Cureus Journal of Computer Science

Open Access Review Article

A Systematic Review on Blockchain-Based Framework for Storing Educational Records Using InterPlanetary File System

Received 02/12/2025
Review began 03/27/2025
Review ended 05/06/2025
Published 05/09/2025

DOI: https://doi.org/10.7759/s44389-024-00745-y

Aryan A. Ayare [1] , Vaishnavi A. Jadhav [1] , Mustafa K. Banatwala [1] , Shashank V. Changlere [1] , Aparna Mote [1] , Pranalini Joshi [2]

1. Computer Department, Zeal College of Engineering and Research, PUNE, IND 2. Computer Department, Vishwakarma University, PUNE, IND

Corresponding authors: Aryan A. Ayare, aryan.rujul@gmail.com, Vaishnavi A. Jadhav, vaishuj122003@gmail.com

Abstract

Blockchain is a decentralized and distributed ledger technology that ensures data security, transparency, and immutability, making it a promising solution for academic record management. Currently, academic records are managed through centralized databases controlled by educational institutions, relying on manual processes, institutional servers, and third-party services. These systems are prone to inefficiencies, data breaches, and authentication challenges, often requiring time-consuming verification processes vulnerable to fraud. Blockchain technology addresses these limitations by offering a decentralized, tamper-proof framework that enhances security, accessibility, and trust in academic credential verification. This study reviews various blockchain platforms, consensus mechanisms, and scalability solutions, with a focus on Hyperledger Fabric and Ethereum, assessing their applicability in educational contexts. Furthermore, off-chain storage techniques like InterPlanetary File System, consensus algorithms, and access control mechanisms are analyzed to optimize the efficient and secure management of sensitive academic data. By integrating blockchain technology, educational institutions can modernize record-keeping, streamline verification processes, and enhance trust in academic credentials, ultimately creating a more secure and transparent academic record management system. Statistical analysis further highlights blockchain's growing adoption in education, demonstrating its effectiveness in reducing fraud, improving accessibility, and ensuring data integrity.

Categories: Blockchain and Cryptocurrency, Scalability, Storage Systems
Keywords: blockchain, academic records, data integrity, hyperledger fabric, ethereum, off-chain storage, consensus algorithms, smart contracts

Introduction And Background

The education sector, much like healthcare, is a domain where vast amounts of sensitive data is generated, stored, and accessed daily [1]. Universities maintain broad records, which include student transcripts, degree certificates, and other official documents essential for academic and professional validation. For example, universities frequently update student records during admissions, course registrations, grade submissions, and graduation processes, which create large volumes of sensitive data that need to be securely managed and protected [2]. However, the traditional practice of managing these records in physical form poses significant challenges. Records can be misplaced, physically damaged, or may get compromised, which results in the loss of data integrity and security. These issues often lead to delays, document duplication, and other challenges in verifying the authenticity of records [3].

Due to these challenges, there is an increasing need for a secure and reliable system, which manages and safeguards educational records. Blockchain technology presents a promising solution to address the limitations of traditional record management by ensuring data integrity, security, and accessibility [4]. This enhances trust among stakeholders, which include students, university administrators, and external organizations, but also ensures seamless and reliable access to validated records. As universities strive to modernize their record-keeping processes, the implementation of blockchain-based systems can play a crucial role in streamlining academic data management.

What is blockchain?

Blockchain is a decentralized and distributed ledger technology. It consists of a chain of blocks, where each block represents a set of transactions. In this context, each record is referred to as a ledger, and a single exchange of data is considered a transaction. Blockchain uses a distributed system to validate each transaction, ensuring that no single entity controls the entire process. A key characteristic of blockchain is its immutability, meaning that once a block or transaction is added to the blockchain, it remains unaltered and cannot be edited. This ensures that the integrity and trustworthiness of the data are maintained. Figure *1* shows the workings of blockchain technology.

What makes blockchain a preferred solution?

The following sections will explore the concepts of centralized [5], decentralized [6], and distributed systems [7], highlighting their unique characteristics and applications. Additionally, blockchain technology will be introduced as a decentralized and distributed system, emphasizing its ability to combine the strengths of both architectures while addressing their limitations.

Centralized Systems

A centralized system operates on a structure where a single, central node acts as the primary authority and coordination hub, streamlining control and management processes. All clients connect to this central node, which is responsible for handling key functions such as processing, storage, and resource allocation. This design offers notable advantages in terms of operational efficiency, as there is no need for coordination among multiple nodes, and maintenance tasks are centralized [8]. However, this model also introduces critical vulnerabilities. A prominent concern is the single point of failure, where a failure of the central server can result in a complete system collapse [9]. Additionally, centralized systems are inherently more susceptible to data breaches and exhibit limited scalability, as the central node may become overwhelmed as the number of connected clients increases. These systems often adopt a star topology, where all clients connect directly to the central server, simplifying both setup and troubleshooting processes but rendering the system heavily reliant on the central node. Common examples include client-server web applications such as email servers and websites, banking systems with centralized databases, and cloud-based services like Google Drive and Dropbox [10]. While centralized systems provide simplicity and efficiency, their constraints in terms of reliability and scalability underscore the need for more robust and adaptable solutions.

Decentralized Systems

Decentralized systems address the inherent limitations of centralized systems by distributing control across multiple nodes, each operating with autonomy and authority. This design significantly enhances resilience by reducing reliance on a single point of failure. Authority is distributed among nodes, with each responsible for managing specific tasks or regions, thereby improving fault tolerance and ensuring system functionality even in the event of node failures. These characteristics make decentralized systems particularly reliable for applications such as blockchain and peer-to-peer networks. Despite their advantages, decentralized systems face notable coordination challenges, as nodes must synchronize to maintain data consistency and integrity. This often necessitates the use of consensus mechanisms such as proof of work (PoW) or proof of stake (PoS) [11]. Although decentralized systems generally scale more effectively than their centralized counterparts, increasing the network size introduces additional complexity in coordination efforts. Commonly, decentralized systems employ tree or partial mesh topologies, where nodes are connected in hierarchical or partially interconnected structures. These topologies offer enhanced resilience, as alternate paths can prevent total system failure if one node becomes non-functional. However, they also introduce greater communication overhead and complexities in managing synchronization. Examples of decentralized systems include blockchain networks like Bitcoin and Ethereum, which leverage distributed ledgers, and content delivery networks that distribute data across multiple servers to ensure faster delivery and redundancy [12].

Distributed Systems

A distributed system completely removes central authority, with all nodes assuming equal responsibility and collaborating to achieve common objectives. In this peer-to-peer model, each node functions simultaneously as a client and a server, sharing resources and processing tasks on an equal footing [13]. This architecture provides exceptional fault tolerance, as the system can remain operational even in the face of multiple node failures by dynamically redistributing tasks among the remaining nodes. Distributed systems are also highly scalable, with the addition of nodes increasing overall capacity and efficiency rather than placing strain on a centralized entity. Despite these advantages, managing distributed systems presents significant challenges due to their complexity. Robust algorithms are required to ensure consistency, security, and synchronization, often utilizing techniques such as eventual consistency, sharding, and replication. Typically, distributed systems employ full mesh or ad hoc topologies, where nodes are interconnected in non-hierarchical configurations. These designs offer high redundancy and efficient load distribution but also introduce substantial communication overhead and demand sophisticated algorithms for routing and maintaining data integrity. Examples of distributed systems include the InterPlanetary File System (IPFS) [14] for distributed file sharing, torrent networks like BitTorrent [15] for direct peer-to-peer file transfers, and large-scale infrastructures such as distributed databases (e.g., Apache Cassandra) and global search engines.

Understanding network topology and system architecture is critical to designing robust systems. Centralized systems suit simpler applications with minimal fault tolerance requirements. Decentralized systems strike a balance between control and resilience. Distributed systems, though complex, excel in large-scale, fault-tolerant, and collaborative environments. The choice depends on specific needs such as

reliability, scalability, and complexity tolerance [16]. Table *1* gives a comparison of mentioned network topology architectures.

Aspect	Centralized (Star Topology)	Decentralized (Tree/Partial Mesh)	Distributed (Full Mesh)
Control	Single central authority	Shared across nodes	Fully distributed and collaborative
Fault Tolerance	Low	Moderate	High
Scalability	Limited	Moderate	High
Complexity	Low	Moderate	High
Examples	Client-server, banking systems	Blockchain, CDNs	IPFS, BitTorrent

TABLE 1: Comparison of Architectures With Network Topology

CDNs, Content Delivery Networks; IPFS, InterPlanetary File System

Decentralization in Educational Record Context

Decentralization holds a critical position in transforming how educational records are managed, offering a range of significant advantages as follows [6]:

Enhanced security: Traditional centralized models often leave sensitive data susceptible to hacking attempts and other breaches. By decentralizing record storage through blockchain or similar distributed technologies, data is dispersed across multiple nodes, substantially minimizing the risk of a single point of failure. This architecture bolsters overall security by making it exceedingly challenging for malicious actors to compromise the entire system.

Data integrity: In centralized frameworks, a single authority governs record creation and verification, which can lead to inconsistencies or inaccuracies. Through decentralization, every transaction is documented on a shared ledger, and any modifications require consensus among the participating nodes. This process ensures that once an educational record is confirmed, it remains immutable, guaranteeing the authenticity and reliability of the data.

Increased transparency: Decentralized systems foster greater transparency by allowing various stakeholders such as students, educational institutions, and employers to verify records in real time without relying on a central repository. This open verification process helps build trust, as all parties can be confident in the accuracy and legitimacy of the documented credentials.

Improved accessibility: In decentralized models, records are not confined to a single institutional server or database. Instead, they can be retrieved by authorized users from any location at any time. This capability removes the bottlenecks commonly found in centralized systems and is particularly advantageous for individuals who require cross-border validation of their academic credentials or need immediate access for professional purposes.

Reduced administrative burden: The shift toward decentralization can streamline operations by automating key processes, such as credential issuance and verification. Relying on a distributed network reduces manual intervention, mitigating the risk of human error while simultaneously lowering administrative overhead for educational institutions.

Collaborative ecosystem: A decentralized infrastructure encourages broader collaboration among educational entities, students, and other stakeholders. By sharing a common network, institutions can seamlessly exchange information, maintain consistent records, and ensure that validated credentials are readily accessible across various platforms. This cooperative environment ultimately contributes to a more efficient and robust educational landscape.

Blockchain technology [17] revolutionizes trust and control by leveraging a decentralized and distributed architecture, addressing centralization issues such as single points of failure and reliance on intermediaries. Decentralization ensures authority is shared among participants, eliminating the need for a central entity and fostering a trustless environment where transactions are securely validated through cryptographic protocols and consensus mechanisms [11,18,19]. The immutable ledger guarantees data integrity by ensuring that recorded information cannot be altered without collective agreement. Blockchain's distributed framework enhances resilience by storing a copy of the ledger across multiple

nodes, ensuring high fault tolerance and global accessibility while preventing overload and enabling scalability. In managing educational records, this approach eliminates vulnerabilities tied to centralized databases, ensuring that credentials are secure, immutable, and accessible for real-time validation [20]. Automation through smart contracts further streamlines processes like credential issuance and verification, reducing errors and delays. By combining decentralization's fairness and transparency with the resilience and scalability of distributed architecture, blockchain provides a robust, efficient, and collaborative solution to the challenges of centralization and trust. Figure *1* shows the working of blockchain technology [21].

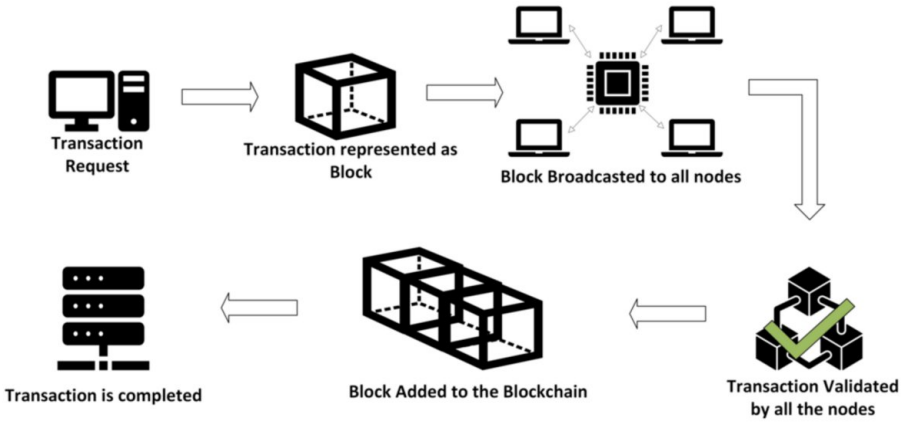

FIGURE 1: Working of Blockchain Technology

Evolution of blockchain

The development of blockchain technology can be summarized into three major phases:

Blockchain 1.0

Blockchain 1.0 is marked as the first stage of blockchain's evolution, which was primarily focused on the use of cryptocurrencies for financial purposes. Bitcoin [22] , the first implementation of Blockchain 1.0, introduced a decentralized way to transfer and store digital currency without the need for mediators like banks. This was revolutionary, but its scope was largely limited to financial use cases [23].

