ITcon
www.itcon.org - Journal of Information Technology in Construction - ISSN 1874-4753

DECENTRALIZED DATA NETWORKS FOR LIFECYCLE MANAGEMENT IN THE BUILT ENVIRONMENT

SUBMITTED: November 2024
REVISED: May 2025
PUBLISHED: May 2025
EDITOR: Robert Amor
DOI: 10.36680/j.itcon.2025.034

David F. Bucher, Chair of Circular Engineering for Architecture
Institute of Construction and Infrastructure Management, ETH Zurich, Zurich, Switzerland
ORCID: https://orcid.org/0000-0002-1210-1066
bucher@ibi.baug.ethz.ch

Jens J. Hunhevicz Dr. Urban Energy Systems Lab
Empa, Dübendorf, Switzerland
ORCID: https://orcid.org/0000-0002-1296-2384
jens.hunhevicz@empa.ch

Brandon S. Byers, Chair of Circular Engineering for Architecture
Institute of Construction and Infrastructure Management, ETH Zurich, Zurich, Switzerland
ORCID: https://orcid.org/0000-0002-8622-8529
byers@ibi.baug.ethz.ch

Meliha Honic, Sustainability Manager
STRABAG
ORCID: https://orcid.org/0000-0002-8466-7122
meliha.honic-eser@strabag.com

Catherine De Wolf, Assistant Professor, Chair of Circular Engineering for Architecture
Institute of Construction and Infrastructure Management, ETH Zurich, Zurich, Switzerland
ORCID: https://orcid.org/0000-0003-2130-0590
dewolf@ibi.baug.ethz.ch

Daniel M. Hall, Assistant Professor, Design and Construction Management
Faculty of Architecture and the Built Environment, Delft University of Technology, Delft, Netherlands
ORCID: https://orcid.org/0000-0002-0957-484X
d.m.hall@tudelft.nl

SUMMARY: *The management of lifecycle data poses significant challenges for the built environment, hindering effective transformation toward important concepts such as a circular economy. Many recent scholars propose blockchain technologies as a solution; however, there is almost no investigation into decentralized data networks, which also offer significant potential for lifecycle data management. This might be due to a lack of clarity in understanding the fundamental characteristics and potential use cases for decentralized data networks. Therefore, this paper combines a comprehensive review with inductive reasoning to classify three functional typologies—immutable, comprehensive, and privacy-centric – of decentralized data networks. Through testing with material passport data, we evaluate the practical implications of these typologies for lifecycle data management in the built environment. The findings highlight that decentralized data networks can improve data sovereignty and interoperability, but their effectiveness depends on use-case-specific trade-offs, such as mutability, access control, and storage location control. To navigate these trade-offs, the paper derives a decision framework that guides practitioners and researchers in selecting the most suitable decentralized data network. These insights contribute to a better understanding of decentralized technologies beyond blockchain and provide actionable recommendations for the future of data management in the built environment.*

KEYWORDS: *Lifecycle Data Management, Decentralized Data Networks, Blockchain, Circular Economy, Web3.*

REFERENCE: *David F. Bucher, Jens J. Hunhevicz, Brandon S. Byers, Meliha Honic, Catherine De Wolf & Daniel M. Hall (2025). Decentralized Data Networks for Lifecycle Management in the Built Environment. Journal of Information Technology in Construction (ITcon), Vol. 30, pg. 826-851, DOI: 10.36680/j.itcon.2025.034*

1. INTRODUCTION

Improving systematic data management is critical for the built environment and its actors. Better data management can improve both economic performance, where data-driven processes optimize the entire lifecycle and facilitate stakeholder collaboration at various stages (Mêda, Sousa and Hjelseth, 2020; Halttula, Haapasalo and Silvola, 2020), and ecological impacts, especially in the context of a circular economy (Bellini and Bang, 2022). In this paper, lifecycle data management is referred to as the integrative assembly of information through all phases of an asset in the built environment (Hu, 2008; Rezgui, Beach and Rana, 2013; Haider, 2015; Ahonen *et al.*, 2022). Lifecycle data management extends beyond the single lifespan of elements in the linear value chain of the current industry; it also indicates the significant role of data in the transition to a digital circular economy paradigm (De Wolf, Çetin and Bocken, 2024). By collecting and analyzing lifecycle data, the embodied and operational environmental impacts of built assets can be better understood and the areas for system-level improvement identified (Berglund-Brown *et al.*, 2022; Honic and De Wolf, 2023).

However, managing lifecycle data is very difficult in the built environment. The construction industry is recognized as one of the most fragmented sectors, and most construction efforts are one-off collaborations between a diverse set of stakeholders (Rutten, Dorée and Halman, 2009; Hall, 2018; Berg *et al.*, 2021). This industry structure is in contradiction to the current technology landscape, which is dominated by centralized data management approaches with complex, multilayered architectures, centralized platforms, and cloud-based systems (Das, Tao and Cheng, 2021). As a result, construction companies lack incentives to maintain long-term data management systems similar to those implemented in the manufacturing or aerospace industries (Halttula, 2020). In other words, there is a conflict between short-term project thinking and long-term data storage needs. The result is that it is often complex to integrate and store the data created during the design and construction stages of a project with information from maintenance and operations. Managing lifecycle data in this context presents various challenges, such as data exchange within discrete silos (Singh, Gu and Wang, 2011; Corry *et al.*, 2014) and the need for interoperability to mitigate information discrepancies (Pauwels and Terkaj, 2016) (see Section 2.1.

Decentralized data networks (DDNs) offer a promising solution to the challenges of managing lifecycle data in the built environment. DDNs distribute data storage and management across multiple nodes, and control and governance of the system is shared among multiple participants (Bucher *et al.*, 2024). These networks are built on protocols that resemble blockchain technology but differ in focus, emphasizing distributed data storage, management, and exchange without relying on a central authority rather than prioritizing transactional trust mechanisms or consensus. Compared to centralized data solutions, these networks decentralize both data storage and interactions. There have been a number of recent studies that suggest decentralization is particularly relevant to the built environment (Böckel, Nuzum and Weissbrod, 2021; Bambacht and Pouwelse, 2022; Hunhevicz *et al.*, 2023; Naderi, Ly and Shojaei, 2024), as a way to deal with the highly fragmented, decentralized industry structure.

However, many of these previous studies have focused mainly on blockchain-based approaches to decentralization, often emphasizing trust mechanisms between participants rather than the broader challenges of decentralized data storage and management. While blockchain provides transparency, it does not inherently solve issues related to long-term lifecycle data accessibility, storage and ownership in the built environment. As a result, there is a need to move beyond a singular focus on blockchain and consider other forms of decentralized technologies, such as DDNs, which have been underexplored despite their strong potential to address issues of data fragmentation in the built environment.

Therefore, in order to provide better clarity, this paper presents a detailed overview and analysis of decentralized data networks. The structure of this paper is as follows. First, two key background concepts are introduced: lifecycle data management and decentralized data networks (Section 2). Next, the research gap and the scope of the contributions are outlined (Section 3). Next, existing decentralized data networks are reviewed and classified to understand their intrinsic attributes according to three core perspectives (Section 4). This is followed by a use case exploration to evaluate the potential of distinct types of systems for lifecycle data management (Section 5). Next, we present a decision framework designed to guide researchers and practitioners in selecting and implementing appropriate decentralized data networks, aligning specific system requirements with practical solution approaches (Section 6). Finally, the findings are discussed, and the work is summarized (Section 7).

2. BACKGROUND

2.1 Lifecycle Data Management

The objective of lifecycle data management in the built environment is to enhance the collaboration and availability of information in a unified system throughout the lifecycle of the asset. This occurs both during the initial stages of project development (Jiao *et al.*, 2013) and over the long-term of the asset's life, also referred to as the operating and maintenance stages (Zhang *et al.*, 2017). Specifically, this includes managing various types of data, including geometric details of an asset, as well as associated properties such as material specifications, performance metrics, and operational data (Jiao *et al.*, 2013; Wang, Cho and Kim, 2015). In addition, this necessitates the establishment of a data governance system, which can be defined as a conceptual model and set of rules that define and specify the policy for the integrative assembly in a manner suitable for implementation (Rezgui, Beach and Rana, 2013). For example, the definition of how to ensure the validity of the data and its intended use.

The importance of lifecycle data management extends beyond improving the existing value chain; it plays an important role in the concept of a circular economy, a paradigm that has received considerable attention in contemporary discourse (De Wolf, Pomponi and Moncaster, 2017; Kirchherr, Reike and Hekkert, 2017; Pomponi and Moncaster, 2017). The circular economy is a system that promotes sustainable development by minimizing resource input, waste, emissions, and energy leakage of products over time, while securing resources. Its goal is to achieve a more sustainable and efficient use of resources (Hariembrundtland, 1985; Bocken *et al.*, 2016, 2021; Konietzko, Bocken and Hultink, 2020). A key transition that is essential to move from a linear economy to a circular economy is a refined approach to circular data management (Wijewickrama *et al.*, 2021; Wijewickrama, Rameezdeen and Chileshe, 2021; Varriale *et al.*, 2024). The preservation and reuse of existing structures and materials (Ginga, Ongpeng and Daly, 2020) require comprehensive documentation and information management methodologies (Dantas *et al.*, 2021). Solutions have been proposed, such as the material passport, which is a comprehensive record that details the origin, quality, and composition, among other key attributes, of materials used in construction or infrastructure projects (Caldas *et al.*, 2022; Byers *et al.*, 2024). However, achieving this in the built environment is challenged by data dispersion and the need for interoperability between multiple stakeholders (Munaro and Tavares, 2021), underscoring the need for an integrated data system tailored to the needs of the circular economy (Tomczak *et al.*, 2024).

Today, the built environment comprises a variety of technological approaches and systems that address the objective of lifecycle data management (Oluleye, Chan and Antwi-Afari, 2023; Schöggl *et al.*, 2023). There are different variants in this context. In some cases, the focus is on the technology itself, while in others, the emphasis is on an engineering tool enabled by the technology (Bucher *et al.*, 2024). A first example of this is Building Information Modelling (BIM), an industry-wide tool that can manage building data, including geometric and material specifications (M Honic, Kovacic and Rechberger, 2019). BIM is particularly beneficial for lifecycle data management, as it provides a comprehensive digital representation of the physical and functional characteristics of an asset, facilitating better coordination, communication between stakeholders and the sharing of data throughout the lifecycle (Borrmann *et al.*, 2018a; Olawumi and Chan, 2020).

Another example is web-based platforms designed to function as data management systems during various project phases, such as common data environments (CDEs) (Jang *et al.*, 2021) and digital twin platforms (DTs) (Borrmann *et al.*, 2018b). They provide a single source of truth for lifecycle data, linking physical products to their digital counterparts (Heisel and Rau-Oberhuber, 2020; Hunhevicz, Motie and Hall, 2022). As a result, they provide a common data repository and consistency for cross-phase data management. Moreover, there are systems that are based on linked data and semantic web principles. They enable the creation of a network of machine-readable data that is accessible for inspection and use (Pauwels, Costin and Rasmussen, 2022). This approach offers vendor-neutral data access and interoperability across different silos or systems. In summary, different approaches offers viable solutions to improve lifecycle data integration and advance data availability by consolidating information.

However, inherent limitations still persist. In particular, these systems lack an ownership mechanism that gives asset owners unrestricted autonomy over their information rather than centralizing data control (Nawari, 2021). This tends to lock owners into platform dependencies (Mulhall *et al.*, 2022), deprive them of any potential revenue from their own data, and fail to guarantee continued accessibility over the lifetime of the asset (Tao *et al.*, 2018). Moreover, these constraints significantly impede data-driven innovation by limiting the ability of asset owners to leverage their data for novel applications and value creation.

Furthermore, establishing incentive systems for long-term data maintenance or provenance mechanisms is not feasible, as these centralized entities retain full control over data interactions (Yan *et al.*, 2022). This limitation also impacts sustainability efforts, where there is a pressing need to implement robust oversight to ensure long-term accountability and resource stewardship. In the absence of such data stewardship, aligning data management practices with sustainable development goals becomes a significant challenge.

In addition, they lack a strong data governance framework that guides data validation and an incentive structure for effective data exchange and stakeholder participation (Soman, Molina-Solana and Whyte, 2020). Finally, data silos and interoperability issues remain key limitations to effective circular data management (Demirkesen and Tezel, 2021).

2.2 Decentralized Data Networks

Decentralized data networks (DDNs) are a category of systems that distribute data storage and management across multiple nodes, thereby eliminating the need for a central authority (Bucher *et al.*, 2024). A comparison can be drawn between these systems and blockchain technology, as both employ decentralized architectures and peer-to-peer networking to improve transparency. The essential distinction between these two systems lies in their respective objectives. Blockchain prioritizes maintaining an immutable ledger of transactions through consensus mechanisms. In contrast, decentralized data networks focus on the storage, retrieval, and management of data without necessarily requiring such consensus protocols. Unlike traditional centralized data management approaches, DDNs enable data to be stored redundantly across a network, enhancing data availability, resilience, and security (Yu and Jajodia, 2007; Zahed Benisi, Aminian and Javadi, 2020). As a result, they empower data creators with control over their data, fostering data sovereignty and facilitating more collaborative and transparent data management practices.

A prominent example of a DDN is the InterPlanetary File System (IPFS). IPFS is a peer-to-peer protocol designed for storing and sharing hypermedia in a distributed file system (Benet, 2014). In IPFS, files are split into smaller units, encrypted with a cryptographic hash, and distributed across network nodes. Each file or piece of data is identified by a unique content identifier (CID), derived from its content, ensuring data integrity and enabling content-addressed retrieval. This means that any node storing the data can serve it upon request without the need for a central server.

