sydney, NSW, Australia, 28 April–2 May 2025; pp. 1761–1767.
59. Wang, P. Fedchain: An efficient and secure consensus protocol based on proof of useful federated learning for blockchain. *arXiv* **2023**, arXiv:2308.15095. [CrossRef]
60. Qu, X.; Wang, S.; Hu, Q.; Cheng, X. Proof of federated learning: A novel energy-recycling consensus algorithm. *IEEE Trans. Parallel Distrib. Syst.* **2021**, *32*, 2074–2085. [CrossRef]
61. Laatikainen, G.; Li, M.; Abrahamsson, P. A system-based view of blockchain governance. *Inf. Softw. Technol.* **2023**, *157*, 107149. [CrossRef]
62. Karim, M.M.; Van, D.H.; Khan, S.; Qu, Q.; Kholodov, Y. Ai agents meet blockchain: A survey on secure and scalable collaboration for multi-agents. *Future Internet* **2025**, *17*, 57. [CrossRef]