Blockchain 2.0

With Blockchain 2.0, the technology moved beyond simple financial transactions. This phase introduced the concept of smart contracts, which are self-executing contracts where the terms of the agreement are written directly into code [23]. Smart contracts enable the automation of complex processes without the need for mediators [24]. Ethereum is a notable example of Blockchain 2.0 as it allows blockchain architects to develop decentralized applications (dApps) and manage digital assets, also known as smart properties [25]. These assets can be controlled through blockchain-based platforms. Ethereum's programming language, Solidity, became popular for writing these contracts.

Smart contracts: Contracts that have been smart are just self-executing codes stored into a block; they automatically invoke the strictness of the agreement once a condition is satisfied [24,25]. They now become allied with legal agreements to develop ultimate automated services that enable decentralized, trustless, and transparent execution. After deployment, smart contracts operate autonomously, executing the coded instructions without needing any human intervention whenever specified conditions are met. They intrinsically alter their codes with logic and such, so that post-deployment, they cannot be changed even at the behest of an interested party, thus ensuring the sanctity of agreements. That way, building on blockchain makes clearing and settlement rely on decentralized networks for execution and validation rather than intermediaries - an exceptional quality promoting trustless transactions. In addition, such transparency enables all participants to view the coding of the contract itself, thus making it obvious and accountable, whereas the automation brings about speediness and reduces errors.

Smart contracts operate on fairly simple terms these days. The developers encode clause terms into a language of programming compatible with the blockchain, such as Solidity for Ethereum or Rust for Solana. When the smart contract goes live, this unique location is attached to the contract - the address it will have for calling up. In cases where various triggers, like transfer of payment or verification of the data, take place, the contract executes the actions as per the definition, that is, transfer of funds or updating a database. All these actions are recorded on the blockchain, consequently creating an unalterable audit

trail. The elements of the system include the blockchain ledger for transaction storage, contract code defining the logic, state database for real-time updates, execution environments like Ethereum virtual machine, consensus mechanisms to validate transactions, and miners or validators who ensure agreement across the network [25].

Efficient smart contracts will boost the automation of processes, minimize delays, and eliminate human interventions [26]. They are highly cost-effective due to the absence of intermediaries, while blockchain technology ensures safety against tampering or fraud. All actions are recorded clearly and accountably in smart contracts, holding everyone responsible. Most critically, they self-execute based on predefined terms without relying on third parties. These features transform smart contracts into powerful tools for blockchain applications by incorporating trustless automation, security, and transparency [27].

Blockchain 3.0

Blockchain 3.0 represents the current state of development in blockchain technology where it expands its applications into non-financial domains [23]. This includes industries such as distribution chain management, healthcare, and education. For instance, in the education sector, blockchain can be used for document verification, ensuring that academic records, certificates, and transcripts are tamper-proof and easily verifiable. Additionally, the use of non-fungible tokens in Blockchain 3.0 allows for unique digital representations of assets, which could include educational credentials or intellectual property [28]. Blockchain 3.0 is characterized by its focus on solving real-world problems beyond finance, with applications that improve efficiency. Figure *2* shows the evolution of blockchain [23].

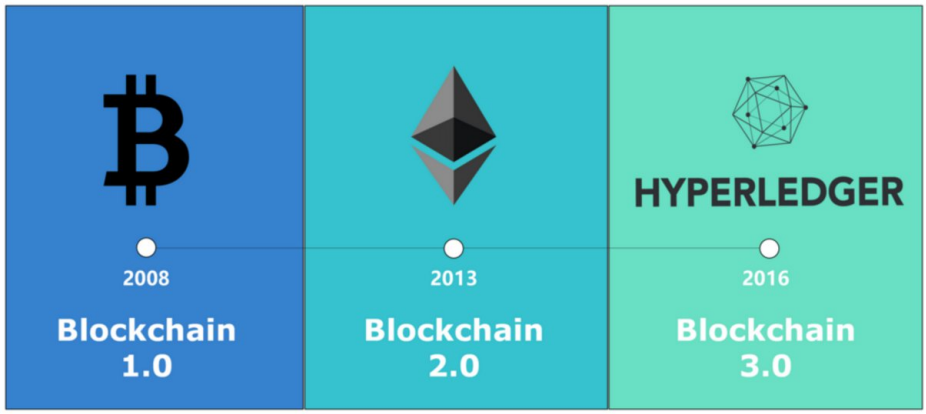

FIGURE 2: Evolution of Blockchain

Smart contracts for various applications

Smart contracts are most helpful in making the automation and securing of transactions streamlined across all industry verticals. Financial services applications include the smart contracts now forming the backbone of decentralized finance, which presupposes loans, insurance, and asset trading without the need for intermediaries [29,30]. This eliminates quite a few layers of complexity and comes with lower operational costs. Escrow contracts provide another advantage in terms of smart contracts, automatically releasing funds on condition of satisfaction, thus minimizing dispute and fraud risks.

What smart contracts now have done has actually created a real-time view of goods moving within the supply chain but again automating payment based on delivery milestones. These two create a window for increased transparency across the supply chain and significantly reduced administrative cycle times. In real estate, an entire transfer of property or lease agreements can be carried out through smart contracts in very simple legal frameworks, speeding up the legal procedures and making trust easier for both involved parties [31]. For health, such contracts are used to manage patient records, guarantee privacy, and automate insurance claims, thus decreasing the load of providing healthcare [32].

Within the domains of dApps, smart contracts act as a backbone to decentralized exchanges and gaming ecosystems platforms [33]. They facilitate peer-to-peer transaction exchanges without central authorities and encourage innovation and greater control over users' assets and data.

Therefore, smart contracts automate the performance of ordinary contracts and make them less dependent on external information. They are useful in a variety of areas-they promise to offer truly remarkable improvements in efficiency, transparency, and security. Yet such a widespread international adoption of these contracts would require overcoming some hurdles such as code vulnerabilities,

scalability issues, and changing legal frameworks. As the development of blockchain heads, smart contracts would not be under the outside consideration of the evolution of the digital economy.

Types of Blockchain

Blockchain technology generally comes in three main varieties, which differ based on who can join and how they participate: Public, Private, and Consortium [34]. Each type has its own architecture, typical use cases, and network structure. Below is a closer look at each category and how their networks are set up.

Public Blockchain

A public blockchain is completely open to anyone that means people can join freely, take part in transactions, and help verify the data on the network [35]. There's no central authority; instead, all members share the same level of control. Figure *3* shows the Network topology of the Public Blockchain. These blockchains often use systems like Proof of Work (PoW) or Proof of Stake (PoS) to ensure security without needing to trust a single entity, relying instead on cryptographic methods [11]. One of the key features here is total transparency every transaction and each block is out in the open, for anyone to see. Because there's no single point of failure, public blockchains tend to be very fault-tolerant, so they keep running smoothly even if some parts of the network go down or get attacked [9].

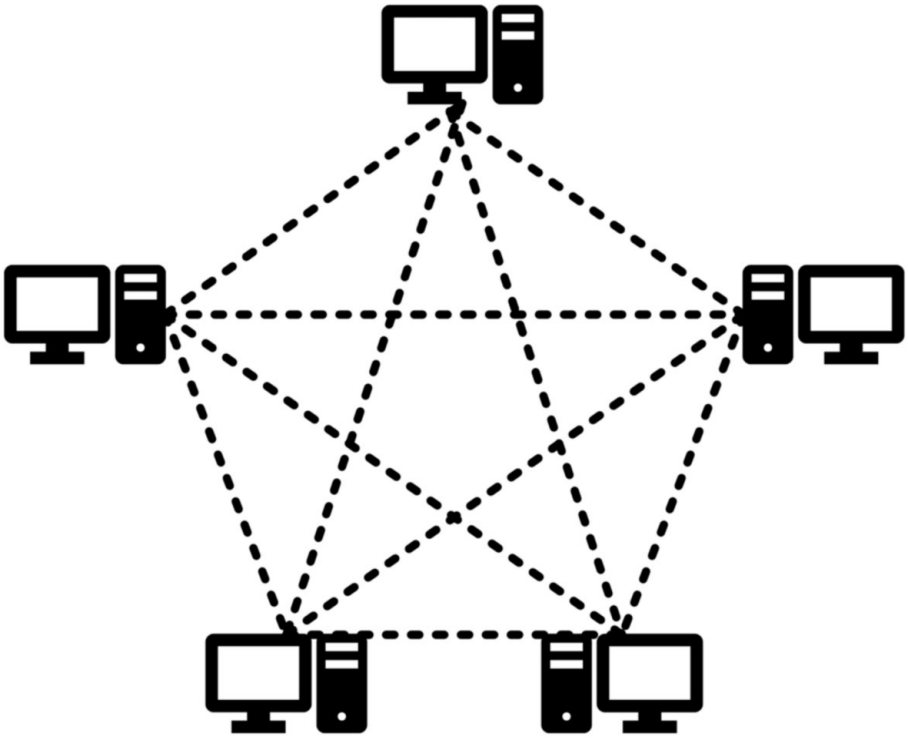

FIGURE 3: Public Blockchain Network Topology

However, there are still a few downsides to think about [36]. For example, many public blockchains can struggle with speed, as you can see in Bitcoin's slower transaction processing. Some of them, especially those using PoW, also require substantial amounts of electricity. On the technical side, public blockchains typically rely on a full mesh topology, where each node connects to many others. While this approach removes single points of failure, it does raise communication costs and complexity. Well-known examples include Bitcoin, the original decentralized digital currency, and Ethereum, a platform built to support dApps and smart contracts [22,23].

Private Blockchain

A private blockchain is a permissioned network that allows authorized participants only, thus making it best suited for organizations or enterprises. In this system, a central administrative authority is involved in participation for some limited number of participants with access to read and write on the blockchain. Figure *4* shows the Network topology of Private Blockchain. With respect to public blockchains, the private ones are the least decentralized as they would have a single organization governing them. However, they are very efficient with transactions taking a short time, with energy savings attributed to

consensus mechanisms like Practical Byzantine Fault Tolerance (PBFT) or Reliable, Replicated, Redundant, And Fault-Tolerant (RAFT) [18,19,37]. Privacy is where only approved participants view transactions and data, thus being confidential. On the downside, since the mechanism is not decentralized, it tends to present less transparency and increase the risk of manipulation by the controlling agent. Topologically, most private blockchains use a star or hierarchical structure where a central node or a small group of nodes meditates everything while controlling the access. This arrangement simplifies management and leads to better performance at the cost of lower fault tolerance and single point of failure risks. Examples include Hyperledger Fabric, a platform supporting modular architecture and private transactions for enterprise use, and Corda, a blockchain tailored for financial institutions, prioritizing security and privacy.

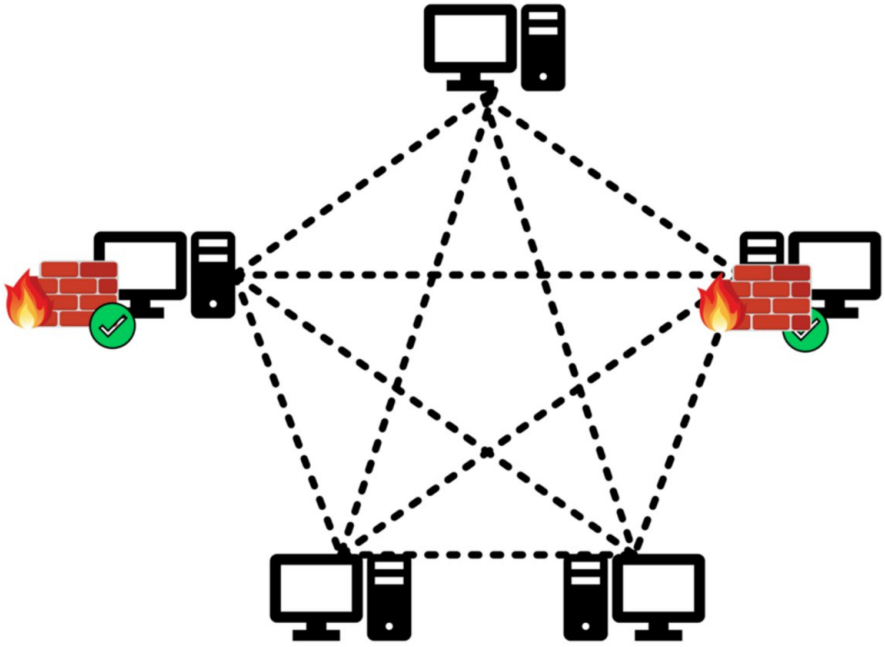

FIGURE 4: Private Blockchain Network Topology

Consortium Blockchain

A consortium blockchain is a partially permissioned network, which is jointly controlled by several organizations sharing the features of public blockchains and private blockchains [38]. Figure *5* shows the network topology of consortium blockchain. The governance for the other parties is defined before contriving it in order to eliminate any one dominating party and to build confidence in all members. It's semi-decentralized to give you a bargain between the features of efficiency that private blockchains afford while still being -of one's own- decentralized. Consortium blockchains will have the best consensus mechanism as it uses PBFT due to the limited size of the network and trust developed between the participants [19].