However, it is important to recognize that there are various forms of decentralized data networks that offer distinctive functionalities. A detailed review of different networks is required because the current literature lacks a deep understanding of the range of available technologies and their capabilities, highlighting the need for further research and documentation in this area.

On a high-level, decentralized data networks in the context of lifecycle data management are usually associated with the following promises. First, decentralized data networks promise the same integration of data across different phases of the lifecycle and distributed origins as traditional systems (Cucko and Turkanovic, 2021). This is achieved through an autonomous network without a central entity. Second, they facilitate traceability and transparency through trusted provenance mechanisms, which are especially crucial for effective lifecycle data management (Mendling *et al.*, 2018; Dietz, Putz and Pernul, 2019). Third, open interfaces improve the accessibility and interaction of data. This is critical for seamless data exchange and collaboration, especially in lifecycle data management, where information must move through multiple stages, systems, and stakeholders (Lin, 2013). The two key characteristics of decentralized data networks are distributed ownership and access control. Within these mechanisms, data ownership can be distributed among different stakeholders, embodying, for example, collaborative accountability for products or materials throughout their lifecycle. This enhances data sovereignty for data creators. This is particularly interesting for data management because, unlike current centralized solutions, it offers the opportunity for an incentive system to share and maintain relevant data more broadly (Jaiman, Pernice and Urovi, 2022; Wang *et al.*, 2023). However, a more detailed exploration of DDNs and their intersection with lifecycle data management is not yet available.

2.3 Differentiating DDNs from Blockchain and the Web3 Paradigm

Decentralized Data Networks align with the emerging decentralization trend that offers a fundamental rethink of the architecture of the Internet. This approach is commonly referred to as Web 3.0 (Gan *et al.*, 2023; Guan *et al.*, 2023). Initially, the notion of Web 3.0 was linked to the Semantic Web, which suggested a decentralized structure founded on a collection of concepts and protocols developed to enhance machine-readable and semantically prosperous web (Cena, Farzan and Lops, 2009; Nath, Dhar and Basishtha, 2014). However, with the development of decentralized data technologies, sometimes referred to as Web3, the terminology of Web 3.0 shifted more towards an overarching term consisting of semantic web approaches and decentralized data technologies.

Initially, the vision for Web3 focused on the decentralization of monetary value, with a focus on blockchain and associated cryptocurrencies such as bitcoin (Nandi *et al.*, 2021; Murray, Kim and Combs, 2023). However, the present view of Web3 has expanded to include objectives such as data storage, data sovereignty, scalability, privacy, transparency, and trust-minimized logic (Belk, Humayun and Brouard, 2022; Hackl, Lueth and Bartolo, 2022). In the context of data-centric solutions, Web3 proposes decentralized data networks based on specific peer-to-peer protocols. As a result, these networks enable an architecture that distributes data storage, management, and processing across multiple nodes or participants. This eliminates reliance on a centralized server or entity.

Decentralized data networks can include an attached trustless interaction protocol, e.g., blockchain. This is not strictly necessary, but there is synergy between decentralized data networks and blockchain. Decentralized data networks provide a reliable infrastructure for information management, which can effectively interact with blockchain smart contracts (Turk and Klinc, 2017). However, blockchain researchers in the built environment have been limited in their exploration of this connection. Instead, researchers tend to use one of three approaches to deal with data management:

1. In the ***on-chain approach***, researchers use a blockchain-based backend for smart contracts in combination with a static frontend. All of the information and data is stored on the blockchain itself, referred to as "on-chain". However, this approach has limited capabilities, especially since on-chain data storage can be costly (Elghaish, Abrishami and Hosseini, 2020). Furthermore, the usage of appended data protocols generally requires multiple on-chain transactions (Tao *et al.*, 2021). This can result in a study contribution that oversimplifies the data framework.
2. In the ***centralized server approach***, researchers store data on centralized servers that are trusted. The blockchain smart contracts then interact with this data. However, this results in users losing a certain amount of ownership and control over their data. This model may be suitable when factors such as trust and user specificity are not critical. While this method may offer practical results, it requires compromises regarding data ownership and control that may limit the full potential of such solutions. There are few studies that improve functionality while maintaining key elements such as trust and persistent access.
3. Finally, some researchers in the built environment have used the ***IPFS approach***, which is grounded in this specific decentralized data network. It has been proposed to use it for storing data within blockchain-centric use cases (Oraskari, Beetz and Törmä, 2018; Dounas, Jabi and Lombardi, 2020; Gao and Zhong, 2022; Tao *et al.*, 2023). However, many researchers do not explain their decision-making process for choosing IPFS over other decentralized data networks or may not know which type of decentralized data network best fits their specific data situation.

3. RESEARCH APPROACH

3.1 Research Gap & Questions

In today's built environment, data management occurs through the existing computational infrastructure of the Internet. The present era is marked by the emergence of a centralized web ecosystem, a phenomenon that started taking shape in the 2010s and envelops both infrastructural components and regulatory platform elements (Boeva, Braun and Kropp, 2024). This ecosystem is characterized by the interconnected notions of datafication and platformization (Singh, Gu and Wang, 2011), which represents an evolution from earlier, more static web pages to dynamic, user-generated content and the proliferation of social media platforms. This approach to data and platforms is often referred to as Web 2.0 (Klinc, Dolenc and Turk, 2009; Shang *et al.*, 2011). The resulting value stems from the collaborative activities of stakeholders, unified data formats, and centralized data repositories. Notable examples of the centralized web ecosystem in the built environment are collaborative platforms such as CDEs and BIM-based process management systems (Patacas, Dawood and Kassem, 2020).

To summarize the background, decentralized data networks can serve as the data system necessary for efficient lifecycle data management. While many studies acknowledge the promise of this novel paradigm for structuring lifecycle data (Yu and Jajodia, 2007; Werbrouck *et al.*, 2019; Perera *et al.*, 2020), the nuanced aspects of implementation remain largely unaddressed in academic discussion. For example, there is little investigation of characteristics such as data sovereignty and mutability in a decentralized data infrastructure, especially in the context of lifecycle data management. Overall, there is a lack of compelling theoretical categorization and a corresponding decision-making process to guide the selection of a specific protocol based on the unique needs of a given use case. Therefore, in this paper, the following main research questions are addressed:

What is the potential of decentralized data networks for lifecycle data management in the built environment?

- SQ 1. How can decentralized data networks assist in lifecycle data management?
- SQ 2. How to determine the most appropriate type of DDN for specific lifecycle use cases?

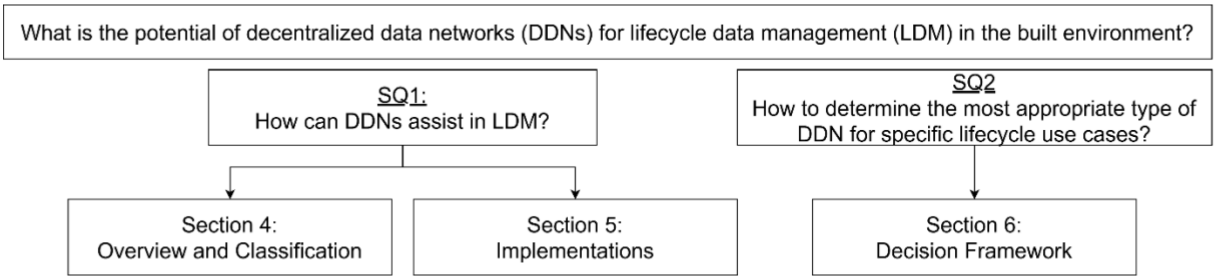


Figure 1: Research approach with the main research question and subsequent two subquestions (SQ) and their corresponding sections.

This paper addresses the two subquestions in separate subsections, which are explained in further detail below. Figure 1 provides an overview of this research approach and a guide to where the corresponding findings can be found for each subquestion.

3.2 Methodology

To answer the above questions, this paper conducts a comprehensive review combined with inductive reasoning. In the initial phase, existing knowledge was gathered from academic databases such as Scopus and Google Scholar. This process involved a systematic selection of papers using keywords such as "*decentralized data networks*", "*decentralized systems*", "*blockchain*", "*lifecycle data management*", and "*built environment*". The selection criteria focused on the relevance to the research scope and to the extent to which the contributions can qualitatively supplement the relevant body of knowledge (Booth, 2016).

In addition, we investigated networks that have been incorporated into prototypes developed by startups as well as those explored in academic research. This allowed us to refine a first body of knowledge by identifying practical applications and emerging trends. This analysis was designed to determine if certain solutions were consistently adopted, which would suggest their potential for broader implementation in lifecycle data management. Based on this first body of knowledge, the following three steps were carried out, which led to the three sections shown in Figure 1.

To build a foundation for the first subquestion, the current technology landscape within the body of knowledge is reviewed to identify existing decentralized data networks and their intrinsic properties (see Section 4). Here, seven main networks are selected – IPFS, Filecoin, Swarm, Gaia, Ceramic, Storj, Arweave – that are frequently used in practice and have already gained some resonance in research. In addition, these networks were selected because they vary slightly in their technological approach to address data management challenges. Therefore, this first analysis is mainly based on technical properties. This results in a classification of these networks based on their unique characteristics.

Based on this understanding, the next step was to examine in detail how decentralized data networks can assist in lifecycle data management. Because one key area of application for lifecycle data management is related to the context of circular economy, we decided for an implementation that is based on a material passport use case with data segments derived from a previous study focusing on a residential building (Meliha Honic, Kovacic and Rechberger, 2019) (see Section 5). In doing so, three different networks were used with an exemplary system implementation to identify their key capabilities. The selection of the three networks implemented was informed by the previous classification that resulted in three broader high-level categories, so one network selected for each category. To then compare and identify their individual strengths and weaknesses, the same use case across all implementations of managing a material passport was used.

Third, a decision framework was developed that consolidates the insights from the previous stages and provides an actionable guide for matching use case requirements with appropriate solution strategies, i.e., when to use which network. This decision framework builds up on functional typologies that emphasize the importance of variations in key functionalities in specific networks for effective lifecycle data management (see Section 6).

4. OVERVIEW AND CLASSIFICATION OF DECENTRALIZED DATA NETWORKS

This section provides a review and classification of decentralized data networks to establish a foundation for selecting networks suitable for lifecycle data management in the built environment (see Figure 1, SQ1). Overall, the review identifies two distinct categories: distributed file systems (DFS) and decentralized storage networks (DSNs). It has been found that three functional typologies can be identified based on a combination of three key perspectives: data sovereignty, data mutability, and universal addressing. In the following, more details on the two distinct categories is given (see Section 4.1), how they can be further divided based on the three key perspectives (see Section 4.2), and how this results in the classification of typologies (see Section **Error! Reference source not found.**).

4.1 The Two Main Categories of Decentralized Data Networks

In principle, decentralized data networks can be divided into two distinct categories. First, there are distributed file systems (DFS), which use a peer-to-peer protocol to distribute files across a multitude of nodes. Once stored, these files remain essentially unchanged (Benet, 2014). Any changes require the modified file to be uploaded again. To increase transparency, certain systems use the blockchain as an anchoring mechanism. Specifically, content is cryptographically aggregated and then timestamped through blockchain recording. This facilitates the validation of both authenticity and integrity. File access in distributed systems is facilitated by universal addressing, which requires knowledge of the location of the file. This often results in data originators relinquishing full data ownership due to a lack of control over file locations within the system.

In contrast, decentralized storage networks (DSNs) emphasize both distributed storage and decentralized ownership. This ensures that the originator has exclusive rights to manipulate data, whether by deleting, modifying, or managing it through a sophisticated identity management architecture. Additionally, these networks seek to cultivate an interconnected data ecosystem by harmonizing the data distributed within it. Such an ecosystem is intended to be accessible to external service providers, thus promoting value creation.

Table 1: Summarizing Table showing the categorization (distributed file system (DFS) or decentralized storage network (DSN)) and the functional typologies of different decentralized data networks based on three key perspectives: Data Sovereignty, Data Mutability, Universal Addressing.

Network	Categorization	Functional Typology	Data Sovereignty	Data Mutability	Universal Addressing
IPFS	DFS	Immutable	(✓)	X	✓
FILECOIN	DFS		(✓)	X	✓
CERAMIC	DSN	Comprehensive	✓	✓	✓
SWARM	DSN		✓	✓	✓
ARWEAVE	DFS		(✓)	(✓)	✓
GAIA	DSN	Privacy-centric	✓	(✓)	X
STORJ	DSN		✓	(✓)	X

4.2 More Detailed Technical Classification

The above two categorizations are a useful starting point. However, the review indicates that they are insufficient to fully explain the variety of decentralized data networks. Therefore, a simplified matrix form is used next to present functional typologies based on three key perspectives that emerge across the two categories: data sovereignty, data mutability, and universal addressing (see Table 1). Data sovereignty refers to the level of control users have over their data. Mutability refers to the ability of the network to modify data without the need to track changes through a tree-structured method involving reuploads. Universal addressing refers to the ability to retrieve data from anywhere on the network (ubiquitous access), for example, through a globally unique identifier (GUID).