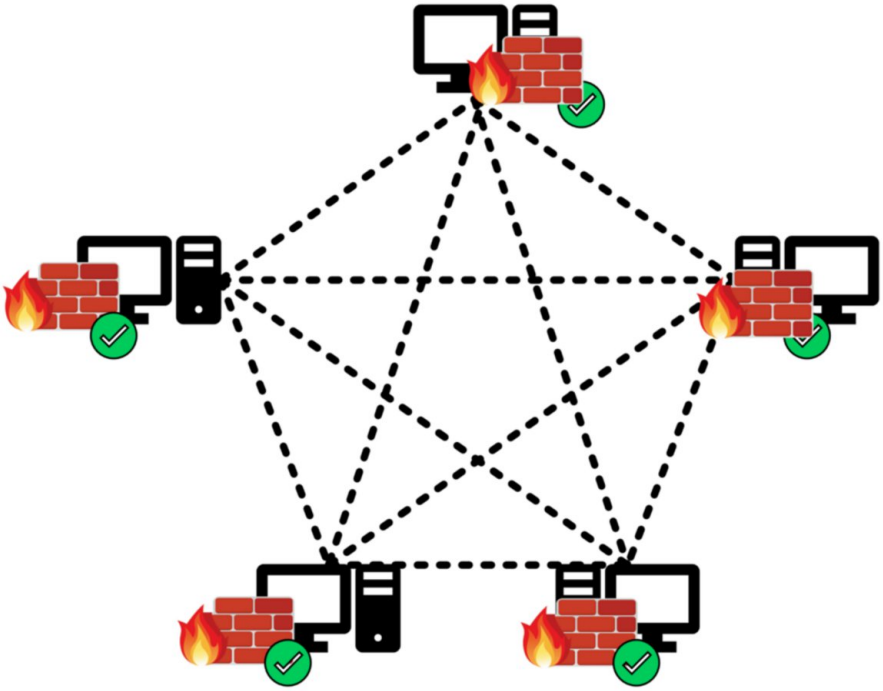

FIGURE 5: Consortium Blockchain Network Topology

Data is accessible only to members of the consortium so that there is privacy yet transparency. However, coordinating with member organizations can be a hassle and often can cause trust issues because the governance is being viewed as biased. Architecturally, consortium blockchains typically use a partial mesh topology, where all nodes are interconnected but not all communicate directly. Certain nodes act as validators or leaders, streamlining operations. This will balance the redundancy needed to avoid failure at node levels using an efficient topological layout to facilitate scalability compared to full mesh topologies and less strength than fully decentralized systems. Energy Web Blockchain is a consortium blockchain created to bring participants from the energy sector together to collaborate in resource management, and R3 Corda is a collaborative decentralized oriented consortium bloc of financial institutions to address the secure and shared transaction processing among financial institutions [39]. Table *2* gives comparison of types of blockchain [40].

Feature	Public Blockchain	Private Blockchain	Consortium Blockchain
Traits	- Trustless environment	- Trusted environment	- Trusted environment
	- Pseudo-anonymity	- Well-known parties	- Well-known parties
	- Public transactions	- Private transactions	- Federated consensus
	- Unknown parties	- Single organization	- Consortium of organizations
Access Control	Open to everyone	Restricted to specific entities	Restricted to consortium members
Governance	Decentralized	Centralized	Collaborative
Consensus	PoW, PoS	PBFT, RAFT	PBFT, RAFT
Privacy	Transparent	Confidential	Semi-transparent
Fault Tolerance	High	Medium	Medium
Performance	Low	High	Moderate

TABLE 2: Comparison of Types of Blockchain

PoW, Proof of Work; PoS, Proof of Stake; PBFT, Practical Byzantine Fault Tolerance; RAFT, Reliable, Replicated, Redundant, And Fault-Tolerant

Choosing between public, private, or consortium blockchains depends on the use case. A public blockchain works best for open or trustless systems such as cryptocurrencies. Private blockchains are best suited for organizations that consider control and confidentiality to be the most important factors, while consortium blockchains serve to create cooperation between multiple entities. The right architectural topology also affects performance, scalability, and resilience considerations; such a need is echoed more by designing one's system to match the specific use cases.

A consortium blockchain is an ideal solution for managing educational records as a secure file-sharing mechanism due to its balance of decentralization, access control, and efficiency. In this context, multiple entities, such as universities, accreditation bodies, students, and banks, require controlled access to sensitive student data, which consortium blockchains facilitate through restricted participation and collaborative governance. Unlike public blockchains, consortium blockchains ensure privacy while maintaining transparency among trusted participants, addressing the confidentiality needs of educational institutions. The partial mesh topology enhances scalability and fault tolerance, ensuring smooth operations and resource optimization. This design enables member institutions to collaboratively validate and share academic credentials, transcripts, and other records while safeguarding data and fostering trust among stakeholders. Consortium blockchains effectively combine privacy, trust, and efficiency, making them the optimal choice for educational record management.

Systematic literature review methodology

In order to achieve the objective of exploring the use of blockchain technology for storing educational records, this study employs a systematic literature review method. An effective review is based on analysis of literature, finding the limitations and research gap in a particular area. The essential purpose of conducting a systematic review is to explore and conceptualize the extant studies, identification of the themes, relations and gaps, and the description of the future directions accordingly.

To ensure a comprehensive evaluation, the review incorporates the following key criteria for analyzing blockchain platforms:

Platform used: Assessing usability and compatibility with educational systems.

Off-chain capabilities: Evaluating how the platform handles transactions and data outside the main blockchain for scalability and efficiency.

Consensus algorithm: Analyzing the mechanism for achieving network agreement and security.

Smart contract functionality: Reviewing the platform's ability to execute automated, self-enforcing contracts.

Data privacy: Examining how the platform protects sensitive educational records.

Cost of transaction: Determining the affordability and economic feasibility of using the platform.

Main use case: Identifying the primary purpose and target audience of the platform.

Access control: Assessing how the platform manages permissions and ensures authorized access to educational records.

Thus, the identified reasons are matched with the aim of this study. This research applies the established strategies for conducting a systematic review, which consists of three phases: organizing the review, managing the review, and reporting the review. Each phase has specific activities, which are described below.

Review Protocol

The review protocol provides the foundation and mechanism to undertake a systematic literature review on the use of blockchain technology for storing educational records. The protocol outlines the research question, objectives, and scope of the review. It also defines the search strategy, inclusion and exclusion criteria, and the methods for selecting and synthesizing the studies. The protocol was designed to ensure transparency, consistency, and reproducibility of the review process. Table *3* presents the inclusion and exclusion criteria used for the review.

Criteria Type	Inclusion Criteria	Exclusion Criteria
Scope	Studies that proposed or implemented a system for storing educational records using blockchain technology	Studies that did not propose or implement a system for storing educational records using blockchain technology
Language	Studies that were published in English	Studies that were not published in English
Publication Year	Studies that were published between 2015 and 2025	Studies that were not published between 2015 and 2025
Availability	Studies that were available in full-text format	Studies that were not available in full-text format
Duplicates	N/A	Studies that were duplicates or had overlapping content

TABLE 3: Inclusion and Exclusion Criteria

Search Strategy Process

The search strategy process involved a comprehensive search of reputable databases, including IEEE Xplore, ScienceDirect, Springer, Scopus, and Google Scholar. The search was conducted using a combination of keywords and search terms, including "Blockchain," "Educational Records," "IPFS," "Consensus Mechanisms," "Access Control," and "Decentralized Storage." Advanced search filters were used to narrow down the search results, including date range, document type, and subject area. The search process was iterative, with the search terms and filters being refined and updated as needed.

Selection Process

The selection process involved a two-stage process. In the first stage, the titles and abstracts of the search results were screened against the inclusion and exclusion criteria. Studies that met the inclusion criteria were selected for further evaluation. In the second stage, the full-text articles of the selected studies were evaluated against the inclusion and exclusion criteria. Studies that met the inclusion criteria were selected for data extraction and synthesis.

Review

Blockchain platforms and consensus algorithms

Overview of Blockchain Platforms

Blockchain technology offers a decentralized and secure framework for managing data across various industries. Leading platforms like Hyperledger Fabric [41], Ethereum [24], and Corda [42] have distinct architectures designed for specific needs. Figure *3* shows different blockchain platforms.

Hyperledger Fabric is a permissioned blockchain designed for enterprises. Its modular setup allows organizations to customize features such as consensus and membership services [43]. It supports private channels and does not require cryptocurrencies. By avoiding gas fees, Fabric is cost-effective and ideal for sectors like supply chain, healthcare, and education, where privacy is essential.

Ethereum is a public blockchain that introduced smart contracts that are self-executing agreements written in code. Its decentralized system ensures transparency, making it suitable for applications with broad network participation. However, transactions require gas fees in Ether (ETH), which can become expensive with high-volume use. Despite this, Ethereum's support for dApps has made it popular in finance, real estate, healthcare, and education [24].

Corda was initially developed for finance but is now also used in industries requiring strict privacy and security. Unlike typical blockchains, Corda only shares transaction data with relevant parties, ensuring confidentiality. This makes it perfect for finance, healthcare, and legal services, where data privacy is critical.

A review of various studies and implementations indicates that Hyperledger Fabric and Ethereum are the most frequently utilized platforms. Hyperledger Fabric is often preferred due to the absence of transaction fees and adaptable permissioned architecture [43]. The elimination of gas fees in Hyperledger Fabric provides a cost-effective solution for enterprises, whereas Ethereum's gas fees associated with each transaction can present a financial obstacle for certain applications.

Overview of Consensus Algorithms

Blockchain networks are based on consensus algorithms to authenticate transactions and append new blocks to ensure that only valid data is recorded. This mechanism involves nodes engaging in consensus processes, commonly referred to as "miners" in proof-based systems. Consensus algorithms are generally divided into two primary categories: proof-based algorithms and vote-based algorithms. Figure *6* shows different types of consensus algorithms [18,19].

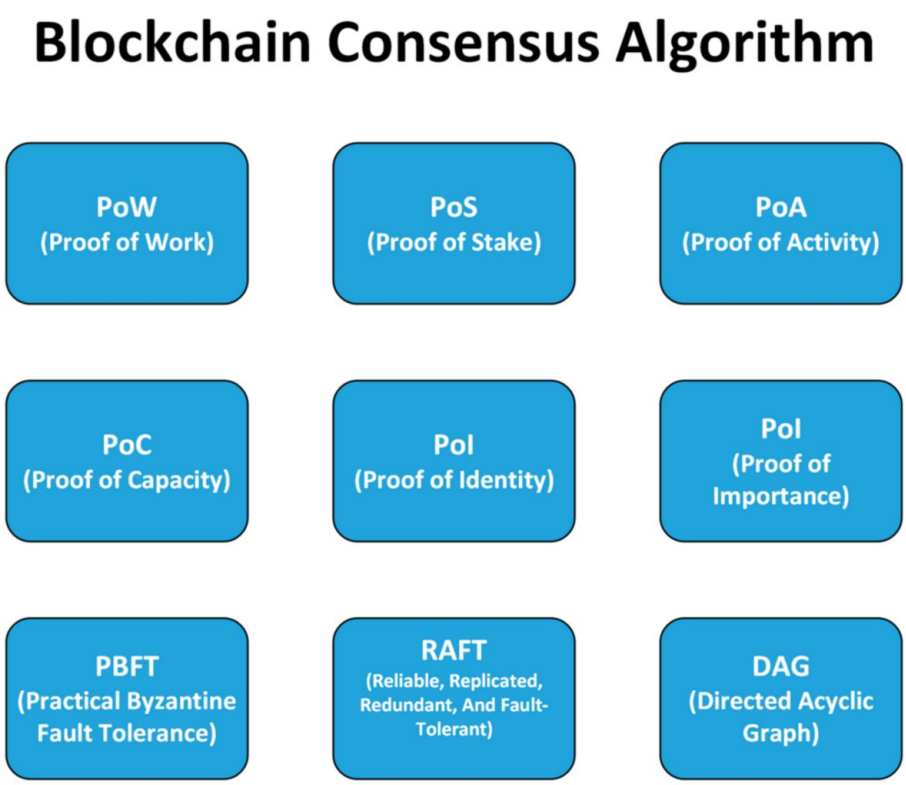

FIGURE 6: Types of Consensus Algorithms

Proof-based algorithms require nodes to present a form of proof to validate transactions:

- PoW: In this method, miners tackle intricate mathematical mysteries (puzzles), consuming considerable computational power and energy. The first miner to solve the puzzle gains the authority to add the new block and receive a reward. While PoW offers robust security, and it is energy-intensive and can be relatively slow [22]. While PoW ensures strong security and immutability, its energy-intensive nature makes it less ideal for educational systems that aim for sustainability and speed.

- PoS: Validators are chosen to create new blocks based on the amount of cryptocurrency they possess and are willing to "stake" as collateral. This approach diminishes energy usage and enhances efficiency compared to PoW, but may result in centralization if a limited number of participants hold substantial stakes. However, the potential for centralization may pose challenges in maintaining a decentralized and equitable system, especially in public educational networks.

Vote-based algorithms achieve consensus through a voting mechanism among nodes:

- PBFT: Employed in permissioned networks with known participants, PBFT involves nodes exchanging messages to collectively agree on the validity of transactions. It provides a high throughput and low latency, therefore making it ideal for enterprise applications. However, the efficiency of PBFT can degrade as the number of nodes increases due to the communication overhead.