The detailed analysis of the seven networks that results in Table 1 can be found in the Appendix A. Here, only the key insights are discussed. Data sovereignty (see Table 1) is present in all solutions, although with varying levels of emphasis. This range of variability indicates a spectrum of data control, from networks that give users exclusive access to their data, ensuring absolute privacy and control, to those where data access is theoretically limited to the user but can be extended to others who know the location of the data. Data mutability (see Table 1) and Universal Addressing (see Table 1) are not always included. For example, Swarm and Ceramic allow users to modify or update stored data, which is useful for use cases that require content evolution over time. However, other networks, such as IPFS lack this feature, which preserves the immutability of data once it is uploaded, thereby enhancing security and integrity but limiting flexibility. Similarly, Arweave, for example, embraces universal addressing, which enables straightforward data retrieval from anywhere in the system, promoting ease of access and interconnectivity. However, other networks, such as Gaia, are avoiding using universal addressing due to prioritizing privacy and security, reducing technical complexity, or enhancing data integrity. Looking at the resulting overall classification in Table 1 suggests that there are three distinguishable typologies each using the same combination of the three perspectives. The following subsection will provide further elaboration on these three typologies.

4.3 Summary: The Three Emerging Typologies of Decentralized Data Networks

This first part of the analysis explored various forms of classification for decentralized data networks, first into two categories (see Section 4.1), then regarding three identified technical perspectives (see Section 4.2). Looking at the resulting classification, three main typologies emerge. The objective here was to address the first part of the first subquestion (see Figure 1, SQ1).

The first typology, termed the ***immutable typology***, is characterized by a lack of data mutability. The immutable typology is best exemplified by IPFS, but also Filecoin based on IPFS fits in this typology. IPFS focuses on data publishing (files or directories) and content addressing, showcasing a unique stance on these dimensions. It emphasizes the absence of inherent data mutability due to its reliance on content-based addressing. Changes in data require a new upload, resulting in a new unique hash. IPFS supports universal addressing through its content identifier system, enabling data to be retrieved from anywhere in the network. The approach to data sovereignty is nuanced. Data are publicly accessible and addressed through cryptographic hashes rather than user- or ownership-based controls.

The second typology, referred to as the ***comprehensive typology***, demonstrates a comprehensive perspective that includes data sovereignty, data mutability, and universal addressing. These include Ceramic, Swarm, and, to some extent, Arweave, allowing for dynamic data updates and interoperability across applications through its blockchain-agnostic framework and the use of decentralized identifiers (DIDs). This ensures that users retain control of their data, while facilitating easy access and modification. Its architecture allows the creation, updating, and sharing of verifiable data streams for a wide range of applications, reflecting its comprehensive and adaptable network design. However, the integration of additional functionality, including access mechanisms, shows significant variability within this typology, potentially increasing complexity, and the likelihood of being considered overly complex.

The third typology termed the ***privacy-centric typology***, has a lack of universal addressability. The privacy-centric typology is represented by Gaia and Storj. These networks emphasize user control and data adaptability, allowing dynamic content management and ownership verification through integration with a blockchain. However, due to the lack of universal addressability, data retrieval depends on specific access mechanisms rather than a global identifier system. Gaia's architecture prioritizes data privacy and control, making it suitable for applications that require secure and user-governed data storage solutions.

The three above typologies are useful in determining which main classes of decentralized data networks exist and could be applied to data management in the built environment. However, this is only a theoretical indication of the potential of the three network typologies and the networks they contain. Therefore, section 5 aims to explore their applicability in the built environment and how they can contribute to lifecycle data management. To do so, for each typology, a network is selected that best exemplifies the defining characteristic as a system type. The selection encompasses IPFS to represent the immutable typology, Ceramic for the comprehensive typology, and Gaia as the type of the privacy-centric typology.

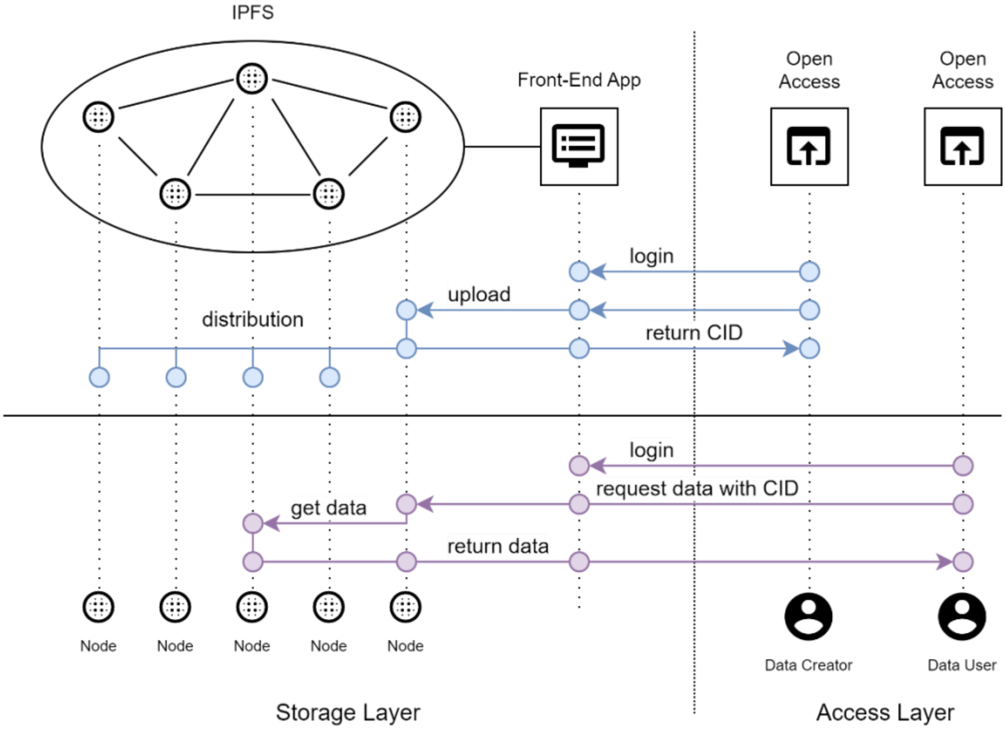


Figure 2: IPFS-based implementation with two technology stack layers showing the workflow of uploading and downloading data.

5. IMPLEMENTATIONS OF DECENTRALIZED DATA NETWORKS FOR A MATERIAL PASSPORT USE CASE

This section demonstrates the implementations for a material passport use case of three different networks representing their typology as a system type (see Section 4.3), i.e. IPFS, Ceramic, and Gaia. The material passport details the components that make up the building envelope and their respective quantities at the structural level. It allows not only discussion of implementation-specific differences but also critical evaluation of how these differences impact their application to lifecycle data management in the built environment. This then serves as the basis for delineating the required decisions to select one network typology over another.

5.1 Typology 1: Immutable Decentralized Data Networks

The first implementation representing the immutable typology (see Section 4.3) uses exemplarly the Interplanetary File System (IPFS), the details of which can be found in Appendix A.2 Interplanetary File System (IPFS). As shown in Figure 2, the architecture and workflows of the system can be explained through the access and storage layers. The access layer includes the associated creators and users. The system is unique in its distributed, decentralized, and permissionless nature. Within the IPFS framework, files are distributed across a network of nodes, without a centralized entity that manages the data. Additionally, the system is designed so that any individual can contribute files that are then distributed throughout the network. After data upload, a distinct content identifier (CID) is assigned to allow users to retrieve the desired information.

These interactions are facilitated by a simple decentralized application (dApp) that has been implemented. However, it is also possible to access the data directly on the IPFS network, bypassing the need for a user interface. Figure 3 illustrates the implemented application that allows users to upload and access content. This process is depicted by the second blue arrow in Figure 2. Each file is assigned a specific CID, which can be used to download the respective file.

As IPFS was used as the foundational network for implementation, it offers several advantages, particularly in terms of data distribution and accessibility. The network emphasizes decentralized and distributed data storage, coupled with content discovery through unique content identifiers (CIDs). Such an approach ensures improved data availability and built-in redundancy mechanisms.

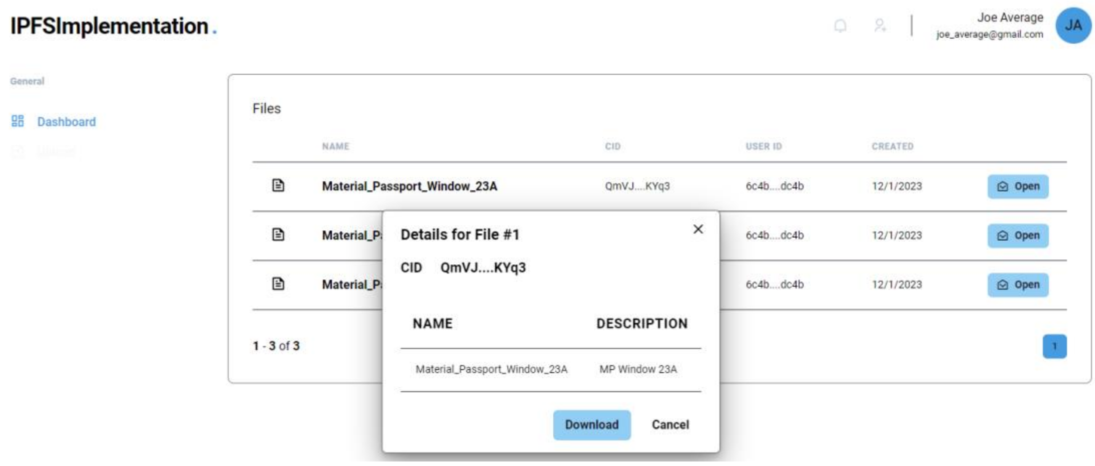


Figure 3: User Interface (UI) of the first implementation based on the Interplanetary File System (IPFS).

However, it presents challenges that counteract some of the promises mentioned above. In particular, the mechanisms for data ownership and access control are rather vague. Although the data is stored in a decentralized and secure manner, its control and governance are insufficient due to centralized mechanisms. This is particularly troubling in the context of lifecycle data management, where tight control over data access and ownership is critical. In addition, data may be uploaded without rigorous controls and without a definitive model for validation. This laxity can lead to inconsistencies in the data, compromising its overall quality. Therefore, when considering IPFS for specific use cases, it is critical to identify potential roadblocks related to ownership, access control, and data model validation.

5.2 Typology 2: Comprehensive Decentralized Data Networks

For the second implementation that represents the second typology (see Section 4.3), the objective was to improve the attribution of data ownership for easy data traceability within the system. For this implementation Ceramic was used, a decentralized storage network (see Appendix A.6 Ceramic). In particular, this implementation has a four-tier architecture, unlike the earlier two-tier IPFS-based prototype. As demonstrated in Figure 4, there exists a computational and logic layer located between the two layers discussed above.

The data network has an enhanced state management system with computational functions, which is essential as any data modification is an actual change and not just a relinking to a previous version, as observed in IPFS-based systems. Additionally, the system's logic tier supports the creation of decentralized identities utilizing private key mechanisms, serving as the fundamental element for user authentication. This element is similar to a blockchain component. However, the approach focuses primarily on identity management without employing the complex methodologies commonly found in blockchain-based automated logic. This feature is depicted in Figure 4, where the user generates a decentralized identity (DID) through the blockchain layer, which is linked to his wallet address. Through the use of his private key for a signature, the computational peer-to-peer network is able to authenticate the user. This is demonstrated in Figure 5b. After the user logs in, their DID must be signed with their private key, i.d. his wallet used.

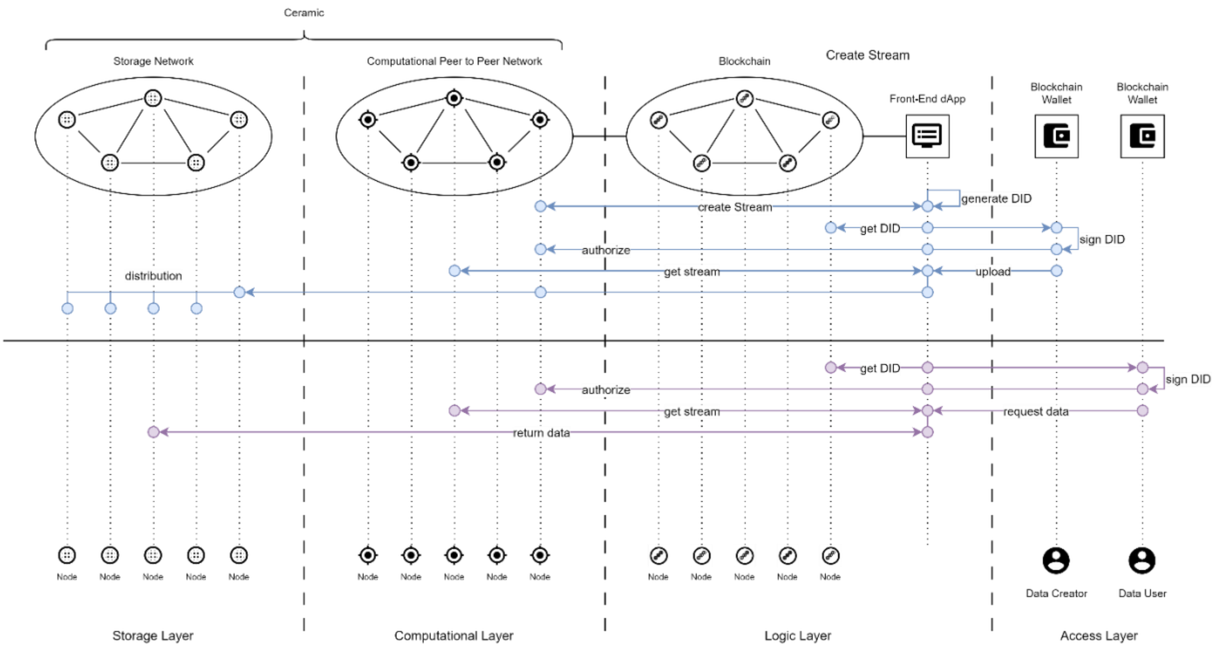
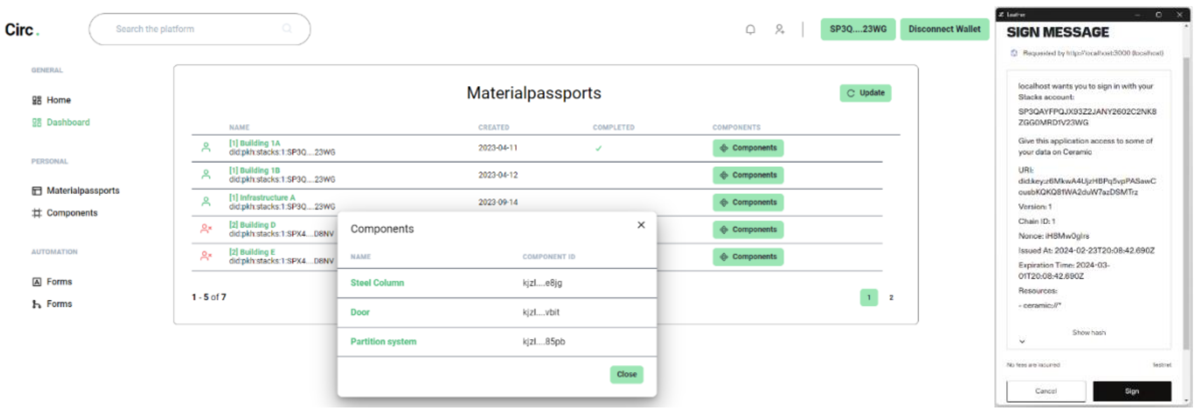


Figure 4: Ceramic-based implementation with four technology stack layers showing the workflow of uploading and downloading data.