- RAFT: A leader-based consensus algorithm used in distributed systems [18]. A leader is elected to manage log replication, and if the leader fails, a new one is chosen. RAFT offers simplicity and fault tolerance but is typically used in smaller, trusted networks due to its reliance on leader coordination.

There is a necessity to have systems implement RAFT consensus, ensuring efficient log replication and leader election for fault tolerance. This will guarantee consistency while maintaining high availability in a distributed environment. Table *4* shows the comparative analysis of four widely used consensus mechanisms.

Consensus Algorithm	PoW	PoS	PBFT	RAFT
Security	High	Moderate	High	Moderate
Performance	Low	Medium	High	High
Decentralization	High	Medium	Low	Low
Energy Efficiency	Low	High	High	High
Governance	Decentralized	Token-based	Centralized	Centralized
Fault Tolerance	Low	Medium	High	Medium

TABLE 4: Comparison of Different Consensus Mechanisms

Source: [44]

PoW, Proof of Work; PoS, Proof of Stake; PBFT, Practical Byzantine Fault Tolerance; RAFT, Reliable, Replicated, Redundant, And Fault-Tolerant

Comparative analysis of consensus mechanisms

Performance and Energy Efficiency

The table highlights the differences in performance and energy efficiency across various consensus algorithms: PoW, PoS, PBFT, and RAFT [11,18,19]. In terms of performance, PoW is rated as low due to its reliance on computationally intensive cryptographic puzzles, which significantly slow down transaction throughput. In contrast, PoS improves scalability by selecting validators based on their stake, resulting in medium performance. PBFT and RAFT excel in performance, and both are rated as high due to their streamlined mechanisms. PBFT achieves quick consensus by minimizing computational requirements, while RAFT efficiently coordinates leader-based systems, making them ideal for environments prioritizing speed and efficiency.

When it comes to energy efficiency, PoW is the least favorable, rated as low, as it demands immense energy for mining activities, raising environmental concerns. On the other hand, PoS offers a high energy efficiency by avoiding energy-intensive computations and relying on stake-based validation. Similarly, PBFT and RAFT are both rated as high for energy efficiency, as they focus on consensus methods that minimize unnecessary computational work. PBFT uses lightweight protocols to maintain high throughput, and RAFT's leader-based design reduces computational overhead, making both algorithms energy-efficient and sustainable options for modern distributed systems.

Security and Fault Tolerance

The table provides an analysis of security and fault tolerance across consensus algorithms: PoW, PoS, PBFT, and RAFT. In terms of security, PoW and PBFT are rated as high due to their robust mechanisms. PoW achieves security by requiring miners to solve computationally intensive puzzles, making attacks like double-spending extremely difficult. Similarly, PBFT ensures high security by requiring multiple nodes to agree on transactions, thereby safeguarding against malicious actors. On the other hand, PoS and RAFT are rated as moderate, with PoS relying on the honesty of validators and RAFT prioritizing simplicity over complex cryptographic guarantees.

For fault tolerance, PBFT leads the way with a high rating as it can handle up to one-third of the nodes behaving maliciously without compromising the network's operations. RAFT and PoS are rated as medium in fault tolerance. RAFT achieves this through its leader-based consensus, which can recover from leader failures but is less resilient to extensive faults. PoS, while moderately fault-tolerant, depends on the distribution of stakes; if a significant number of validators collude, the network's integrity can be compromised. PoW, however, scores the lowest in fault tolerance, as its dependence on computational power makes it vulnerable to 51% attacks where a majority of miners could potentially control the network.

Governance and Decentralization

The table highlights differences in governance and decentralization across various consensus algorithms: PoW, PoS, PBFT, and RAFT. Regarding governance, PoW is characterized by a decentralized governance model, where decisions are made collectively by miners and network participants. PoS, on the other hand, employs a token-based governance system, wherein validators with more tokens have a greater influence, which can lead to a semi-centralized governance structure. PBFT and RAFT both follow a centralized governance approach. PBFT relies on a predetermined set of validators, limiting governance to selected nodes, while RAFT's leader-based architecture inherently centralizes decision-making processes within the system.

In terms of decentralization, PoW is rated as high, reflecting its open and widely distributed network where any participant can contribute computational power to validate transactions. PoS achieves a medium level of decentralization as validator selection is often influenced by token ownership, potentially consolidating power among a smaller group of participants. PBFT and RAFT, however, are rated low in decentralization. PBFT relies on a fixed group of validators, making it less inclusive and more centralized. Similarly, RAFT's leader-centric consensus mechanism prioritizes efficiency and simplicity over decentralization, as decisions depend on the leader node and a small set of follower nodes. This trade-off between decentralization and control is particularly evident in enterprise or permissioned blockchain systems, where governance and decentralization are tailored to meet specific operational needs.

RAFT-based consensus is an excellent choice for managing educational records due to its simplicity, efficiency, and fault tolerance, making it more suitable than other consensus mechanisms. Unlike PoW, which is resource-intensive and slow due to its high computational and energy demands, RAFT ensures efficient and timely log replication with minimal overhead. While PoS reduces energy consumption, its reliance on participants' stakes can lead to centralization risks, which are undesirable in systems aiming for equitable governance among educational institutions. RAFT's leader-based approach simplifies coordination by designating a leader to manage log replication and maintain consistency across nodes, ensuring accurate and up-to-date records which is a critical requirement for trustworthy academic credential management. Its fault-tolerance mechanism, which facilitates seamless leader re-election in case of failure, guarantees high availability and minimizes disruptions. Moreover, RAFT is tailored for smaller, trusted networks, making it ideal for a consortium of educational institutions or permissioned systems with pre-verified participants. With its low communication overhead compared to vote-based mechanisms like PBFT, RAFT offers scalability and reliability, ensuring consistency and integrity in distributed educational record systems [45].

Addressing scalability challenges and off-chain storage solutions

Understanding Scalability Issues in Blockchain

Blockchain networks are prone to scalability challenges when handling large amounts of transactions and data, such as images, certificates, and records. While blockchain ensures data security and integrity, its decentralized structure can create bottlenecks, slowing down the processing of high transaction volumes.

A major limitation is the fixed block size and generation time on most platforms, which restricts the number of transactions processed per second. For instance, both Bitcoin and Ethereum can handle only a limited number of transactions per block, leading to reduced throughput as volumes rise. This results in delays, higher fees, and slower processing times [46].

Adding large data types like images or certificates increases network strain, making on-chain storage inefficient and further slowing down the validation process and increasing storage needs for all nodes.

These scalability issues affect performance and user experience, causing significant delays that obstruct real-time applications like supply chain tracking or academic record verification. This high congestion and longer confirmation times degrades the user experience therefore limiting blockchain's use in high-volume environments.

To support complex applications, particularly in the education sector where large datasets and real-time access are vital, addressing scalability is crucial [47]. Solutions like off-chain storage and layer-2 scaling are being explored to improve performance and user satisfaction.

Introduction to Off-Chain Storage

Off-chain storage involves keeping large data like images, documents, and certificates outside the blockchain while maintaining the transaction details on-chain. This approach helps address blockchain scalability issues, as storing large datasets directly on-chain can slow down performance.

The main advantage of off-chain storage is reducing blockchain bloat. By storing only essential data on-chain, the block size remains smaller and more efficient, improving overall transaction speed [48]. This

method is ideal for applications like educational certificates or digital assets that need to be referenced but do not require full on-chain storage. This makes the blockchain lightweight and scalable. Off-chain solutions are generally centralized or decentralized/distributed systems.

Centralized off-chain storage: A centralized off-chain storage is where the data actually stays on a central server or cloud infrastructure with the blockchain holding a reference towards it either by a hash or a URL. Examples include services such as Amazon S3, Google Cloud, and Microsoft Azure [49,50].

This system lets users upload large files to a centralized server. The server generates a file hash-a cryptographic fingerprint that is unique for the given file-for data integrity purposes. The hash and metadata related to the file, such as its location, are recorded in the blockchain to generate a safe and immutable reference. When a user needs access to the file, it will have to retrieve that file immediately from the central server by using the reference stored on the blockchain for efficient data retrieval, as well as for validation purposes if any.

Centralized off-chain storage is an ideal solution for high-performance and cost-effective applications where decentralization is not a vital concern, but due to dependence on a central authority, it is unsuitable for such applications that prioritize resilience, censorship resistance, or trust minimization. It is therefore centralized off-chain storage. On the other hand, such decoupling will result in decentralized/distributed off-chain storage as an alternative that can overcome all these disadvantages; however, it does come with some trade-offs in complexity as well as costs.

Decentralized off-chain storage: Decentralized off-chain storage systems represent a transformative approach to data management by leveraging peer-to-peer architectures rather than relying on a single central server [51]. These systems distribute data across a network of nodes, ensuring resilience, scalability, and enhanced user control. Unlike centralized models, where a central authority governs data storage and access, decentralized systems empower users with full ownership of their data, allowing them to define access rights and usage policies. Additionally, the integration of blockchain or distributed ledger technologies reinforces data integrity and transparency by recording metadata such as file hashes and ownership details, ensuring permanent immutability and accountability [50]. Table *4* shows the performance metrics comparison of different decentralized off-chain storages.

The core functionality of decentralized off-chain storage revolves around dividing files into smaller encrypted shards and distributing them across multiple nodes in the network. Each shard is stored redundantly on various nodes, guaranteeing data availability even in the event of node failures or network disruptions [52]. Blockchain technology complements this architecture by securely storing metadata, such as cryptographic fingerprints and location pointers, creating a tamper-proof reference for data retrieval. When users request access to a file, the system assembles the required shards, decrypts them, and reconstructs the original file, ensuring both privacy and security. Incentive mechanisms, such as token-based rewards, motivate nodes to contribute storage and bandwidth, fostering a sustainable ecosystem.

The advantages of decentralized off-chain storage are numerous, including high availability, enhanced data sovereignty, and resistance to censorship. By eliminating central points of failure, these systems ensure data accessibility even during outages or attacks. Encryption safeguards privacy, while blockchain integration ensures transparency and trust in data-related transactions. However, the decentralized approach comes with challenges, such as increased complexity in system management, potential delays in data retrieval, and reliance on robust incentive mechanisms to maintain network participation. Despite these trade-offs, decentralized storage offers a compelling alternative for applications prioritizing resilience, security, and user control over conventional centralized solutions [53].

Off-chain solutions enhance performance by reducing the computational load on blockchain nodes, resulting in faster transactions, lower fees, and a better user experience [22]. Technologies like Hadoop Distributed File System (HDFS), IPFS, and MySQL are commonly used storage methods to manage large datasets off-chain while blockchain ensures data validation and security [44,54].

Decentralized systems like IPFS provide secure off-chain storage by storing cryptographic hashes of data on-chain, ensuring verifiability without overloading the blockchain [2,14]. IPFS has been successfully used in implementations like in the study by Chen et al. [55] to create a secure file system.

Existing Decentralized Storage Solutions

Arweave [56] is built on Blockweave technology and specializes in permanent data storage. Its standout features include encryption and data permanence, making it ideal for storing archival data. With high privacy measures and blockchain utilization, Arweave ensures that data remains secure and immutable. It operates on a payment model involving AR tokens, offering a sustainable ecosystem. However, its lack of versioning support and focus on user-level data control limits certain functionalities. Arweave is steadily growing in community adoption, becoming a reliable choice for applications requiring immutable and

transparent data storage. Table *5* gives the comparison of all the Decentralized Storage Platforms.

OrbitDB [57] is a distributed database built on IPFS and Libp2p, designed for decentralized applications (dApps), blockchain-based solutions, and offline-first web applications. Its architecture prioritizes cryptographic security, offering public-key-based access control to safeguard data integrity. Although data is publicly accessible by default, users can implement their encryption for sensitive data. While it does not have a built-in payment model, OrbitDB empowers users with full control over their data, ensuring that every peer hosts its instance. With versioning support via CRDTs and a growing community of contributors, OrbitDB is an emerging player in decentralized storage solutions.

IPFS (InterPlanetary File System) [14] is a peer-to-peer network designed for decentralized file sharing and content addressing. It ensures moderate privacy with features like file identifiers to enhance accessibility and security. While IPFS does not integrate blockchain inherently, it complements blockchain projects by enabling efficient off-chain storage. As an open-source project, it operates on a free payment model, granting user-level control over data. With wide adoption and support for versioning, IPFS has become a cornerstone in the decentralized storage ecosystem, particularly for content distribution and retrieval use cases.

Sia [58] leverages blockchain technology to provide decentralized storage solutions, emphasizing encryption and file splitting for enhanced security. It is highly private and utilizes blockchain for secure file contracts, ensuring reliable and trustless transactions. Operating on a payment model involving Siacoin, Sia allows users to monetize their unused storage while granting end-users complete control over their data. Its support for versioning and niche adoption makes it a viable option for applications requiring secure and cost-effective decentralized storage.

Storj [59] is a decentralized cloud storage platform that uses encryption and data sharding to ensure data security and privacy. Built on a blockchain infrastructure, it operates on a payment model using the STORJ token. Storj provides user-level data control, where data is distributed across multiple nodes for redundancy. With support for versioning and wide community adoption, Storj is particularly suitable for businesses and individuals looking for a scalable and decentralized alternative to traditional cloud storage providers.