The next step involves uploading the user's data, returning the specific stream where their data is stored, and distributing it by the storage network. In addition, the computational layer establishes a predetermined data schema, ensuring the integrity and conformity of the input data to the specified schema. Figure 5a illustrates the implementation of the prototype. Components, with most of the information contained in them, can be linked to a specific material passport.



(a) *(b)*

Figure 5: (a) User Interface (UI) of the second implementation based on Ceramic. (b) Generation of a decentralized identity (DID) by signing with a wallet.

The implementation of such an architectural framework leads to improved governance of data input and access and a refined ownership paradigm. This is done by creating the structure mentioned above, called a stream, with the application's DID as the proprietary identifier. As users enter data into the stream, their unique DID is registered simultaneously. This allows the decentralized application, acting as a trusted entity, to accurately determine access permissions. The authorization process can then be handled using a token-based or a DID-centric approach (address-based) (Hunhevicz *et al.*, 2023).

In this system, enhancements are evident for the data provenance mechanisms. Leveraging such decentralized data networks ensures that the source of data can be traced, bolstered by the trustless attribution of the system. Two primary factors contribute to this. First, data ownership and subsequent modifications can be attributed to a specific entity, and second, the characteristics of these networks ensure distributed data. As a result, information is accessible from multiple nodes rather than being confined to a single network location, and only authorized entities have access to it.

5.3 Typology 3: Privacy-Centric Decentralized Data Networks

The next implementation, which represents the third typology (see Section 4.3) focuses on the methods available for managing data storage locations. Previous implementations have shown that analyzing data ownership attribution alone, without considering the storage location in specific use cases, is insufficient. Although previous solutions allow for decentralized storage without central control over the data, they do not provide the ability for originators to specify the exact physical location of their data storage. The Gaia network, used for the third typology in this example, provides functionalities to accurately determine the data storage location within a network (see Appendix A.5 Gaia).

This is especially critical in use cases where sensitive data needs to be processed in accordance with the latest data protection regulations. Furthermore, this enables the continued use of legacy storage systems (see storage layer in Figure 6), simplifying the integration of current systems. In practice, this means that existing storage systems can still be used. However, a computational and logical layer is added to enhance functionality, similar to previous approaches.

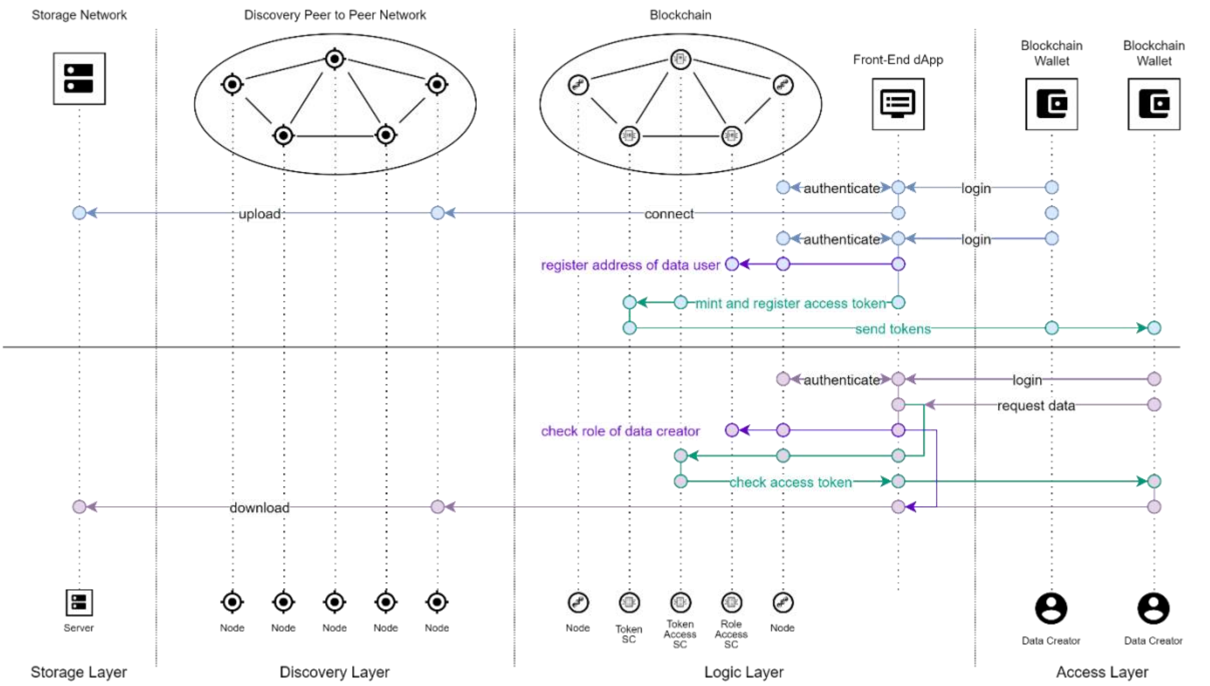


Figure 6: Gaia-based implementation with four technology stack layers showing the workflow of uploading and downloading data.

As shown in Figure 6, it is possible to achieve independence from the access mechanisms of the network, enabling the customization of the access methodologies (Hunhevicz *et al.*, 2023). This includes the possibility of granting access to stored data through token-based or address-based mechanisms. More details can be found in the relevant research papers. By precisely determining the storage location, use cases can be addressed that require streamlined ownership management, while also necessitating custodial characteristics. This is facilitated by making the information publicly accessible via an URL, but fundamentally encrypting the data, in contrast to IPFS, for example. Decryption then occurs only through a specific private key of the creator or the responsible decentralized application.

However, there are certain limitations to the control of the data formats. Unlike the second prototype, this system does not allow full control or global standardization of data formats. This can pose challenges, especially when aggregating data from multiple sources, as it may lead to the development of data silos, albeit decentralized ones. Furthermore, the data is still stored on centralized servers, which means that redundancy and the risk of a single point of failure remain.

5.4 Summary

The previous section shows the impact of the three tested decentralized data networks for lifecycle data management in the built environment. The three solutions were assessed for their ability to provide a system for transparent data and peer-to-peer exchange that operates autonomously from centralized entities. All three implementations meet this requirement by providing a system for non-custodial data storage. However, testing of these implementations has revealed differences in capabilities across networks. For example, the first network IPFS offers limited mechanisms for information storage and governance. The third network, Gaia, provides more advanced mechanisms, where the control of the storage location allows for easier implementation of access control mechanisms. Traceability capabilities also differ, with Gaia and the third network, Ceramic, providing partial support for provenance mechanisms. However, IPFS requires the integration of additional features, such as timestamping, to enhance this functionality. In conclusion, despite the initial classification based on the Table 1 showing differences in data sovereignty, data mutability, and universal addressing, the evaluation of their implementation for a specific use case has highlighted two capabilities that are particularly important for lifecycle data management applications. First, data access mechanisms are critical to enable data sharing despite its distributed nature. Second, the ability to control where data is stored is critical, especially when managing sensitive information that requires strict security and privacy measures.

These insights indicate that a decision framework would serve as an essential tool for researchers and practitioners in the field. This is especially relevant as the use of such systems, particularly in the early stages, raises complex questions about the interaction between chosen technology and specific requirements. Therefore, in the following section, a framework is presented designed to assist in navigating the diverse technological landscape (see Section 6).

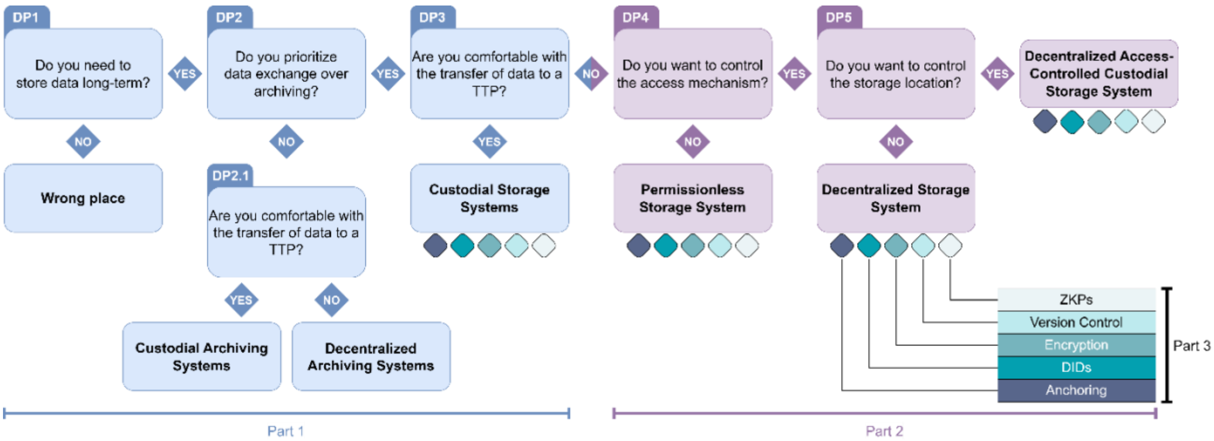


Figure 7: Decision Framework consisting of three parts with five guiding Decision Points (DPs).

6. A DECISION FRAMEWORK FOR PRACTITIONERS USING DDNS

Practitioners in the built environment need to select the most appropriate type of decentralized data network for a use case (see Figure 1, SQ2). The following decision framework is developed based on the insights developed in the previous two sections. The aim is to allow stakeholders to select a decentralized data network that aligns with the requirements of a specific use case of lifecycle data management in the built environment and is easier to navigate than trying to derive a suitable network based on functional typologies, as shown in Table 1.

The proposed decision framework shown in Figure 7 consists of three parts in which navigation through the various stages is facilitated by the so-called decision points (DP). The following sections further elaborate on each DP, structured according to the three parts. Part 1 determines whether the use of a decentralized data network is necessary (see Section 6.1); Part 2 evaluates which of the three typologies is the most suitable (see Section 6.2); and Part 3 discusses additional functionalities that can be added if necessary (see Section 6.3).

6.1 Part 1: Do you need a decentralized data network?

6.1.1 DP1: Do you need to store data long-term?

Step 1 evaluates whether data storage is necessary at all (see Figure 7 DP1). Typically, decentralized data networks are designed to support data storage for specific use cases. However, applications such as real-time monitoring or sensor stream processing often do not require long-term storage capabilities.

6.1.2 DP2: Do you prioritize data exchange over archiving?

Step 2 evaluates the infrastructure requirements used for the specific use case in terms of data exchange or archiving (see Figure 7 DP2). It is essential to note that this evaluation encompasses not only the actual exchange of information but also the requirement for storing data in the same system. This assessment essentially determines whether the system requirement prioritizes information exchange over long-term archiving. If the focus is on the latter, the subquestion of whether data transfer to a trusted third party (TTP) is feasible should be considered (see Figure 7 DP2.1). Two potential outcomes emerge: a custodial archiving system, which may employ conventional cloud-based or local solutions or decentralized archiving systems.

Custodial Archiving Systems: The first category of systems specializes in secure storage and preservation of data for long periods. These systems ensure data integrity and availability, providing a stable and reliable repository for historical data and records. This is achieved through custodial archiving systems based on centralized server systems, which are the most widely used solution today.

Decentralized Archiving Systems: A second type of system provides the aforementioned functionalities through a decentralized architecture. This approach can overcome the limitations of centralized systems, such as the involvement of a middleman and loss of data control. Arweave is an example of this latter type, with further details in Appendix A.8 Arweave.

6.1.3 DP3: Are you comfortable with the transfer of data to a trusted third party (TTP)?

Step 3 evaluates whether data can be exchanged with the use of a trusted third party (TTP) (see Figure 7 DP3). This refers to a system's component that can act as an intermediary to guarantee secure and efficient data exchange. Although it retains direct control over the data, it also provides services that reinforce trust in the network itself. Custodial storage systems, including cloud-based solutions, can be used if the presence of a TTP is considered appropriate. Otherwise, a shift to the second part of the framework, which focuses on decentralized systems, is made.