Swarm [60] is a decentralized storage and communication system built on Ethereum's Web3 stack. It emphasizes encryption and redundancy to ensure secure and resilient data storage. By leveraging blockchain technology, Swarm provides token-based payment mechanisms using BZZ tokens. Its architecture enables high privacy and user-level control over data. Although it lacks versioning support, Swarm is gaining traction in the decentralized ecosystem due to its focus on reliable and scalable solutions for data storage and management.

Filecoin [61] is a blockchain-based decentralized storage network designed to enable secure and efficient data storage and retrieval. It employs encryption and file contracts to ensure data integrity and privacy. Filecoin uses its native cryptocurrency, Filecoin (FIL), for transactions, creating an incentive model for storage providers. With user-level data control and growing community adoption, Filecoin is a popular choice for applications needing robust and decentralized file storage solutions. However, it does not offer versioning support, which may limit some use cases.

BitTorrent [62] is a widely adopted peer-to-peer file-sharing platform known for its simplicity and efficiency. Its security features are basic, relying on the source of the files being shared. Privacy levels are moderate, making it less suited for sensitive data sharing. BitTorrent operates on a free, ad-supported payment model, with users retaining control over their data. Although it does not support versioning or blockchain integration, its very wide community adoption ensures it remains a popular choice for general file-sharing purposes.

Platform	Security	Cost	Speed	Transparency	Scalability	Anonymity	Immutability	Complexity	Latency
Arweave	High	Moderate	Moderate	High	High	Moderate	Very High	Moderate	Moderate
OrbitDB	High	Free/Variable	Moderate	High	Moderate	High	Moderate	Moderate	Moderate
IPFS	High	Moderate	Moderate	High	High	Moderate	High	Moderate	Moderate
Sia	High	Variable	High	Moderate	Moderate	High	High	Moderate	Moderate
Storj	High	Variable	High	Moderate	High	Moderate	High	Moderate	Low
Swarm	High	Variable	Moderate	Moderate	Moderate	Moderate	High	Moderate	Moderate
Filecoin	High	Variable	High	High	High	Moderate	High	Moderate	Low
BitTorrent	Moderate	High (low for users)	High (popular files)	Moderate	High (popular files)	Low	Moderate	Easy	Low (for popular files)

TABLE 5: Comparison of Off-Chain Storage

Source: [4,3,51]

Comparison of performance metrics

Speed

Speed varies across platforms and significantly impacts user experience. High-speed platforms such as Sia, Storj, and Filecoin are optimized for swift data transfer and processing, making them suitable for applications requiring real-time or frequent access. Conversely, platforms like Arweave, OrbitDB, and IPFS operate at moderate speeds, which may suffice for less time-sensitive use cases, such as archival storage or distributed content delivery. Speed efficiency often hinges on network architecture, data distribution mechanisms, and caching strategies implemented by the platform.

Latency

Latency measures the delay in responding to user requests and varies across the platforms in the table. Platforms like Storj and Swarm demonstrate low latency, making them ideal for applications needing immediate responsiveness, such as media streaming or interactive applications. In contrast, platforms like Arweave, OrbitDB, and IPFS exhibit moderate latency, which could impact their performance in high-demand scenarios. Interestingly, BitTorrent shows low latency specifically for popular files, leveraging its peer-to-peer nature and swarm dynamics to enhance responsiveness for frequently accessed data.

Security

Security is a cornerstone of decentralized storage platforms, ensuring that user data is protected from unauthorized access and breaches. Most platforms in the table, such as Arweave, OrbitDB, IPFS, Sia, Storj, Swarm, and Filecoin, are rated with high security. These platforms employ advanced cryptographic techniques, such as encryption and hashing, to secure data during storage and transfer. Additionally, decentralized architectures distribute data across multiple nodes, reducing the risk of single points of failure. BitTorrent, however, is rated moderate in security, as its peer-to-peer model, while effective for data sharing, lacks robust encryption and authentication mechanisms, making it more susceptible to vulnerabilities.

Transparency

Transparency is another key attribute, highlighting how openly a platform operates and the extent to which users can audit its processes. Platforms like Arweave, OrbitDB, and IPFS excel in transparency, offering users a clear view of data storage and retrieval processes, ensuring trustworthiness in their decentralized frameworks. Filecoin also stands out with high transparency, leveraging its blockchain-based storage verification mechanisms. Conversely, platforms like Sia, Storj, and Swarm are rated moderate in transparency, as their operations may involve more obfuscated processes or require higher technical expertise for users to monitor. BitTorrent, with moderate transparency, reflects its focus on ease of use rather than auditability.

Cost

The cost of decentralized storage platforms varies widely based on their pricing models and the resources required for operation. Platforms like OrbitDB offer "free/variable" costs, making them attractive for users seeking affordable or scalable solutions. Others, such as Sia, Storj, Swarm, and Filecoin, follow variable pricing models that depend on storage needs, bandwidth usage, or market demand, offering flexibility to users. Arweave and IPFS provide moderate cost options, balancing affordability with features. BitTorrent, on the other hand, stands out with "high" cost efficiency for users, especially for sharing popular files, as the platform relies on its peer-to-peer structure, which minimizes operational expenses for individual participants.

Complexity

The complexity of a platform determines its ease of use and the technical expertise required to operate or integrate it into applications. Most platforms in the table, including Arweave, OrbitDB, IPFS, Sia, Storj, and Swarm, are rated with moderate complexity, reflecting a balance between advanced features and usability. These platforms may require some understanding of decentralized networks and APIs, making them ideal for users with a basic technical background. Filecoin also falls into the moderate complexity category, owing to its blockchain-based storage validation and token incentives. BitTorrent is rated "easy," as its user-friendly interface and straightforward functionality make it highly accessible, even for those with minimal technical expertise.

Scalability

Scalability is a critical parameter for decentralized storage platforms as it measures their ability to handle increasing amounts of data and users. Platforms like Arweave, IPFS, and Filecoin score high in scalability, offering robust infrastructures capable of managing vast storage requirements and user bases. These platforms leverage advanced architectures to distribute and replicate data efficiently. In contrast, OrbitDB and Sia are rated as "moderate," indicating that while they can scale, their capacity may be limited by specific constraints such as network size or resource requirements. Storj, with its decentralized cloud storage model, achieves high scalability by distributing data across nodes, making it suitable for enterprise-level applications. Swarm also demonstrates moderate scalability, primarily optimized for medium-scale deployments. BitTorrent excels in scalability for popular files, as its peer-to-peer architecture ensures efficient data sharing even with massive user participation.

Anonymity

Anonymity reflects the degree of user privacy a platform provides. OrbitDB and Sia stand out with high anonymity, safeguarding user identities and data exchanges through encryption and decentralized operations. Platforms like Arweave, IPFS, and Filecoin are rated "moderate" in anonymity, offering privacy features but not at the level of more privacy-focused platforms. Swarm and Storj also maintain moderate anonymity, balancing data accessibility with user privacy. BitTorrent, with its low anonymity rating, does not prioritize user privacy, as its peer-to-peer nature can expose user IP addresses and file-sharing activities. Ensuring anonymity remains essential for platforms that aim to attract privacy-conscious users or operate in sensitive industries.

Immutability

Immutability is a cornerstone of decentralized storage, ensuring that stored data remains tamper-proof and verifiable. Arweave achieves "very high" immutability, leveraging its permanent storage model to preserve data indefinitely. Platforms like IPFS, Sia, Storj, Swarm, and Filecoin provide "high" immutability, using content-based addressing and blockchain-based validation to prevent unauthorized alterations. These platforms are ideal for use cases requiring secure and verifiable data storage, such as legal records or digital archives. OrbitDB's immutability is rated "moderate," as its real-time database updates may not guarantee the same level of permanence as other platforms. Similarly, BitTorrent also scores moderate, as it is not primarily designed for ensuring data immutability but rather for efficient file sharing. Immutability is crucial for maintaining trust and integrity in decentralized systems, particularly in applications demanding long-term data reliability.

IPFS is the ideal choice for storing educational records due to its high security, immutability, scalability, and transparency, which are essential for safeguarding sensitive data like transcripts, diplomas, and certificates. Its decentralized architecture ensures data availability and resilience by eliminating single points of failure, while its content-based addressing guarantees tamper-proof and verifiable records. Unlike other platforms, IPFS offers a cost-effective solution with moderate speed and latency, suitable for archival storage and retrieval needs. Furthermore, its seamless integration with blockchain technologies, such as Hyperledger Fabric, enables enhanced data integrity and provenance tracking, making it superior to alternatives like Arweave, OrbitDB, and Filecoin for educational contexts.

Best practices for scalability in educational contexts

Educational institutions frequently generate large volumes of sensitive data, such as student records, certificates, and other materials which require scalable and efficient systems. To ensure smooth operation and scalability in educational blockchain implementations, specific strategies and storage solutions must be carefully selected.

Strategies for Optimizing Performance

Hybrid on-chain/off-chain models: Store key data like transaction details on-chain, while large files like certificates are kept off-chain. This reduces blockchain size and improves performance. Systems like IPFS or HDFS can manage large files while cryptographic hashes ensure on-chain verification of records [2,63].

Layer-2 scaling solutions: Use layer-2 solutions such as sidechains or state channels to decrease the parent blockchain's load. This speeds up transaction processing while maintaining security and scalability [64,65].

Efficient consensus mechanisms: Implement lightweight consensus algorithms like PBFT or PoS for permissioned networks. These reduce energy use and increase the transaction speed, hence making them ideal for education systems that need efficient validation.

Data access management and user permissions

Importance of Access Control

Access control is truly essential to protect user privacy, secure sensitive information, and comply with regulations like Family Educational Rights and Privacy Act and General Data Protection Regulation [66]. In today's connected world, various entities such as universities for academic verification or banks for loan assessments, need access to student records.

Controlling which entity can access these records is crucial to maintaining their security. For example, access control systems make sure that only authorized entities can view or verify the information [67]. A hierarchical system gives students the most control over their data, allowing them to decide who can access it and for what purpose and therefore addressing the issue of too much administrative control in traditional systems [68].

Organizations such as banks or universities generally have read-only access to verify data without making changes, although some may need to add documents, like certificates, to a student’s record. By using tiered access controls, educational institutions can securely share data while allowing students to manage their information.

Access Control Mechanisms in Blockchain Systems

In several permissioned blockchains like Hyperledger Fabric and Corda, the access control is fundamental to managing who can access and perform actions within the network. Unlike public blockchains, permissioned systems restrict access to only verified participants making them well-suited for handling sensitive data, including educational records.

Role-Based Access Control: Role-Based Access Control assigns permissions according to specific roles. In education, students may have full control over their data, while external entities like banks or universities might receive read-only access for record verification. Certification bodies could be allowed to upload documents but not modify existing records. This type of structure ensures clear and secure role distinctions and access control.

Attribute-Based Access Control: Attribute-Based Access Control offers a way more flexibility by granting access based on user attributes like role, location, or time. For instance, a university might be restricted to accessing transcripts within a certain timeframe, or only specific credentials could be needed for scholarship verification. ABAC enables a dynamic, context-sensitive access control which improves both security and adaptability.

Many proposed systems lack robust access control mechanisms [2,69], highlighting the need for comprehensive solutions in this area.

Identity Management: Blockchain platforms like Hyperledger utilize digital identities to authenticate participants, ensuring that only authorized entities can access the network. This is essential for maintaining privacy while allowing trusted parties to securely verify or interact with the educational data [4]. Effective identity management is a foundation for secure access control, providing a reliable means of

authentication and authorization within blockchain-based educational record systems.

By leveraging these access control mechanisms, blockchain systems can offer a secure, flexible, and efficient management of educational records which ensures that data remains protected while still being accessible to authorized users.

Table *6* shows comparison of reviewed systems.

Fartitchou et al.'s [70] implementation BlockMEDC, built on the Ethereum platform, integrates IPFS for off-chain storage, offering decentralized and efficient handling of large-scale data. The system adopts the PoS consensus algorithm, which ensures energy efficiency and scalability while maintaining a secure and tamper-proof ledger. However, BlockMEDC does not explicitly detail the specific implementation or parameters of PoS, which is critical to understanding its resilience against attacks and its overall security model in permissioned blockchain environments. The inclusion of smart contract functionality enables automated workflows, ensuring transparency and immutability in academic processes like record verification and permission management. Data privacy is bolstered through a combination of permissioned blockchain architecture and encryption, protecting sensitive academic records and institutional data from unauthorized access. These features make BlockMEDC a robust solution for managing academic records and implementing permissioned access frameworks. The system's medium transaction cost positions it as a cost-effective solution for academic use cases, striking a balance between performance and affordability. Additionally, BlockMEDC emphasizes academic records access and permissions management, with built-in access control mechanisms that allow for secure and granular governance of user rights.

Wang et al.'s [71] implementation, built on the Ethereum platform, integrates IPFS for off-chain storage, ensuring scalability and decentralized data management. The system employs the PoW consensus mechanism, which guarantees robust security and immutability. However, the energy-intensive nature of PoW, coupled with slower transaction speeds, limits its suitability for high-throughput applications. Additionally, while the PoW algorithm enhances trust and security, its inefficiency in permissioned or semi-private environments creates a bottleneck for systems requiring fast operations, such as academic document verification and sharing. The system supports smart contracts, allowing the automation of data access workflows and permissions management. Its focus on fine-grained access control ensures that permissions can be precisely tailored to individual users or roles, which is essential for academic scenarios involving sensitive records. Data privacy is achieved through encryption, protecting sensitive academic records during transmission and storage. Despite these strengths, it incurs medium transaction costs, making it moderately efficient but not optimal for budget-conscious implementations. Its primary use cases include data storage, sharing, and academic record access, with a strong emphasis on permission management to ensure secure access to sensitive data. While Wang et al.'s implementation emphasizes permission management, it lacks a comprehensive focus on institutional roles.

Turkanovic et al.'s [72] implementation EduCTX is based on the Ark blockchain platform, employs the delegated proof of stake consensus mechanism, which enhances scalability, reduces energy consumption, and ensures high transaction throughput. Delegated proof of stake leverages a limited number of trusted delegates to validate transactions, resulting in faster block generation times. However, the reliance on a small group of validators can raise concerns about decentralization and vulnerability to collusion, which could affect the trustworthiness of the network in academic environments. EduCTX integrates smart contracts, enabling the automation of workflows such as academic credit transfers and record management. The system adopts a permissionless framework, allowing broader access but potentially exposing it to unauthorized activities. While EduCTX provides adequate data privacy measures, it lacks encryption mechanisms that are critical for securing sensitive academic records, especially in environments where malicious actors may exploit permissionless access. EduCTX is positioned as a medium-cost solution and focuses primarily on credit transfer and academic record management, offering streamlined processes for issuing and verifying student credentials. It includes access control capabilities, enabling users to manage and govern access rights effectively, ensuring a degree of security in credential management workflows.

Gresch et al.'s [73] implementation utilizes the Geth blockchain platform, incorporating IPFS for off-chain storage to handle large data files efficiently while maintaining a decentralized architecture. The system employs the proof of authority (PoA) consensus mechanism, which ensures fast and energy-efficient transaction validation. PoA relies on a limited number of trusted validators, providing high throughput and low latency. However, the centralized nature of PoA raises concerns about the network's resistance to censorship and single points of failure, which could affect trust in highly sensitive academic environments. The inclusion of smart contracts in the proposed framework enables the automation of tasks such as diploma verification and transparency in document handling. Data privacy is enhanced through encryption combined with selective disclosure, allowing specific data fields to be revealed while keeping other information private. This selective approach to data sharing is particularly advantageous in scenarios where students need to share specific credentials without exposing all their academic records. Additionally, the proposed system is a low-cost solution, making it suitable for institutions seeking to

implement a blockchain-based academic document verification system without significant financial overhead. The system's primary use case revolves around diploma verification and ensuring transparency in academic credentials. Access control is a core feature, enabling the governance of who can view, verify, or interact with the academic records, thereby maintaining data security and integrity.

Tariq et al.'s [74] implementation is built on the Ethereum platform but does not utilize any off-chain storage solutions, limiting its ability to handle large data files efficiently. The system employs the PoA consensus mechanism, prioritizing transaction speed and energy efficiency. While PoA provides high throughput and low operational costs, it inherently sacrifices decentralization, which could make the system susceptible to validator collusion or single points of failure. This tradeoff may pose challenges for systems handling sensitive academic records, where trust and resilience are critical. The integration of smart contracts enables the automation of credential verification processes, reducing manual effort and ensuring transparency. Data privacy is maintained through encryption combined with selective disclosure, allowing users to reveal specific information as needed while safeguarding the rest. This approach is particularly useful for academic credential sharing, where students may need to disclose specific qualifications without exposing their entire academic history. With a medium transaction cost, this implementation strikes a balance between affordability and performance, making it a practical choice for institutions with moderate operational budgets. Its primary use case revolves around credential verification, ensuring the authenticity of academic records and facilitating secure interactions between students and external stakeholders. Additionally, the system incorporates access control mechanisms, allowing administrators to manage permissions and maintain the integrity of sensitive data.

Daraghmi et al.'s [75] implementation is built on the Ethereum platform and utilizes the PoA consensus mechanism [75]. PoA is designed for fast and energy-efficient transaction validation by relying on a limited number of trusted validators. While this mechanism ensures low latency and high throughput, it sacrifices decentralization, which could lead to vulnerabilities such as validator collusion or single points of failure. These limitations pose challenges for applications requiring robust security and trust, especially in academic ecosystems where sensitive records must be safeguarded. Although this implementation does not explicitly mention off-chain storage, it supports smart contracts, enabling the automation of processes such as academic record access and permissions management. This automation ensures transparency and reduces manual intervention in document handling. Data privacy is strengthened through encryption combined with permission management, allowing administrators to enforce fine-grained access control to protect sensitive information. The system is cost-efficient, with low transaction costs, making it an attractive choice for institutions with budget constraints. Its primary use case focuses on academic record access and permissions management, ensuring that records are securely shared and accessed within the educational ecosystem. Additionally, it incorporates access control mechanisms to manage user roles and permissions effectively, which is essential for maintaining data integrity and security.

Marhane et al.'s [2] implementation utilizes the Hyperledger Fabric platform and leverages IPFS for off-chain storage, providing an efficient way to handle large files while maintaining decentralization. The consensus mechanism employed is RAFT, which ensures fault tolerance and efficiency in validating transactions. RAFT is particularly suitable for permissioned blockchain networks as it emphasizes consistency and leader-based decision-making. However, the system does not fully explore more decentralized consensus mechanisms like Byzantine Fault Tolerance (BFT), which may better address scenarios requiring higher levels of security and resistance to adversarial behavior. Smart contracts are a key feature, automating workflows such as document sharing and verification. Data privacy is maintained through a combination of permissioned blockchain architecture and encryption, ensuring that access to sensitive information is restricted to authorized users. This makes it well-suited for secure file-sharing applications within educational institutions or similar ecosystems. The system is a cost-efficient solution, with low transaction costs making it accessible to institutions with limited budgets. Its primary use case centers around secure file sharing for universities, enabling efficient and secure document exchange among stakeholders. However, it lacks granular access control mechanisms, limiting its ability to enforce role-based permissions for different user categories, which is a significant drawback in complex academic environments.

Srivastava et al.'s [76] implementation is built on the Ark platform and does not specify an off-chain storage solution. This limitation can restrict its ability to manage large datasets, such as academic records, efficiently. The system employs the PoW consensus mechanism, ensuring a high level of security and decentralization. However, PoW is resource-intensive and may lead to increased energy consumption and slower transaction speeds compared to more lightweight consensus mechanisms such as PoA or RAFT. The system integrates smart contracts, enabling automation of tasks such as document verification and permission management. Data privacy is achieved through a combination of permissioned access and encryption, which ensures that sensitive academic data is accessible only to authorized users. It is focused on credit transfer and academic record management, addressing the need for secure and reliable mechanisms to verify and share educational credentials across institutions. A key strength of this implementation lies in its low transaction cost, making it a cost-effective solution for institutions aiming to digitize academic processes. Additionally, access control mechanisms are included, ensuring that

permissions can be managed to restrict or allow specific users to perform certain actions. However, the absence of an off-chain storage solution and the reliance on PoW present scalability and efficiency challenges, particularly in handling large volumes of academic records.

Kaneriya and Patel's [69] implementation is built on the Geth platform and utilizes IPFS for off-chain storage, ensuring efficient handling of large academic records while maintaining a decentralized structure. The system employs the PoA consensus mechanism, which enables fast and low-cost transactions by relying on a set of pre-approved validators. However, while PoA enhances efficiency, it compromises decentralization and security, as trust is concentrated among a limited number of validators. This makes the system vulnerable to potential collusion or centralized control, which can be a concern in academic environments where integrity and transparency are crucial. Smart contracts play a crucial role in this system, facilitating automated operations such as credential verification, access management, and permission enforcement. The system prioritizes data privacy through encryption combined with selective disclosure, allowing users to share only the necessary portions of their records while keeping the rest confidential. This feature is particularly useful in academic settings, where students may need to verify specific credentials without revealing their entire educational history. One of the system's strengths is its low-cost transaction model, making it financially viable for educational institutions. The primary use case revolves around diploma verification, ensuring that academic credentials are securely stored and easily verifiable by third parties. Additionally, access control is integrated to regulate document interactions, ensuring that only authorized entities can read or modify records.

Mothukuri et al.'s [54] implementation BlockHDFS, built on the Hyperledger platform, demonstrates a robust approach by integrating HDFS for off-chain data storage. This combination leverages HDFS's high reliability and scalability to manage large datasets efficiently, aligning well with applications requiring extensive file-sharing. While BlockHDFS mentions the use of a permissioned model for achieving security and privacy, it does not explicitly state the consensus algorithm used. This omission is significant, as the choice of consensus mechanism is a critical factor in determining the security, scalability, and overall reliability of the blockchain network. In addition to supporting smart contracts, which enable the automation of business processes and logic, BlockHDFS enforces data privacy through a permissioned model. This approach restricts participation to authorized entities, enhancing security and ensuring compliance with organizational or regulatory requirements. The permissioned nature allows for better control over data access and modifications, making it a reliable choice for enterprises prioritizing secure collaboration. BlockHDFS's transaction speed is notably fast, a critical feature for use cases where high throughput and low latency are essential, such as real-time file sharing. Moreover, the cost of transactions remains low, which makes it an economically viable solution for large-scale operations. However, the lack of clarity around the consensus algorithm is a potential drawback, as it leaves uncertainties regarding the system's ability to handle Byzantine faults or prevent potential security threats. Furthermore, BlockHDFS lacks built-in access control mechanisms, which could otherwise provide finer-grained control over user-specific permissions. Despite these limitations, BlockHDFS is particularly suited for file-sharing applications, where speed, cost efficiency, and data privacy are prioritized.

Li and Han's [47] EduRSS, built on the Ethereum platform, leverages MongoDB for off-chain data storage, ensuring scalable and flexible handling of data beyond the blockchain. Ethereum's adoption of the PoW consensus mechanism ensures robust security through computationally intensive mining, but this comes at the cost of slower transaction speeds and higher energy consumption. However, it is important to note that the choice of PoW significantly impacts the system's scalability and energy efficiency, making it less suitable for high-throughput applications. EduRSS supports smart contracts, which automate predefined processes and enable transparent execution of agreements, making it a valuable tool for decentralized applications. Additionally, EduRSS ensures data privacy through encryption, providing an extra layer of security for sensitive information. This feature is particularly useful for educational use cases, where confidentiality of student and institutional data is paramount. Despite its robust security and privacy measures, EduRSS exhibits slower transaction speeds and medium transaction costs, limiting its efficiency in scenarios requiring high-frequency transactions. Moreover, it lacks explicit access control mechanisms, which are critical for applications where role-based or user-specific permissions are essential. While the system is primarily designed for educational purposes, the absence of detailed access control and the reliance on a computationally expensive consensus mechanism may present challenges for its broader adoption in education-centric blockchain implementations.

Liang et al.'s [4] EduChain, utilizing Hyperledger as its platform, integrates IPFS or MySQL for off-chain storage, offering flexible and decentralized handling of data. While it mentions employing the BFT consensus mechanism, this approach is significant as it enhances fault tolerance and ensures consistency in distributed networks, particularly in permissioned blockchains. However, it is crucial to highlight that EduChain does not explicitly detail the specific variant of the BFT algorithm used, which can be a critical determinant of the system's security, scalability, and resilience to Byzantine faults. EduChain supports smart contracts, enabling the automation of processes and ensuring transparent and tamper-proof execution of rules and agreements, a vital feature for decentralized education systems. Data privacy is enforced through encryption, ensuring that sensitive educational records and institutional data are protected against unauthorized access. These features make EduChain a strong candidate for educational

use cases requiring secure and efficient document management. The system achieves medium transaction speeds at a low cost, balancing performance with economic feasibility. This combination is well-suited for educational applications where cost efficiency is a priority. However, EduChain lacks built-in access control mechanisms, which is a notable limitation for scenarios requiring granular role-based permissions. While tailored for education, the absence of specific access control and detailed consensus mechanisms may affect its ability to address the unique needs of complex academic environments comprehensively.

Jahid Alam et al.'s [3] implementation, based on the Hyperledger platform, integrates with IPFS for off-chain storage, providing a decentralized and efficient method for managing large volumes of educational data. It employs the RAFT consensus algorithm, which is leader-based and ensures reliable and efficient consensus in permissioned blockchain networks. However, it is important to note that EduSecure does not provide extensive details about the implementation of RAFT, a key aspect that influences the security and performance of the system. EduSecure supports smart contracts, enabling the automation of critical processes and ensuring transparent and tamper-proof execution of educational workflows. Data privacy is ensured through encryption, making it well-suited for handling sensitive academic records and institutional data. This robust privacy mechanism is particularly valuable in educational environments, where data confidentiality is paramount. While EduSecure exhibits slower transaction speeds and higher transaction costs compared to some alternatives, its primary advantage lies in its incorporation of built-in access control mechanisms. These mechanisms enable granular, role-based permissions, allowing it to differentiate between user roles such as students, faculty, and administrators. This makes EduSecure an ideal solution for educational use cases requiring secure, transparent, and role-specific access to data and documents. Despite its relatively higher costs and slower performance, EduSecure’s focus on access control and privacy makes it a strong candidate for blockchain-based education systems.