Custodial Storage Systems: These systems are data storage solutions that serve as centralized platforms for managing and storing data for users while providing ease of use, robust security, and reliable access. They come in a variety of forms, including relational and object-oriented databases.

6.2 Part 2: What decentralized data network do you need?

6.2.1 DP4: Do you want to control the access mechanism?

Step 4 evaluates the need for access control mechanisms. As indicated earlier in this section, this is one of the two critical factors essential for lifecycle data management requirements. With the advent of decentralized systems, the network inherently offers this capability through the logic included in its codebase. Here, there are two possibilities for distinguishing between permissionless and permissioned solutions. The former embodies fundamental openness, permitting unrestricted publication of data, even in relation to specific data models. However, this transparency implies that the location of the data is easily identifiable, making it readily accessible and challenging to monetize.

Permissionless Storage Systems: Permissionless storage systems are networks that allow users to store, access, and manage data without requiring permission from a central authority or being restricted by specific network rules.

The second option of permissioned access requires an authentication process to check if a particular user has access to the system. The access method is organized through identity management. Therefore, proper identification, particularly facilitated by a blockchain-based component, is required to use an account that is manifested within the system as a wallet or a unique private key. A more detailed discussion of identity management can be found in Section 6.3.

6.2.2 DP5: Do you want to control the storage location?

Step 5 ultimately assesses the need for storage location control. After DP4, a permissioned system is in place where authentication and identity management form the basis. The current focus is on the system's capability to pinpoint and determine the specific storage location, such as a server or node. If this is not confirmed, indicating that control over storage location is not required, then the system falls into the category of decentralized storage systems. These systems, as demonstrated in the second implementation in Section 5.2, operate fully decentralized. All data-related actions, including access, storage, and exchange, are automated without any intermediaries.

Decentralized Storage Systems: These systems allow decentralized storage and access without a central entity, relying solely on the encoded rules of the network. However, they do not allow for the definition of storage location, so information is distributed throughout the system according to its technical architecture.

If DP5 is answered positively, the system is categorized as a decentralized, access-controlled custodial storage system. They enable migration from conventional legacy systems by utilizing the existing centralized storage infrastructure but still preserve benefits such as trusted data sharing because access and sharing are decentralized.

Decentralized Access Controlled Storage System: These systems are very similar to decentralized storage systems, with distinct differences in the exact storage location that can be determined by the user. While this has the advantages mentioned above, it also has the drawback of reducing the decentralization of information.

6.3 Part 3: Additional Functionalities for Improved Performance, Security, and Usability

After concluding the preliminary assessment of the need for decentralized data networks and comparing the available solutions, it is possible to move on to the final part of the decision framework. This stage entails the examination of additional features that can improve the functionality and user experience of the selected network, independent of the chosen network.

Anchoring uses blockchain technology to establish trust by improving information traceability. The method generates a cryptographic hash of the data, embedding it into a blockchain transaction. This approach assigns a definitive timestamp to the hash, ensuring authenticity. It is not possible to modify the dataset without detection as any changes will create a different hash value from the original one stored.

Decentralized Identifiers (DIDs) support a globally verifiable and decentralized digital identity and can establish identification for any entity, be it a person, organization, or abstract entity. The creation and management of a DID does not depend on a central authority but is organized in a decentralized manner. This approach is typically implemented using a blockchain component, more specifically through a blockchain address or account verification. Different blockchains have the ability to create such DIDs.

Encryption is an essential part of protecting the confidentiality of data. Cryptographic algorithms help guarantee the privacy and integrity of data from unauthorized access or eavesdropping. Data safety becomes even more critical when working with sensitive information due to its storage being distributed across multiple locations. Although some solutions may have a basic level of encryption, the data is still unencrypted when stored in a hidden location away from public view. However, this basic level of encryption may not be sufficient to meet the demands of enhanced security. Consequently, it is crucial to integrate advanced encryption protocols into these solutions to strengthen confidentiality and protect sensitive data from exposure.

Version control is a feature that provides the infrastructure to monitor and archive each iteration of a data set. Preserving and retrieving historical data states enables auditing of changes and restoring previous states, which is especially useful in situations where traceability of data is crucial. When implementing a use case, assessing whether a solution can support data versioning practices is essential. Equally important is determining whether additional version control systems are needed to maintain the required data reliability standards.

Zero-Knowledge Proofs (ZKPs) offer an advanced cryptographic technique that improves data interaction's privacy and security within decentralized data networks. They verify the presence or accuracy of data against prespecified standards while keeping the information confidential. In particular, ZKPs significantly benefit scenarios that require the validity of the data to be checked before transactions can be initiated. For example, when participants need to prove that they possess specific data or comply with certain criteria, ZKPs can provide evidence without revealing the actual data. This feature improves the reliability of exchanges within the network, which subsequently promotes data integrity and establishes a trusted ecosystem.

7. DISCUSSION

This section discusses the contribution and limitations of the work structured according to this paper's three previous sections 4, 5, and 6. In addition, an outlook is provided on potential further applications of decentralized data networks in the built environment.

7.1 Overview and Classification of Decentralized Data Networks

In the first part of the paper (see Section 4), the limited research described in section 2 in this domain is addressed. This is done by outlining the classification of decentralized data networks. Although all networks share similar high-level promises and aim to enable data integration and advance data availability (Zahed Benisi, Aminian and Javadi, 2020), they exhibit distinct technical differences that must be considered for implementation.

The three identified typologies are based on perspectives of data sovereignty, data mutability, and universal addressing. This nuanced classification is a first attempt to categorize these networks in a systematic way that will aid in the selection of a decentralized data network for specific system implementations. Upon examination of the three current research approaches - on-chain, centralized server, IPFS (see Section 2.2), it is becoming apparent that they are not technologically congruent, contributing to the complexity of the domain.

The key perspective for the typologies of the classification is the concept of data sovereignty, a feature that varies between all the networks examined. However, when considering the other perspectives of data mutability and universal addressing, networks differ in their capabilities. Some networks excel in both areas, while others specialize in one or the other. This distinction highlights the unique benefits and potential limitations of each network. The emerging focus on data sovereignty is not particularly unexpected. It is a fundamental feature of decentralized data technologies and, at the same time, represents a significant challenge in lifecycle data management (El Arass and Souissi, 2018).

Finally, it is important to recognize the dynamic nature of this novel technology landscape (Belk, Humayun and Brouard, 2022; Park *et al.*, 2023). Review and classification efforts are based on the current state of the art, acknowledging the rapid pace of innovation within this domain. Ongoing revision and refinement will be necessary to maintain the relevance of the outlined networks as the field evolves. However, it is expected that the typologies persist because they comprehensively cover the characteristics of the domain. This suggests that even as specific technologies evolve or migrate between typologies, the framework will continue to provide valuable guidance for navigating the complex landscape of decentralized data networks.

7.2 Implementation of Decentralized Data Networks in the Built Environment

Drawing on the classification presented in the first part of the paper, this research demonstrates how different decentralized data networks can be applied to a use case in the built environment. This demonstration makes a contribution to research and practice that goes beyond current applications, which are generally limited to systems implemented with IPFS (Balduf *et al.*, 2022).

The three distinct implementations, based on the three different typologies, confirm the expected limitations of a theoretical classification. The identified differences are significant and affect both the technical execution and their respective contributions to the limitations of lifecycle data management (see Section 2.1). Using material passports as a common use case highlighted these differences through actual implementation, providing insights for the subsequent development of a decision framework.

The analysis indicates that each of the three typologies has potential advantages and disadvantages for deployment. Therefore, no single system type may be universally suitable for all use cases. More sophisticated networks that offer a wider array of functionalities, such as Ceramic, require a trade-off in terms of increased complexity in implementation and potential maintenance. Furthermore, the implemented system relies on blockchain components that require on-chain transactions and associated fees.

This study did not investigate the exact costs associated with the implementation and operation of these networks or evaluate their performance. The main objective was to outline the structural and functional differences. More research is needed to examine the differences in cost and performance.

The paper mentioned the integration of blockchain technology with decentralized data networks at the beginning (see Section 2.2). In general, it should be noted that this research has shown how the challenge of interoperable lifecycle data management in the built environment can be addressed with decentralized technologies beyond blockchain. This is because blockchain-centric approaches, in particular, are limited by a lack of fine-grained data provenance and peer-to-peer sharing without the performance bottlenecks of consensus-based ledgers (Dakhli, Lafhaj and Mossman, 2019; Bucher *et al.*, 2024). However, the implementations have shown that a wider range of functionalities can only be enabled through this combination. Nevertheless, it is incorrect to assume that they inherently belong together as a standard, and the specific technology stack used must be examined on a case-by-case basis.

Finally, it is important to emphasize that the implementations conducted were not directly applied or tested in the industry. Future research should aim to gather more empirical insights, focusing on usability and impacts on the industry, to enrich the foundational knowledge established by this study.

7.3 Decentralized Data Network Decision Framework

The resulting decision framework consolidates all findings into an accessible tool for future research and industry applications. It is a foundation for comprehensive engagement with emerging technologies, facilitating informed experimentation in various contexts. Based on the analyses of the preceding sections of the paper, the framework synthesizes these insights, offering a structured approach to navigate the complexities of decentralized data networks. However, it is important to recognize that the framework is dynamic and should be adapted in response to changes.

The decision framework can provide significant assistance not by proposing particular solutions but by illustrating how the requirements of a use case can align with the best system type. It has a generic approach, which implies the potential for further elaboration, such as transforming it into a more detailed flowchart that provides deeper insights. However, it exclusively focuses on decentralized data networks and does not offer a comparative analysis with existing technologies and approaches described in section 2.1. Future research should explore the integration of such networks with other technologies used in the built environment.

7.4 Future Outlook on Decentralized Data Networks in the Built Environment

Based on the initial experimentation with decentralized data networks, insights into how such networks could support the challenges of lifecycle data management are presented, particularly in the context of a circular economy. Decentralized data networks can support advanced applications that require the decentralization of information and access mechanisms independent of central entities (Wang, Zhang and Zhang, 2018). Examples of this include incentive structures as well as more complex applications such as decentralized data marketplaces (F. Bucher and M. Hall, 2022). This concept relies on decentralized storage as the core component and includes additional features such as data exchange, interoperability, and evaluation.

Such a platform operates without a central authority that regulates the market participants, creating a trustless and autonomous environment. Data creators maintain complete control over their data, as there is no centralized data repository (Ramachandran, Radhakrishnan and Krishnamachari, 2018). Data can be exchanged for monetary value, and embedded mechanisms allow for incentivizing data sharing and verification. As described above, this is likely to be facilitated by the integration of additional blockchain-based components. Additionally, participants can retain decentralized control over their monetary assets while implementing governance mechanisms to ensure data quality and integrity and to penalize dishonest practices (Wang *et al.*, 2019).

This initial understanding of decentralized data networks in the built environment is just the beginning. These systems require further exploration, testing with stakeholders, and verification. Although this Web3-based technology, primarily supported by decentralization and trustless interrelation mechanisms, offers a new approach to interactions, it raises a fundamental question: is this technology useful in the built environment where actors are often known or contractually linked, potentially negating the need for a trustless system? On the other hand, it could be argued that these technologies more accurately reflect the characteristics of the built environment (Dubois and Gadde, 2002; Hunhevicz, Dounas and Hall, 2022), as data creation occurs in a decentralized manner with various participants contributing to an integral information model throughout the lifecycle of an asset. Such open questions need to be addressed in future contributions within this research stream.

Specifically, the development of decentralized data technologies is progressing slowly but steadily. Building on the foundations presented in this paper, future research should work on iterative prototypes, conduct pilot studies in real-world environments, and carry out evaluations in close coordination with industry. These practical steps are essential to validate the contributions and ensure their adaptability to the industrial environment. However, due to the rapid pace of technological development, researchers should be aware that it will not always be possible to use the latest solutions. For the industry, these rapid advances require a focus on flexibility and modularity. Rather than demanding the immediate implementation of the latest features, a more optimal approach would be to design a technological landscape that allows the seamless integration of new components or functionalities with minimal effort.

8. CONCLUSION

In response to the overarching research question about the potential of decentralized data networks for lifecycle data management in the built environment, this paper presents a comprehensive overview and classification, exploring their characteristics and integration for implemented systems. First, it provides the reader with a detailed understanding of the different available networks, highlighting their intrinsic properties and supporting capabilities for lifecycle data management. This includes an identification of three typologies of decentralized data networks relevant to lifecycle data management: the immutable typology, the comprehensive typology, and the privacy-centric typology, as well as an exploration of their implementation for lifecycle data management. This allows for a first understanding of the types of decentralized data networks and partially addresses the first subquestion relating to the assistance of these networks for lifecycle data management. In addition, the decision framework developed from the second subquestion is an attempt to standardize functional typologies and should help future research select a decentralized data network for a use case and discuss more nuanced options with respect to other available options. In doing so, this paper is one of the first to explore how these networks have the potential to transform ownership management and strengthen data sovereignty for data creators in the built environment.

The research is significant because it demonstrates that decentralized systems have the potential to transform lifecycle data management practices, primarily by enabling novel methods for storage and access, and hence also ownership models. An interesting key insight is that access at the network level can be achieved without additional blockchain-based components, and data management can be made more decentralized without the use of blockchain. However, it is important to note that advanced ownership models and authentication capabilities, such as decentralized identifiers (DIDs), are fundamentally improved by blockchain technology. In the end, a combination of both decentralized data networks and blockchain is probably most likely to fulfill more advanced use case requirements of lifecycle data management in the built environment.