Parameter	Platform	Off-Chain	Consensus Algorithm	Smart Contract	Data Privacy	Cost of Transaction	Main Use Case	Access Control
[70]	Ethereum	IPFS	PoS	Yes	Permissioned Blockchain + Encryption	Medium	Academic Records Access & Permissions Management	Yes
[71]	Ethereum	IPFS	PoW	Yes	Encryption	Medium	Data Storage and Sharing, Academic Records Access & Permissions Management	Yes
[72]	Ark	-	DPoS	Yes	Permissionless	Medium	Credit Transfer & Academic Record Management	Yes
[73]	Geth	IPFS	PoA	Yes	Encryption + Selective Disclosure	Low	Diploma Verification & Transparency	
[74]	Ethereum	None	PoA	Yes	Encryption + Selective Disclosure	Medium	Credential Verification	Yes
[75]	Ethereum	-	PoA	Yes	Encryption + Permission Management	Low	Academic Records Access & Permissions Management	Yes
[2]	Hyperledger Fabric	IPFS	RAFT	Yes	Permissioned Blockchain + Encryption	Low	Secure File Sharing for Universities	No
[76]	Ark	-	PoW	Yes	Permissioned + Encryption	Low	Credit Transfer & Academic Record Management	Yes
[69]	Geth	IPFS	BFT	Yes	Permissioned + Encryption	Low	Student Credential Verification	Yes
[54]	Hyperledger	HDFS	-	Yes	Permissioned	Low	File-sharing	No
[47]	Ethereum	MongoDB	PoW	Yes	Encryption	Medium	Education	Yes
[4]	Hyperledger	IPFS/MySQL	BFT	Yes	Permissioned + Encryption	Low	Education	No
[3]	Hyperledger	IPFS	RAFT	Yes	Permissioned + Encryption	High	Education	Yes

TABLE 6: Comparative Analysis of Reviewed Systems

BFT, Byzantine Fault Tolerance; DPoS, Delegated Proof of Stake; HDF, Hadoop Distributed File System; IPFS, InterPlanetary File System; PoA, Proof of Authority; PoS, Proof of Stake; PoW, Proof of Work; RAFT, Raft Consensus Algorithm

While the reviewed systems present innovative approaches to blockchain-based academic record management, each has its own limitations, including scalability issues, lack of decentralized storage, inefficient consensus mechanisms, weak access control models, or limited privacy features. A truly robust and comprehensive system should integrate the strengths of these existing solutions while addressing their shortcomings. Such a system must incorporate efficient off-chain storage mechanisms like IPFS to handle large academic files, scalable and secure consensus mechanisms such as BFT or RAFT to ensure fault tolerance and decentralization, and granular role-based access control to define precise permissions for different stakeholders like HODs, faculty, and students. Furthermore, it should enhance data privacy and security by employing advanced encryption techniques and permissioned blockchain frameworks to protect sensitive academic records. To improve usability, the system should also support document categorization, tagging, and efficient retrieval mechanisms for seamless academic record verification. By overcoming the limitations of existing systems, this integrated approach can create a secure, transparent, and efficient academic record management platform tailored to the complex needs of modern educational institutions.

Adoption challenges and implementation barriers

Technical Barriers

Integrating blockchain into university IT systems comes with several challenges. A major issue is that the existing systems, which are often centralized, need to be adapted to work with blockchain, which is a complicated process that can be time-consuming. Another problem is the lack of staff with the right skills, such as knowledge of blockchain technology and cryptography, which can lead to delays, higher costs, and overall security risks. To overcome these challenges, universities and other organizations need to invest in training and hire experts while also creating systems that work well with their current technology to make the transition smoother.

Legal and Ethical Considerations

The use of blockchain in education brings legal and ethical challenges, particularly in data ownership and intellectual property [66]. It is essential to clarify whether the data belongs to students, institutions, or third parties, especially with academic research or student work stored on a blockchain. The immutability of blockchain data also raises concerns when incorrect or outdated information cannot be modified, complicating personal data management and privacy issues, such as the right to be forgotten [69]. Clear legal frameworks are needed to address data ownership, privacy, and security while maintaining blockchain's benefits.

Applications of blockchain

According to Delgado-von-Eitzen et al. [77] and Dhotre [78], blockchain technology has a wide range of applications in the education sector, offering solutions for secure data management, verification, and decentralization. Some key applications are as follows:

Credential Verification: Blockchain enables secure, tamper-proof digital diplomas and certificates, making it easy for employers and institutions to verify qualifications instantly [74].

Academic Records Management: It provides decentralized and permanent storage for student records, ensuring data integrity and privacy, with easy access across institutions [75].

Smart Contracts for Course Management: Blockchain automates tasks like tuition payments, course enrollments, and scholarship distribution based on preset conditions [79].

Global Credit Transfer and Micro-Credentials: Blockchain allows seamless transfer of academic credits between institutions and supports recognition of micro-credentials for lifelong learning [76].

Figure 7 show some of the applications of blockchain in educational sector.

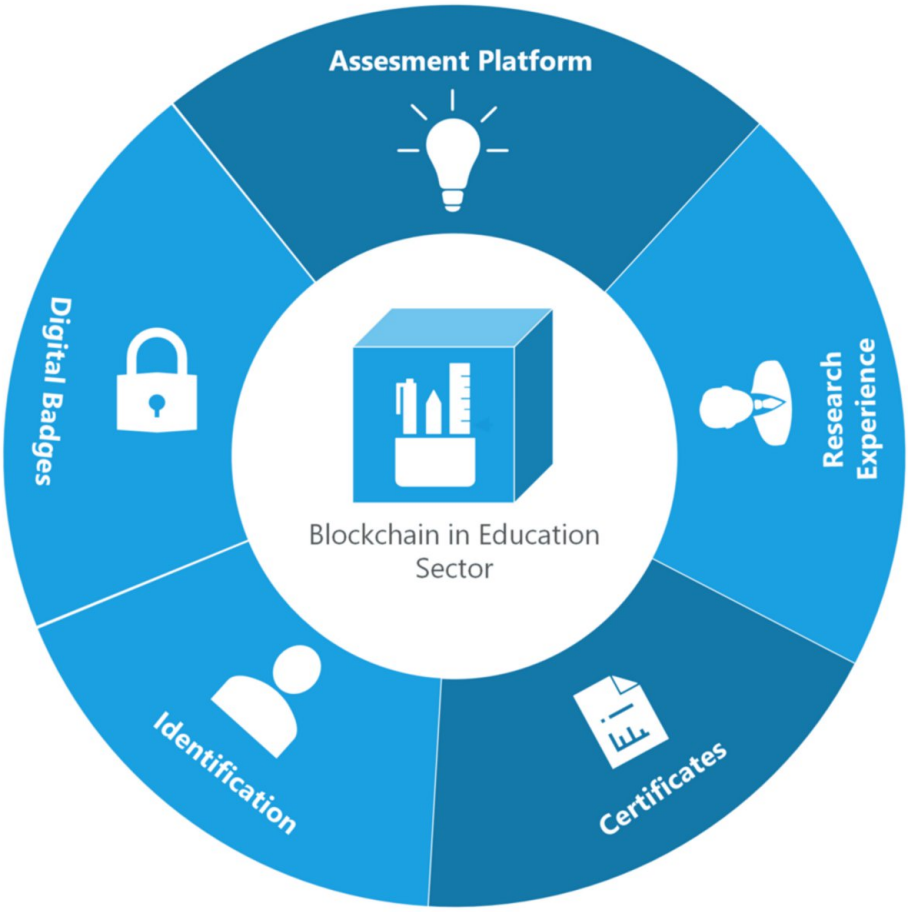

FIGURE 7: Applications of Blockchain in the Education Sector

Conclusions

In conclusion, there is an increasing need for a comprehensive system that integrates the critical features of blockchain technology such as decentralization, immutability, scalability, and robust access control mechanisms to manage educational data efficiently. Such a system would provide a secure, reliable, and tamper-proof environment for storing and verifying academic records and overcoming the limitations of traditional methods. By incorporating off-chain storage solutions, efficient consensus algorithms, and advanced access control methods, blockchain can offer universities and other educational institutions a powerful tool to streamline data management, improve transparency, and ensure the privacy of sensitive information. Implementing such a system would not only enhance operational efficiency but also build trust among students, administrators, and external organizations by providing seamless access to verified academic credentials.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Aryan A. Ayare, Mustafa K. Banatwala, Vaishnavi A. Jadhav, Shashank V. Changlere, Aparna Mote, Pranalini Joshi

Acquisition, analysis, or interpretation of data: Aryan A. Ayare, Mustafa K. Banatwala, Vaishnavi A. Jadhav, Shashank V. Changlere

Drafting of the manuscript: Aryan A. Ayare, Mustafa K. Banatwala, Vaishnavi A. Jadhav, Shashank V. Changlere

Critical review of the manuscript for important intellectual content: Aparna Mote, Pranalini Joshi

Supervision: Pranalini Joshi

Disclosures

Conflicts of interest: In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

References

1. Garcia-Font V: Blockchain: Opportunities and challenges in the educational context. Engineering Data-Driven Adaptive Trust-based e-Assessment Systems. Baneres D, Rodríguez M, Guerrero-Roldán A (ed): Springer, Cham; 2020. 34:133-157. 10.1007/978-3-030-29326-0_7
2. Marhane K, Taif F, Namir A: Secure sharing of university data using Hyperledger Fabric and IPFS system. Procedia Computer Science. 2023, 224:163-168. 10.1016/j.procs.2023.09.024
3. Jahid Alam M, Hossain S, Shekh A, Reno S: Utilizing hyperledger fabric based private blockchain and IPFS to secure educational certificate management. 2022 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE). 2022, 5-11. 10.1109/WIECON-ECE57977.2022.10151082
4. Liang X, Zhao Q, Zhang Y, Liu H, Zhang Q: EduChain: A highly available education consortium blockchain platform based on Hyperledger Fabric. Concurrency and Computation Practice and Experience. 2021, 35:e6330. 10.1002/cpe.6330
5. Chakraborty P, Maitra S, Nandi M, Talnikar S: Centralized systems. Contact Tracing in Post-Covid World. Indian Statistical Institute Series. Springer, Singapore; 2020. 31-70. 10.1007/978-981-15-9727-5_2
6. Tmeizeh M, Rodríguez-Domínguez C, Hurtado-Torres MV: A survey of decentralized storage and decentralized database in blockchain-based proposed systems: Potentials and limitations. Blockchain and Applications, 5th International Congress. Machado JM (ed): Springer, Cham; 2023. 778:204-213. 10.1007/978-3-031-45155-3_21
7. Huang H, Lin J, Zheng B, Zheng Z, Bian J: When Blockchain meets distributed file systems: An overview, challenges, and open issues. IEEE Access. 2020, 8:50574-50586. 10.1109/access.2020.2979881
8. Islam S, Apu KU: Decentralized vs. centralized database solutions in blockchain: Advantages, challenges, and use cases. Global Mainstream Journal of Innovation, Engineering & Emerging Technology. 2024, 3:58-68. 10.62304/jieet.v3i04.195
9. Noveck BS: The single point of failure. Innovating Government. van der Hof S, Groothuis M (ed): T.M.C. Asser Press, Hague, Netherlands; 2011. 20:77-99. 10.1007/978-90-6704-731-9_6
10. Yan C: Cloud Storage Services. Thesis. Centria University of Applied Sciences, Finland; 2017.
11. Fahim S, Katibur Rahman SM, Mahmood S: Blockchain: A comparative study of consensus algorithms PoW, PoS, PoA, PoV. International Journal of Mathematical Sciences and Computing. 2023, 9:46-57. 10.5815/ijmsc.2023.03.04
12. Vakali A, Pallis G: Content delivery networks: Status and trends. IEEE Internet Computing. 2003, 7:68-74. 10.1109/mic.2003.1250586
13. Buford J, Yu H, Lua EK: P2P Networking and Applications. Morgan Kaufmann, Burlington, MA; 2009.
14. Benet J: IPFS - Content addressed, versioned, P2P file system. arXiv:1407.3561. 10.48550/arXiv.1407.3561
15. Xia RL, Muppala JK: A survey of BitTorrent performance. IEEE Communications Surveys & Tutorials. 2010, 12:140-158. 10.1109/surv.2010.021110.00036
16. Vázquez-Rodas A, de la Cruz Llopis LJ: A centrality-based topology control protocol for wireless mesh networks. Ad Hoc Networks. 2015, 24:34-54. 10.1016/j.adhoc.2014.07.026
17. Zheng Z, Xie S, Dai HN, Chen X, Wang H: Blockchain challenges and opportunities: A survey. International Journal of Web and Grid Services. 2018, 14:352-375. 10.1504/ijwgs.2018.095647
18. Fu W, Wei X, Tong S: An improved blockchain consensus algorithm based on Raft. Arabian Journal for Science and Engineering. 2021, 46:8137-8149. 10.1007/s13369-021-05427-8
19. Analysis of the Blockchain Consensus Algorithms. https://appinventiv.com/blog/Blockchain-consensus-algorithms-guide/.
20. Yumna H, Khan MM, Ikram M, Ilyas S: Use of Blockchain in education: A systematic literature review. Intelligent Information and Database Systems. ACIIDS 2019. Nguyen N, Gaol F, Hong TP, Trawiński B (ed): Springer, Cham; 2019. 11432:191-202. 10.1007/978-3-030-14802-7_17
21. Satish Babu BV, Suresh Babu K, Kare DP: Streamlined multi-scenario revocation method leveraging blockchain and auxiliary trees. Bulletin of Electrical Engineering and Informatics. 2024, 13:2547-2554. 10.11591/eei.v13i4.7908
22. Nakamoto S: Bitcoin: A peer-to-peer electronic cash system. 2008.
23. Mukherjee P, Pradhan C: Blockchain 1.0 to Blockchain 4.0—The evolutionary transformation of blockchain technology. Blockchain Technology: Applications and Challenges. Panda SK, Jena AK, Swain SK, Satapathy SC (ed): Springer, Cham; 2021. 203:29-49. 10.1007/978-3-030-69395-4_3
24. Khan SN, Loukil F, Ghedira-Guegan C, Benkhelifa E, Bani-Hani A: Blockchain smart contracts: Applications, challenges, and future trends. Peer-to-Peer Networking and Applications. 2021, 14:2901-2925. 10.1007/s12083-021-01127-0
25. Vitalik B: A next-generation smart contract and decentralized application platform. Ethereum White Paper No. 37. 2014, 1-36.
26. Zheng Z, Xie S, Dai HN, Chen W, Chen X, Weng J, Imran M: An overview on smart contracts: Challenges, advances and platforms. Future Generation Computer Systems. 2020, 105:475-491. 10.1016/j.future.2019.12.019
27. Taherdoost H: Smart contracts in blockchain technology: A critical review. Information. 2023, 14:117.