These characteristics are particularly relevant in applications such as the circular economy, where the exchange of verifiable information that is maintained by an owner over the long term is a key requirement. In addition, the contribution establishes the groundwork by highlighting additional opportunities for a research field that involves the next generation of decentralized data storage and governance in the built environment.

The limitations of the study are related to the evolving landscape of decentralized technologies and the complexity of integrating them into existing data management practices. It is important to note that a holistic consideration of these factors poses a challenge that extends beyond the scope of this work. Furthermore, more research is needed to test and validate the proposed decision framework, verify the potential described, and refine the integration process with existing approaches.

For the built environment, the research is relevant beyond academic discourse and provides tangible utility to practitioners since it offers a new approach to managing lifecycle data, reducing current limitations induced by the centralization of data. This is because decentralized data networks have the potential to significantly improve data storage, enable transparent interactions, and foster collaboration, resulting in more efficient and effective lifecycle data management. In addition, this study adds to the growing scholarship suggesting the need to move away from a centralizing paradigm and towards a decentralized architecture that ensures persistent data accessibility and governance by its creators. The results provide actionable insights and guidance as to how decentralized data networks can emerge as a new infrastructure or backbone for future data management strategies.

REFERENCES

3Box Labs (2022) Ceramic - The Composable Data Network. Available at: https://ceramic.network/ (Accessed: 28 July 2023).

Ahonen, T. et al. (2022) 'Long-Term History Data and Integrated Expert Knowledge in Asset Lifecycle Management', in J. Machado et al. (eds) Innovations in Industrial Engineering. Cham: Springer International Publishing (Lecture Notes in Mechanical Engineering), pp. 1–11. Available at: https://doi.org/10.1007/978-3-030-78170-5_1.

Balduf, L. et al. (2022) 'Monitoring Data Requests in Decentralized Data Storage Systems: A Case Study of IPFS', in 2022 IEEE 42nd International Conference on Distributed Computing Systems (ICDCS). 2022 IEEE 42nd International Conference on Distributed Computing Systems (ICDCS), Bologna, Italy: IEEE, pp. 658–668. Available at: https://doi.org/10.1109/ICDCS54860.2022.00069.

Bambacht, J. and Pouwelse, J. (2022) 'Web3: A Decentralized Societal Infrastructure for Identity, Trust, Money, and Data'. arXiv. Available at: https://doi.org/10.48550/ARXIV.2203.00398.

Belk, R., Humayun, M. and Brouard, M. (2022) 'Money, possessions, and ownership in the Metaverse: NFTs, cryptocurrencies, Web3 and Wild Markets', Journal of Business Research, 153, pp. 198–205. Available at: https://doi.org/10.1016/j.jbusres.2022.08.031.

Bellini, A. and Bang, S. (2022) 'Barriers for data management as an enabler of circular economy: an exploratory study of the Norwegian AEC-industry', IOP Conference Series: Earth and Environmental Science, 1122(1), p. 012047. Available at: https://doi.org/10.1088/1755-1315/1122/1/012047.

Benet, J. (2014) 'IPFS - Content Addressed, Versioned, P2P File System'. arXiv. Available at: http://arxiv.org/abs/1407.3561 (Accessed: 16 April 2024).

Berg, J.B. et al. (2021) 'Reconfiguring the construction value chain: analysing key sources of friction in the business model archetypes of AEC companies in strategic partnerships', Construction Management and Economics, 39(6), pp. 533–548. Available at: https://doi.org/10.1080/01446193.2021.1925134.

Berglund-Brown, J. et al. (2022) 'Assessing Circular Information Flow in Industrialized Construction: a framework for evaluating data for circular construction', IOP Conference Series: Earth and Environmental Science, 1078(1), p. 012005. Available at: https://doi.org/10.1088/1755-1315/1078/1/012005.

Böckel, A., Nuzum, A.-K. and Weissbrod, I. (2021) 'Blockchain for the Circular Economy: Analysis of the Research-Practice Gap', Sustainable Production and Consumption, 25, pp. 525–539. Available at: https://doi.org/10.1016/j.spc.2020.12.006.

Bocken, N. et al. (2021) Circularity as the new normal: Future fitting Swiss business strategies. Switzerland: PwC Switzerland.

Bocken, N.M.P. et al. (2016) 'Product design and business model strategies for a circular economy', Journal of Industrial and Production Engineering, 33(5), pp. 308–320. Available at: https://doi.org/10.1080/21681015.2016.1172124.

Boeva, Y., Braun, K. and Kropp, C. (2024) 'Platformization in the built environment: the political techno-economy of Building Information Modeling', Science as Culture, 33(2), pp. 146–173. Available at: https://doi.org/10.1080/09505431.2023.2237042.

Booth, A. (2016) 'Searching for qualitative research for inclusion in systematic reviews: a structured methodological review', Systematic Reviews, 5(1), p. 74. Available at: https://doi.org/10.1186/s13643-016-0249-x.

Borrmann, A. et al. (eds) (2018a) Building Information Modeling: Technology Foundations and Industry Practice. Cham: Springer International Publishing. Available at: https://doi.org/10.1007/978-3-319-92862-3.

Borrmann, A. et al. (2018b) 'Building Information Modeling: Why? What? How?', in A. Borrmann et al. (eds) Building Information Modeling. Cham: Springer International Publishing, pp. 1–24. Available at: https://doi.org/10.1007/978-3-319-92862-3_1.

Bucher, D.F. et al. (2024) 'From BIM to Web3: A critical interpretive synthesis of present and emerging data management approaches in construction informatics', Advanced Engineering Informatics, 62, p. 102884. Available at: https://doi.org/10.1016/j.aei.2024.102884.

Byers, B.S. et al. (2024) 'From research to practice: A review on technologies for addressing the information gap for building material reuse in circular construction', Sustainable Production and Consumption, 45, pp. 177–191. Available at: https://doi.org/10.1016/j.spc.2023.12.017.

Caldas, L.R. et al. (2022) 'How Different Tools Contribute to Climate Change Mitigation in a Circular Building Environment?—A Systematic Literature Review', Sustainability, 14(7), p. 3759. Available at: https://doi.org/10.3390/su14073759.

Cena, F., Farzan, R. and Lops, P. (2009) 'Web 3.0: merging semantic web with social web', in Proceedings of the 20th ACM conference on Hypertext and hypermedia. HT '09: 20th ACM Conference on Hypertext and Hypermedia, Torino Italy: ACM, pp. 385–386. Available at: https://doi.org/10.1145/1557914.1558002.

Corry, E. et al. (2014) 'Using semantic web technologies to access soft AEC data', Advanced Engineering Informatics, 28(4), pp. 370–380. Available at: https://doi.org/10.1016/j.aei.2014.05.002.

Cucko, S. and Turkanovic, M. (2021) 'Decentralized and Self-Sovereign Identity: Systematic Mapping Study', IEEE Access, 9, pp. 139009–139027. Available at: https://doi.org/10.1109/ACCESS.2021.3117588.

Dakhli, Z., Lafhaj, Z. and Mossman, A. (2019) 'The Potential of Blockchain in Building Construction', Buildings, 9(4), p. 77. Available at: https://doi.org/10.3390/buildings9040077.

Daniel, E. and Tschorsch, F. (2022) 'IPFS and Friends: A Qualitative Comparison of Next Generation Peer-to-Peer Data Networks', IEEE Communications Surveys & Tutorials, 24(1), pp. 31–52. Available at: https://doi.org/10.1109/COMST.2022.3143147.

Dantas, T.E.T. et al. (2021) 'How the combination of Circular Economy and Industry 4.0 can contribute towards achieving the Sustainable Development Goals', Sustainable Production and Consumption, 26, pp. 213–227. Available at: https://doi.org/10.1016/j.spc.2020.10.005.

Das, M., Tao, X. and Cheng, J.C.P. (2021) 'A Secure and Distributed Construction Document Management System Using Blockchain', in E. Toledo Santos and S. Scheer (eds) Proceedings of the 18th International Conference on Computing in Civil and Building Engineering. Cham: Springer International Publishing (Lecture Notes in Civil Engineering), pp. 850–862. Available at: https://doi.org/10.1007/978-3-030-51295-8_59.

De Wolf, C., Çetin, S. and Bocken, N.M.P. (eds) (2024) A Circular Built Environment in the Digital Age. Cham: Springer International Publishing (Circular Economy and Sustainability). Available at: https://doi.org/10.1007/978-3-031-39675-5.

De Wolf, C., Pomponi, F. and Moncaster, A. (2017) 'Measuring embodied carbon dioxide equivalent of buildings: A review and critique of current industry practice', Energy and Buildings, 140, pp. 68–80. Available at: https://doi.org/10.1016/j.enbuild.2017.01.075.

Demirkesen, S. and Tezel, A. (2021) 'Investigating major challenges for industry 4.0 adoption among construction companies', Engineering, Construction and Architectural Management [Preprint]. Available at: https://doi.org/10.1108/ECAM-12-2020-1059.

Dietz, M., Putz, B. and Pernul, G. (2019) 'A Distributed Ledger Approach to Digital Twin Secure Data Sharing', in S.N. Foley (ed.) Data and Applications Security and Privacy XXXIII. Cham: Springer International Publishing (Lecture Notes in Computer Science), pp. 281–300. Available at: https://doi.org/10.1007/978-3-030-22479-0_15.

Dounas, T., Jabi, W. and Lombardi, D. (2020) 'Smart Contracts for Decentralised Building Information Modelling', in Proceedings of 38th Education and research in computer aided architectural design in Europe (eCAADe) conference 2020: anthropologic: architecture and fabrication in the cognitive age. eCAADe 2020: Anthropologic : Architecture and Fabrication in the cognitive age, Berlin, Germany, pp. 565–574. Available at: https://doi.org/10.52842/conf.ecaade.2020.2.565.

Dubois, A. and Gadde, L.-E. (2002) 'The construction industry as a loosely coupled system: implications for productivity and innovation', Construction Management and Economics, 20(7), pp. 621–631. Available at: https://doi.org/10.1080/01446190210163543.

El Arass, M. and Souissi, N. (2018) 'Data Lifecycle: From Big Data to SmartData', in 2018 IEEE 5th International Congress on Information Science and Technology (CiSt). 2018 IEEE 5th International Congress on Information Science and Technology (CiSt), Marrakech: IEEE, pp. 80–87. Available at: https://doi.org/10.1109/CIST.2018.8596547.

Elghaish, F., Abrishami, S. and Hosseini, M.R. (2020) 'Integrated project delivery with blockchain: An automated financial system', Automation in Construction, 114, p. 103182. Available at: https://doi.org/10.1016/j.autcon.2020.103182.

F. Bucher, D. and M. Hall, D. (2022) 'New Ways of Data Governance for Construction? Decentralized Data Marketplaces as Web3 Concept just around the Corner', in Proceedings of the 29th EG-ICE International Workshop on Intelligent Computing in Engineering. The 29th EG-ICE International Workshop on Intelligent Computing in Engineering, EG-ICE, pp. 339–349. Available at: https://doi.org/10.7146/aul.455.c224.

Gan, W. et al. (2023) 'Web 3.0: The Future of Internet', in Companion Proceedings of the ACM Web Conference 2023. WWW '23: The ACM Web Conference 2023, Austin TX USA: ACM, pp. 1266–1275. Available at: https://doi.org/10.1145/3543873.3587583.

Gao, H. and Zhong, B. (2022) 'A blockchain-based framework for supporting BIM-based building code compliance checking workflow', IOP Conference Series: Materials Science and Engineering, 1218(1), p. 012016. Available at: https://doi.org/10.1088/1757-899X/1218/1/012016.

Ginga, C.P., Ongpeng, J.M.C. and Daly, Ma.K.M. (2020) 'Circular Economy on Construction and Demolition Waste: A Literature Review on Material Recovery and Production', Materials, 13(13), p. 2970. Available at: https://doi.org/10.3390/ma13132970.

Guan, C. et al. (2023) 'An ecosystem approach to Web3.0: a systematic review and research agenda', Journal of Electronic Business & Digital Economics, 2(1), pp. 139–156. Available at: https://doi.org/10.1108/JEBDE-10-2022-0039.

Guidi, B., Michienzi, A. and Ricci, L. (2022) 'Evaluating the decentralisation of filecoin', in Proceedings of the 3rd International Workshop on Distributed Infrastructure for the Common Good. Middleware '22: 23rd International Middleware Conference, Quebec Quebec City Canada: ACM, pp. 13–18. Available at: https://doi.org/10.1145/3565383.3566108.

Hackl, C., Lueth, D. and Bartolo, T.D. (2022) Navigating the Metaverse: A Guide to Limitless Possibilities in a Web 3.0 World. John Wiley & Sons.

Haider, A. (2015) 'Asset Lifecycle Data Governance Framework', in W.B. Lee et al. (eds) Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012). Cham: Springer International Publishing (Lecture Notes in Mechanical Engineering), pp. 287–296. Available at: https://doi.org/10.1007/978-3-319-06966-1_27.

Hall, D. (2018) 'Cracks in the mirror: Conceptualizing the ongoing AEC industry re-organization', in Working Paper Proceedings 16th Engineering Project Organization Conference. Engineering Project Organization Society (EPOS), pp. 458–477. Available at: https://doi.org/10.3929/ETHZ-B-000275328.

Halttula, H. (2020) Enhancing data utilisation in the construction project lifecycle through early involvement and integration. Academic dissertation. University of Oulu. Available at: http://jultika.oulu.fi/files/isbn9789526226156.pdf (Accessed: 19 January 2024).

Halttula, H., Haapasalo, H. and Silvola, R. (2020) ‘Managing data flows in infrastructure projects - the lifecycle process model’, Journal of Information Technology in Construction, 25, pp. 193–211. Available at: https://doi.org/10.36680/j.itcon.2020.012.

Hariembrundtland, G. (1985) ‘World Commission on environment and development’, Environmental Policy and Law, 14(1), pp. 26–30. Available at: https://doi.org/10.1016/S0378-777X(85)80040-8.

He, L. (2023) ‘A Comparative Examination of Network and Contract-Based Blockchain Storage Solutions for Decentralized Applications’, in C. Chen et al. (eds) Proceedings of the 3rd International Conference on Digital Economy and Computer Application (DECA 2023). Dordrecht: Atlantis Press International BV (Atlantis Highlights in Computer Sciences), pp. 133–145. Available at: https://doi.org/10.2991/978-94-6463-304-7_16.

Heisel, F. and Rau-Oberhuber, S. (2020) ‘Calculation and evaluation of circularity indicators for the built environment using the case studies of UMAR and Madaster’, Journal of Cleaner Production, 243, p. 118482. Available at: https://doi.org/10.1016/j.jclepro.2019.118482.

Honic, M. and De Wolf, C. (2023) ‘Applying EU Level(s) framework indicators to improve circularity: A case study’, IOP Conference Series: Earth and Environmental Science, 1196(1), p. 012041. Available at: https://doi.org/10.1088/1755-1315/1196/1/012041.

Honic, M, Kovacic, I. and Rechberger, H. (2019) ‘Concept for a BIM-based Material Passport for buildings’, IOP Conference Series: Earth and Environmental Science, 225, p. 012073. Available at: https://doi.org/10.1088/1755-1315/225/1/012073.

Honic, Meliha, Kovacic, I. and Rechberger, H. (2019) ‘Improving the recycling potential of buildings through Material Passports (MP): An Austrian case study’, Journal of Cleaner Production, 217, pp. 787–797. Available at: https://doi.org/10.1016/j.jclepro.2019.01.212.

Hu, W. (2008) ‘Information Lifecycle Modeling Framework for Construction Project Lifecycle Management’, in 2008 International Seminar on Future Information Technology and Management Engineering. 2008 International Seminar on Future Information Technology and Management Engineering (FITME), Leicestershire, United Kingdom: IEEE, pp. 372–375. Available at: https://doi.org/10.1109/FITME.2008.142.

Hunhevicz, J., Dounas, T. and Hall, D.M. (2022) ‘The Promise of Blockchain for the Construction Industry: A Governance Lens’, in T. Dounas and D. Lombardi (eds) Blockchain for Construction. Singapore: Springer Nature (Blockchain Technologies), pp. 5–33. Available at: https://doi.org/10.1007/978-981-19-3759-0_2.

Hunhevicz, J.J. et al. (2023) ‘Web3-based role and token data access: the case of building material passports’, in Proceedings of the 2023 European Conference on Computing in Construction and the 40th International CIB W78 Conference. 2023 European Conference on Computing in Construction and the 40th International CIB W78 Conference, p. 8. Available at: https://doi.org/10.35490/EC3.2023.217.

Hunhevicz, J.J., Motie, M. and Hall, D.M. (2022) ‘Digital building twins and blockchain for performance-based (smart) contracts’, Automation in Construction, 133, p. 103981. Available at: https://doi.org/10.1016/j.autcon.2021.103981.

Jaiman, V., Pernice, L. and Urovi, V. (2022) ‘User incentives for blockchain-based data sharing platforms’, PLOS ONE. Edited by B. Guidi, 17(4), p. e0266624. Available at: https://doi.org/10.1371/journal.pone.0266624.

Jang, K. et al. (2021) ‘Infrastructure BIM Platform for Lifecycle Management’, Applied Sciences, 11(21), p. 10310. Available at: https://doi.org/10.3390/app112110310.

Jiao, Y. et al. (2013) ‘A cloud approach to unified lifecycle data management in architecture, engineering, construction and facilities management: Integrating BIMs and SNS’, Advanced Engineering Informatics, 27(2), pp. 173–188. Available at: https://doi.org/10.1016/j.aei.2012.11.006.

Kirchherr, J., Reike, D. and Hekkert, M. (2017) ‘Conceptualizing the circular economy: An analysis of 114 definitions’, Resources, Conservation and Recycling, 127, pp. 221–232. Available at: https://doi.org/10.1016/j.resconrec.2017.09.005.

Klinc, R., Dolenc, M. and Turk, Z. (2009) ‘Possible benefits of Web 2.0 to Construction Industry’, in 25th W78 Conference “Improving the management of construction projects through IT adoption“. International Conference on Information Technology in Construction, Santiago, Chile: Universidad de Talca, Chile, p. 9. Available at: https://www.academia.edu/download/78672529/w78-2008-2-06.pdf (Accessed: 22 March 2024).

Konietzko, J., Bocken, N. and Hultink, E.J. (2020) ‘A Tool to Analyze, Ideate and Develop Circular Innovation Ecosystems’, Sustainability, 12(1), p. 417. Available at: https://doi.org/10.3390/su12010417.

Li, H. et al. (2023) ‘An Empirical Study of Storj DCS: Ecosystem, Performance, and Security’, in 2023 IEEE/ACM 31st International Symposium on Quality of Service (IWQoS). 2023 IEEE/ACM 31st International Symposium on Quality of Service (IWQoS), Orlando, FL, USA: IEEE, pp. 1–10. Available at: https://doi.org/10.1109/IWQoS57198.2023.10188691.

Lin, Y.-C. (2013) ‘Construction network-based interface management system’, Automation in Construction, 30, pp. 228–241. Available at: https://doi.org/10.1016/j.autcon.2012.11.025.

Mêda, P., Sousa, H. and Hjelseth, E. (2020) ‘Data Templates—Product Information Management Across Project Life-Cycle’, in J.M.P.Q. Delgado (ed.) Sustainable Materials in Building Construction. Cham: Springer International Publishing (Building Pathology and Rehabilitation), pp. 117–133. Available at: https://doi.org/10.1007/978-3-030-46800-2_5.

Mendling, J. et al. (2018) ‘Blockchains for Business Process Management - Challenges and Opportunities’, ACM Transactions on Management Information Systems, 9(1), pp. 1–16. Available at: https://doi.org/10.1145/3183367.

Mulhall, D. et al. (2022) ‘The Product Circularity Data Sheet—A Standardized Digital Fingerprint for Circular Economy Data about Products’, Energies, 15(9), p. 3397. Available at: https://doi.org/10.3390/en15093397.

Munaro, M.R. and Tavares, S.F. (2021) ‘Materials passport’s review: challenges and opportunities toward a circular economy building sector’, Built Environment Project and Asset Management, 11(4), pp. 767–782. Available at: https://doi.org/10.1108/BEPAM-02-2020-0027.

Murray, A., Kim, D. and Combs, J. (2023) ‘The promise of a decentralized internet: What is Web3 and how can firms prepare?’, Business Horizons, 66(2), pp. 191–202. Available at: https://doi.org/10.1016/j.bushor.2022.06.002.

Naderi, H., Ly, R. and Shojaei, A. (2024) ‘From Data to Value: Introducing an NFT-Powered Framework for Data Exchange of Digital Twins in the AEC Industry’, in Construction Research Congress 2024. Construction Research Congress 2024, Des Moines, Iowa: American Society of Civil Engineers, pp. 299–308. Available at: https://doi.org/10.1061/9780784485262.031.

Nandi, S. et al. (2021) ‘Redesigning Supply Chains using Blockchain-Enabled Circular Economy and COVID-19 Experiences’, Sustainable Production and Consumption, 27, pp. 10–22. Available at: https://doi.org/10.1016/j.spc.2020.10.019.

Nath, K., Dhar, S. and Basishtha, S. (2014) ‘Web 1.0 to Web 3.0 - Evolution of the Web and its various challenges’, in 2014 International Conference on Reliability Optimization and Information Technology (ICROIT). 2014 International Conference on Optimization, Reliabilty, and Information Technology (ICROIT), Faridabad, Haryana, India: IEEE, pp. 86–89. Available at: https://doi.org/10.1109/ICROIT.2014.6798297.

Nawari, N.O. (2021) ‘Blockchain Technologies: Hyperledger Fabric in BIM Work Processes’, in E. Toledo Santos and S. Scheer (eds) Proceedings of the 18th International Conference on Computing in Civil and Building Engineering. Cham: Springer International Publishing (Lecture Notes in Civil Engineering), pp. 813–823. Available at: https://doi.org/10.1007/978-3-030-51295-8_56.

Olawumi, T.O. and Chan, D.W.M. (2020) ‘Concomitant impediments to the implementation of smart sustainable practices in the built environment’, Sustainable Production and Consumption, 21, pp. 239–251. Available at: https://doi.org/10.1016/j.spc.2019.09.001.

Oluleye, B.I., Chan, D.W.M. and Antwi-Afari, P. (2023) ‘Adopting Artificial Intelligence for enhancing the implementation of systemic circularity in the construction industry: A critical review’, Sustainable Production and Consumption, 35, pp. 509–524. Available at: https://doi.org/10.1016/j.spc.2022.12.002.

Oraskari, J., Beetz, J. and Törmä, S. (2018) ‘Federated Building Models in collaborative AEC/FM settings using IPFS’, in. International Conference on Computing in Civil and Building Engineering. Available at: https://research.aalto.fi/en/publications/federated-building-models-in-collaborative-aecfm-settings-using-i (Accessed: 16 April 2024).

Ozyilmaz, K.R. and Yurdakul, A. (2019) ‘Designing a Blockchain-Based IoT With Ethereum, Swarm, and LoRa: The Software Solution to Create High Availability With Minimal Security Risks’, IEEE Consumer Electronics Magazine, 8(2), pp. 28–34. Available at: https://doi.org/10.1109/MCE.2018.2880806.

Park, A. et al. (2023) ‘Interoperability: Our exciting and terrifying Web3 future’, Business Horizons, 66(4), pp. 529–541. Available at: https://doi.org/10.1016/j.bushor.2022.10.005.

Patacas, J., Dawood, N. and Kassem, M. (2020) ‘BIM for facilities management: A framework and a common data environment using open standards’, Automation in Construction, 120, p. 103366. Available at: https://doi.org/10.1016/j.autcon.2020.103366.

Pauwels, P., Costin, A. and Rasmussen, M.H. (2022) ‘Knowledge Graphs and Linked Data for the Built Environment’, in M. Bolpagni, R. Gavina, and D. Ribeiro (eds) Industry 4.0 for the Built Environment: Methodologies, Technologies and Skills. Cham: Springer International Publishing (Structural Integrity), pp. 157–183. Available at: https://doi.org/10.1007/978-3-030-82430-3_7.

Pauwels, P. and Terkaj, W. (2016) ‘EXPRESS to OWL for construction industry: Towards a recommendable and usable ifcOWL ontology’, Automation in Construction, 63, pp. 100–133. Available at: https://doi.org/10.1016/j.autcon.2015.12.003.

Perera, S. et al. (2020) ‘Blockchain technology: Is it hype or real in the construction industry?’, Journal of Industrial Information Integration, 17, p. 100125. Available at: https://doi.org/10.1016/j.jii.2020.100125.

Pomponi, F. and Moncaster, A. (2017) ‘Circular economy for the built environment: A research framework’, Journal of Cleaner Production, 143, pp. 710–718. Available at: https://doi.org/10.1016/j.jclepro.2016.12.055.

Psaras, Y. and Dias, D. (2020) ‘The InterPlanetary File System and the Filecoin Network’, in 2020 50th Annual IEEE-IFIP International Conference on Dependable Systems and Networks-Supplemental Volume (DSN-S). 2020 50th Annual IEEE/IFIP International Conference on Dependable Systems and Networks - Supplemental Volume (DSN-S), Valencia, Spain: IEEE, pp. 80–80. Available at: https://doi.org/10.1109/DSN-S50200.2020.00043.

Ramachandran, G.S., Radhakrishnan, R. and Krishnamachari, B. (2018) ‘Towards a Decentralized Data Marketplace for Smart Cities’, in 2018 IEEE International Smart Cities Conference (ISC2). 2018 IEEE International Smart Cities Conference (ISC2), Kansas City, MO, USA: IEEE, pp. 1–8. Available at: https://doi.org/10.1109/ISC2.2018.8656952.

Rezgui, Y., Beach, T. and Rana, O. (2013) ‘A GOVERNANCE APPROACH FOR BIM MANAGEMENT ACROSS LIFECYCLE AND SUPPLY CHAINS USING MIXED-MODES OF INFORMATION DELIVERY’, Journal of Civil Engineering and Management, 19(2), pp. 239–258. Available at: https://doi.org/10.3846/13923730.2012.760480.

Rutten, M.E.J., Dorée, A.G. and Halman, J.I.M. (2009) ‘Innovation and interorganizational cooperation: a synthesis of literature’, Construction Innovation, 9(3), pp. 285–297. Available at: https://doi.org/10.1108/14714170910973501.

Schöggl, J.-P. et al. (2023) ‘Implementation of digital technologies for a circular economy and sustainability management in the manufacturing sector’, Sustainable Production and Consumption, 35, pp. 401–420. Available at: https://doi.org/10.1016/j.spc.2022.11.012.