10.3390/info14020117
28. Zhao R, Chen Z, Xue F: A blockchain 3.0 paradigm for digital twins in construction project management. Automation in Construction. 2023, 145:104645. 10.1016/j.autcon.2022.104645
29. El Hassouni L, Ouchekkir A: Smart contracts: An emerging business model in decentralized finance. Digital Technologies and Applications. ICDTA 2023. Motahhir S, Bossoufi B (ed): Springer, Cham; 2023. 668:197-207. 10.1007/978-3-031-29857-8_20
30. Wang H, Guo C, Chen S: LoC — A new financial loan management system based on smart contracts. Future Generation Computer Systems. 2019, 100:648-655. 10.1016/j.future.2019.05.040
31. Bennett R, Miller T, Pickering M, Kara A-K: Hybrid approaches for smart contracts in land administration: Lessons from three blockchain proofs-of-concept. Land. 2021, 10:220. 10.3390/land10020220
32. Griggs KN, Ossipova O, Kohlios CP, Baccarini AN, Howson EA, Hayajneh T: Healthcare blockchain system using smart contracts for secure automated remote patient monitoring. Journal of Medical Systems. 2018, 42:130. 10.1007/s10916-018-0982-x
33. Min T, Wang H, Guo Y, Cai W: Blockchain games: A survey. 2019 IEEE Conference on Games (CoG). 2019, 1-8. 10.1109/CIG.2019.8848111
34. Paul P, Aithal PS, Saavedra R, Ghosh S: Blockchain technology and its types—A short review. International Journal of Applied Science and Engineering (IJASE). 2021, 9:189-200.
35. Irresberger F, John K, Mueller P, Saleh F: The public blockchain ecosystem: An empirical analysis. SSRN Electronic Journal. 2023, 10.2139/ssrn.3592849
36. Holotescu V, Vasiu R: Challenges and emerging solutions for public blockchains. BRAIN. Broad Research in Artificial Intelligence and Neuroscience. 2020, 11:58-83. 10.18662/brain/11.1/15
37. Pongnumkul S, Siripanpornchana C, Thajchayapong S: Performance analysis of private blockchain platforms in varying workloads. 2017 26th International Conference on Computer Communication and Networks (ICCCN). 2017, 1-6. 10.1109/ICCCN.2017.8038517
38. Dib O, Brousmiche KL, Durand A, Thea E, Ben Hamida E: Consortium blockchains: Overview, applications and challenges. International Journal on Advances in Telecommunications. 2018, 11:51-64.
39. Wang Q, Su M: Integrating blockchain technology into the energy sector — from theory of blockchain to research and application of energy blockchain. Computer Science Review. 2020, 37:100275. 10.1016/j.cosrev.2020.100275
40. Kim JW: Blockchain technology and its applications: Case studies. Journal of System and Management Sciences. 2020, 10:83-93. 10.33168/JSMS.2020.0106
41. Androulaki E, Barger A, Bortnikov V, et al.: Hyperledger fabric: a distributed operating system for permissioned blockchains. EuroSys '18: Proceedings of the Thirteenth EuroSys Conference. 2018, 1-15. 10.1145/3190508.3190538
42. Brown RG: The Corda platform: An introduction. (2018). https://corda-cn.readthedocs.io/zh-cn/latest/_static/corda-platform-whitepaper.pdf.
43. Mahajan P, Ketkar P: Comparative study of permissioned blockchain platform for interoperability and access control of electronic health record. AIDE-2023 and PCES-2023. 2024, 172.
44. Liu Y, Li K, Huang Z, Li B, Wang G, Cai W: EduChain: A blockchain-based education data management system. Blockchain Technology and Application. CBCC 2020. Xu K, Zhu J, Song X, Lu Z (ed): Springer, Singapore; 2020. 3:66-81. 10.1007/978-981-33-6478-3_5
45. Huang D, Ma X, Zhang S: Performance analysis of the raft consensus algorithm for private Blockchains. IEEE Transactions on Systems Man and Cybernetics Systems. 2020, 50:172-181. 10.1109/tsmc.2019.2895471
46. Nyaletey E, Parizi RM, Zhang Q, Choo K-KR: BlockIPFS - Blockchain-enabled interplanetary file system for forensic and trusted data traceability. 2019 IEEE International Conference on Blockchain (Blockchain). 2019, 18-25. 10.1109/Blockchain.2019.00012
47. Li H, Han D: EduRSS: A Blockchain-based educational records secure storage and sharing scheme. IEEE Access. 2019, 7:179273-179289. 10.1109/access.2019.2956157
48. Yang F, Ding Z, Jia L, Sun Y, Zhu Q: Blockchain-based file replication for data availability of IPFS consumers. IEEE Transactions on Consumer Electronics. 2024, 70:1191-120. 10.1109/tce.2024.3364237
49. Xu C, Zhang C, Xu J, Pei J: SlimChain: Scaling blockchain transactions through off-chain storage and parallel processing. Proceedings of the VLDB Endowment. 2021, 14:2314-2326. 10.14778/3476249.3476283
50. Kalajdjieski J, Raikwar M, Arsov N, Velinov G, Gligoroski D: Databases fit for blockchain technology: A complete overview. Blockchain Research and Applications. 2023, 4:100116. 10.1016/j.bcra.2022.100116
51. Merlec MM, In HP: Blockchain-based decentralized storage systems for sustainable data self-sovereignty: A comparative study. Sustainability. 2024, 16:7671. 10.3390/su16177671
52. Yu G, Wang X, Yu K, Ni W, Zhang JA, Li RP: Survey: Sharding in blockchains. IEEE Access. 2020, 8:14155-14181. 10.1109/access.2020.2965147
53. Benisi NZ, Aminian M, Javadi B: Blockchain-based decentralized storage networks: A survey. Journal of Network and Computer Applications. 2020, 162:102656. 10.1016/j.jnca.2020.102656
54. Mothukuri V, Cheerla SS, Parizi RM, Zhang Q, Choo KKR: BlockHDFS: Blockchain-integrated Hadoop distributed file system for secure provenance traceability. Blockchain Research and Applications. 2021, 2:100032. 10.1016/j.bcra.2021.100032
55. Chen J, Zhang C, Yan Y, Liu Y: FileWallet: A file management system based on IPFS and Hyperledger Fabric. Computer Modeling in Engineering & Sciences. 2022, 130:949-966. 10.32604/cmes.2022.017516
56. Williams S, Diordiiev V, Berman L, Raybould I, Uemlianin I: Arweave: A protocol for economically sustainable information permanence. Arweave Yellow Paper. 2019, 67.
57. Peer-to-Peer Databases for the Decentralized Web. (2021). https://orbitdb.org.
58. Vorick D, Champine L: Sia: Simple Decentralized Storage. (2014). Accessed: 20 July 2024: https://sia.tech/sia.pdf.
59. Storj Labs, Inc.: Storj: A Decentralized Cloud Storage Network Framework v3.0. White Paper. (2018). https://static.storj.io/storjv3.pdf.
60. Swarm: SWARM: Storage and Communication Infrastructure for a Self-Sovereign Digital Society. (2021). Accessed: 20 July 2024: https://www.ethswarm.org/swarm-whitepaper.pdf.
61. Protocol Labs: Filecoin: A Decentralized Storage Network. (2017). Accessed: 20 July 2024:

https://filecoin.io/filecoin.pdf.
62. Pouwelse J, Garbacki P, Epema D, Sips H: The Bittorrent P2P file-sharing system: Measurements and analysis. Peer-to-Peer Systems IV. IPTPS 2005. Castro M, van Renesse R (ed): Springer, Berlin, Heidelberg; 2005. 3640:205-216. 10.1007/11558989_19
63. Kumar R, Tripathi R: Blockchain-based framework for data storage in peer-to-peer scheme using interplanetary file system. Handbook of Research on Blockchain Technology. Krishnan S, Balas VE, Julie EG, Robinson YH, Balaji S, Kumar R (ed): Academic Press, Cambridge, MA; 2020. 35-59. 10.1016/B978-0-12-819816-2.00002-2
64. Wei Q, Li B, Chang W, Jia Z, Shen Z, Shao Z: A survey of blockchain data management systems. ACM Transactions on Embedded Computing Systems. 2022, 21:1-28. 10.1145/3502741
65. Athanere S, Thakur R: Blockchain based hierarchical semi-decentralized approach using IPFS for secure and efficient data sharing. Journal of King Saud University - Computer and Information Sciences. 2022, 34:1523-1534. 10.1016/j.jksuci.2022.01.019
66. Mihus I: The main areas of the blockchain technology using in educational management. Economics, Finance and Management Review. 2020, 84-88. 10.36690/2674-5208-2020-4-84
67. Kang P, Yang W, Zheng J: Blockchain private file storage-sharing method based on IPFS. Sensors. 2022, 22:5100. 10.3390/s22145100
68. Ubaka-Okoye M, Azeta AA, Oni AA, Okagbue HI, Nicholas-Omoregbe OS: Securing educational data using agent-based blockchain technology.. International Journal of Scientific & Technology Research. 2020, 9:2936-2938.
69. Kaneriya J, Patel H: A secure and privacy-preserving student credential verification system using blockchain technology. International Journal of Information and Education Technology. 2023, 13:1251-1260. 10.18178/ijiet.2023.13.8.1927
70. Fartitchou M, Lamaakal I, El Makkaoui K, El Allali Z, Maleh Y: BlockMEDC: Blockchain smart contracts for securing Moroccan higher education digital certificates. arXiv:2410.07258. 10.48550/arXiv.2410.07258
71. Wang S, Zhang Y, Zhang Y: A blockchain-based framework for data sharing with fine-grained access control in decentralized storage systems. IEEE Access. 2018, 6:38437-38450. 10.1109/access.2018.2851611
72. Turkanović M, Hölbl M, Košič K, Heričko M, Kamišalić A: EduCTX: A blockchain-based higher education credit platform. IEEE Access. 2018, 6:5112-5127. 10.1109/access.2018.2789929
73. Gresch J, Rodrigues B, Scheid E, Kanhere SS, Stiller B: The proposal of a blockchain-based architecture for transparent certificate handling. Business Information Systems Workshops. BIS 2018. Abramowicz W, Paschke A (ed): Springer, Cham; 2019. 339:185-196. 10.1007/978-3-030-04849-5_16
74. Tariq A, Binte Haq H, Ali ST: Cerberus: A blockchain-based accreditation and degree verification system. IEEE Transactions on Computational Social Systems. 2023, 10:1503-1514. 10.1109/tcss.2022.3188453
75. Daraghmi E-Y, Daraghmi Y-A, Yuan S-M: UniChain: A design of blockchain-based system for electronic academic records access and permissions management. Applied Sciences. 2019, 9:4966. 10.3390/app9224966
76. Srivastava A, Bhattacharya P, Singh A, Mathur A, Prakash O, Pradhan R: A distributed credit transfer educational framework based on blockchain. 2018 Second International Conference on Advances in Computing, Control and Communication Technology (IAC3T). 2018, 54-59. 10.1109/IAC3T.2018.8674023
77. Delgado-von-Eitzen C, Anido-Rifón L, Fernández-Iglesias MJ: Blockchain applications in education: A systematic literature review. Applied Sciences. 2021, 11:11811. 10.3390/app112411811
78. Dhotre S: A survey of blockchain applications beyond cryptocurrency.. International Journal for Research in Applied Science & Engineering Technology. 2019, 7:1102-1105. 10.22214/ijraset.2019.4196
79. Khan M, Naz T: Smart contracts based on blockchain for decentralized learning management system. SN Computer Science. 2021, 2:260. 10.1007/s42979-021-00661-1