Shang, S.S.C. et al. (2011) ‘Understanding Web 2.0 service models: A knowledge-creating perspective’, Information & Management, 48(4–5), pp. 178–184. Available at: https://doi.org/10.1016/j.im.2011.01.005.

Sheikh, S. et al. (2022) ‘Distributed records in the Rohingya refugee diaspora: Arweave and the R-Archive’, Journal of Documentation, 79(4), pp. 813–829. Available at: https://doi.org/10.1108/JD-08-2022-0174.

Singh, V., Gu, N. and Wang, X. (2011) ‘A theoretical framework of a BIM-based multi-disciplinary collaboration platform’, Automation in Construction, 20(2), pp. 134–144. Available at: https://doi.org/10.1016/j.autcon.2010.09.011.

Soman, R.K., Molina-Solana, M. and Whyte, J.K. (2020) ‘Linked-Data based Constraint-Checking (LDCC) to support look-ahead planning in construction’, Automation in Construction, 120, p. 103369. Available at: https://doi.org/10.1016/j.autcon.2020.103369.

Stacks (2024) Gaia: A decentralized high-performance storage system. Available at: https://github.com/stacks-network/gaia (Accessed: 7 March 2024).

Swarm Foundation (2024) Swarm. Available at: https://www.ethswarm.org/ (Accessed: 7 March 2024).

Tao, F. et al. (2018) ‘Digital twin-driven product design, manufacturing and service with big data’, International Journal of Advanced Manufacturing Technology, 94(9–12), pp. 3563–3576. Available at: https://doi.org/10.1007/s00170-017-0233-1.

Tao, X. et al. (2021) ‘Distributed common data environment using blockchain and Interplanetary File System for secure BIM-based collaborative design’, Automation in Construction, 130, p. 103851. Available at: https://doi.org/10.1016/j.autcon.2021.103851.

Tao, X. et al. (2023) ‘Smart contract swarm and multi-branch structure for secure and efficient BIM versioning in blockchain-aided common data environment’, Computers in Industry, 149, p. 103922. Available at: https://doi.org/10.1016/j.compind.2023.103922.

The IPFS project (2024) IPFS Documentation | IPFS Docs. Available at: https://docs.ipfs.tech/ (Accessed: 7 March 2024).

Tomczak, A. et al. (2024) ‘Requiring Circularity Data in Bim With Information Delivery Specification’, Circular Economy, 1(2). Available at: https://doi.org/10.55845/REJY5239.

Turk, Z. and Klinc, R. (2017) ‘Potentials of Blockchain Technology for Construction Management’, Procedia Engineering, 196, pp. 638–645. Available at: https://doi.org/10.1016/j.proeng.2017.08.052.

Varriale, V. et al. (2024) ‘The role of digital technologies in production systems for achieving sustainable development goals’, Sustainable Production and Consumption, 47, pp. 87–104. Available at: https://doi.org/10.1016/j.spc.2024.03.035.

Wang, C., Cho, Y.K. and Kim, C. (2015) ‘Automatic BIM component extraction from point clouds of existing buildings for sustainability applications’, Automation in Construction, 56, pp. 1–13. Available at: https://doi.org/10.1016/j.autcon.2015.04.001.

Wang, S., Zhang, Yinglong and Zhang, Yaling (2018) ‘A Blockchain-Based Framework for Data Sharing With Fine-Grained Access Control in Decentralized Storage Systems’, IEEE Access, 6, pp. 38437–38450. Available at: https://doi.org/10.1109/ACCESS.2018.2851611.

Wang, Y. et al. (2023) ‘Incentive-Aware Decentralized Data Collaboration’, Proceedings of the ACM on Management of Data, 1(2), pp. 1–27. Available at: https://doi.org/10.1145/3589303.

Wang, Z.-J. et al. (2019) ‘Decentralized Data Marketplace to Enable Trusted Machine Economy’, in 2019 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE). 2019 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE), Yunlin, Taiwan: IEEE, pp. 246–250. Available at: https://doi.org/10.1109/ECICE47484.2019.8942729.

Werbrouck, J. et al. (2019) ‘Towards a semantic Construction Digital Twin: Directions for future research’, in Advances in ICT in Design, Construction and Management in Architecture, Engineering, Construction and

Operations (AECO) : Proceedings of the 36th CIB W78 2019 Conference. 36th CIB W78 2019 Conference, pp. 113–123. Available at: http://hdl.handle.net/1854/LU-8633673 (Accessed: 10 September 2023).

Wijewickrama, M.K.C.S. et al. (2021) ‘Information sharing in reverse logistics supply chain of demolition waste: A systematic literature review’, Journal of Cleaner Production, 280, p. 124359. Available at: https://doi.org/10.1016/j.jclepro.2020.124359.

Wijewickrama, M.K.C.S., Rameezdeen, R. and Chileshe, N. (2021) ‘Information brokerage for circular economy in the construction industry: A systematic literature review’, Journal of Cleaner Production, 313, p. 127938. Available at: https://doi.org/10.1016/j.jclepro.2021.127938.

Williams, S. et al. (2019) ‘Arweave: A protocol for economically sustainable information permanence’, arweave.org, Tech. Rep [Preprint]. Available at: https://arweave.org/yellow-paper.pdf (Accessed: 7 March 2024).

Yan, J. et al. (2022) ‘From building to city level dynamic digital Twin: a review from data management perspective’, IOP Conference Series: Earth and Environmental Science, 1101(9), p. 092033. Available at: https://doi.org/10.1088/1755-1315/1101/9/092033.

Yu, T. and Jajodia, S. (eds) (2007) Secure Data Management in Decentralized Systems. Boston, MA: Springer US (Advances in Information Security). Available at: https://doi.org/10.1007/978-0-387-27696-0.

Zahed Benisi, N., Aminian, M. and Javadi, B. (2020) ‘Blockchain-based decentralized storage networks: A survey’, Journal of Network and Computer Applications, 162, p. 102656. Available at: https://doi.org/10.1016/j.jnca.2020.102656.

Zhang, X. et al. (2019) ‘Frameup: An incriminatory attack on Storj: A peer to peer blockchain enabled distributed storage system’, Digital Investigation, 29, pp. 28–42. Available at: https://doi.org/10.1016/j.diin.2019.02.003.

Zhang, Y. et al. (2017) ‘A framework for Big Data driven product lifecycle management’, Journal of Cleaner Production, 159, pp. 229–240. Available at: https://doi.org/10.1016/j.jclepro.2017.04.172.

Zheng, Q. et al. (2018) ‘An Innovative IPFS-Based Storage Model for Blockchain’, in 2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI). 2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI), Santiago: IEEE, pp. 704–708. Available at: https://doi.org/10.1109/WI.2018.000-8.

APPENDIX A

A.1 Overview

To facilitate the overview and classification of different networks (see Section 4), it is essential to first understand the technical basics. It is important to distinguish between the protocol layer and the data network or system. There are several protocols that enable a distributed system. Table 2 summarizes the most important ones.

In particular, the Interplanetary File System (IPFS) is often described as both a protocol and a system. However, it is more accurately described as a bundle of protocols that aim to enhance the resilience of the Web (Daniel and Tschorsch, 2022). However, IPFS is just one of several networks that have been created with different objectives in mind. For example, specific networks offer a data system that is blockchain-agnostic and can be tailored to suit various data models (see Appendix A.5 Gaia).

Other networks are closer tied to a blockchain, resulting in a crypto-economic system with well-defined incentive mechanisms rather than a pure data network (see Appendix A.4 Swarm). Next, there are networks that mirror centralized cloud providers. First, token-based storage, where a user pays for a certain amount of data for a certain period of time, following a pay-as-you-go model (see Appendix A.3 Filecoin). Second, the perpetual storage model, where the user pays a one-time fee to store records permanently (see Appendix A.8 Arweave). In the following, different solutions are examined in detail, while at the end a summary table with the perspectives of data sovereignty is presented, mutability and universal addressing.

A.2 Interplanetary File System (IPFS)

IPFS employs a Merkle DAG (directed acyclic graph) to structure its data for content addressing and fragmented storage, making it one of the most widely used storage systems (Wang, Zhang and Zhang, 2018; The IPFS project, 2024). This enables each file to possess a unique hash value, similar to a fingerprint, enabling larger files to be represented via a content-sensitive tree structure. Consequently, content-based addressing is employed rather than address-based addressing since the content determines the hash and therefore the addressing. While this enhances network performance, it also restricts the ability to modify data after storage without altering addressing, thus making IPFS less appropriate for mutable data that require occasional changes (Zheng *et al.*, 2018). Furthermore, the architecture does not inherently maintain data sovereignty, since the data is published publicly on the network.

A.3 Filecoin

Filecoin provides a supplementary layer that offers a monetary incentive system on top of IPFS (Guidi, Michienzi and Ricci, 2022). This blockchain-based layer records the commitments made by participants in the network in a trustworthy manner. Data creators are charged for storing and retrieving data from storage providers, who are obligated to ensure that the information remains available and stored continuously and transparently. Retrieval miners facilitate information retrieval from stored data, obtained either directly from the customer or from storage miners. The computational effort of both types of miners is compensated with network-specific tokens. In summary, the network ensures universal data integrity through storage guarantees and public transparency, while discoverability is achieved through internal location management (Psaras and Dias, 2020). This ensures that the customer receives what they have stored.

A.4 Swarm

In response to one of the main limitations of blockchain systems, namely their storage capacity, efforts are being made to develop storage networks that complement these systems. Swarm is one such network, which offers a decentralized alternative to centralized cloud storage services and is a key component of the Ethereum ecosystem (Swarm Foundation, 2024). Data storage is facilitated through two mechanisms- global pinning and storage nodes. In the former, the user stores the data on their drive, and the network receives a reference only, while the latter involves storage nodes that provide redundancy and store both fragmented and uploaded content. The postage stamp system is used for this purpose. Although global pinning is free, it poses the risk of data loss due to lack of redundancy. Overall, Swarm facilitates the creation of decentralized applications that ensure that their data are censorship and single-point-of-failure attack resistant, achieved through decentralized storage and microtransactions (Ozyilmaz and Yurdakul, 2019).

A.5 Gaia

Gaia is a storage network that is used in conjunction with the Stacks blockchain (Stacks, 2024). It differs from other solutions, such as IPFS, which are designed for immutable, censorship resistant, and permanent storage but do not offer user control or decentralized access mechanisms over the content (Hunhevicz *et al.*, 2023). Gaia allows users to maintain control over where their information is stored while connecting access to those data to an on-chain decentralized identifier (DID) solution. Gaia is designed to store user application data off the blockchain, thereby improving performance and accessibility for reading and writing, without the need for a central trust authority. Its architecture consists of three layers: the trustless ledger and decentralized identifiers at the base layer, the routing network at the second layer, and the storage system at the final layer. The blockchain governs the ownership of identities within the network. Routing mechanisms are held in a peer-to-peer system that links identities to a chosen storage location in the storage network. Data storage operates as a simple key-value store, and when a user interacts with a decentralized application, the latter stores information on the user's behalf. The hub service writes to the store by requesting a valid authentication token from a requester while the storage is usually a separate, dedicated storage resource provided by a common cloud provider.

A.6 Ceramic

Ceramic is a decentralized data network that aims to provide an open but predictable data layer, where predictable means that the interfaces or tools have prior knowledge of the expected data structure and data model for a given application (3Box Labs, 2022). This allows the creation, updating, and sharing of verifiable and tamper-proof data streams. Additionally, these data streams adhere to a clearly defined data model, allowing for interoperability of data between applications and composable data. Ceramic uses a blockchain-agnostic approach that allows it to work with different blockchain technologies, making it highly adaptable and facilitating a wide range of use cases. At the core of this approach is a cross-chain DID solution that associates structured data with a specific identifier. This is then consolidated into a distributed user table, where each row corresponds to a user identifier.

A.7 Storj

Another network that, like IPFS, emphasizes file-based storage is Storj. The main attribute of the network is its robust and encrypted cloud infrastructure (Li *et al.*, 2023). This allows replicating the functions of modern cloud solutions through the network while ensuring users of similar performance with additional advantages such as improved resilience, reduced expenses, and increased security (Zhang *et al.*, 2019). Sharding is the process used to achieve this, by which the file to be uploaded is split into smaller pieces and then dispersed among multiple storage nodes. This ensures that no single node has complete access to the entire file. Furthermore, prior to uploading, it is standard practice to encrypt the file. To avoid network bloating, only metadata, rather than the complete data set, are anchored on the Ethereum blockchain. However, certain limitations must be taken into account. Storj operates at the file level, which precludes mutability or the ability to change stored data without re-uploading the whole file. Additionally, external applications are unable to access stored data due to lack of control over data location by the user or creator.

A.8 Arweave

Arweave is an archiving network that aims at permanence, providing an unchanging repository of data, unlike platforms that prioritize mutable storage (Williams *et al.*, 2019). Its primary focus is to create a permanent network where data cannot be altered or deleted (He, 2023). Unlike traditional blockchain technologies, which rely on a linear block structure, Arweave adopts a graph-like model in which each block is linked to two prior blocks (Sheikh *et al.*, 2022). This enforces the reliability of the data history. Furthermore, Arweave requires a one-time payment for data storage, which assigns a perpetual endowment to each piece of data. Consequently, it ensures indefinite preservation without the need for ongoing fees. This economic model is based on the assumption that storage costs will decrease over time, meaning that the initial payment will cover long-term data storage costs. Arweave's use of the HTTP protocol allows for seamless integration with existing web infrastructure. This feature enables users to access the stored content directly through web browsers, facilitating easy retrieval of the data. Arweave prioritizes accessibility, with the goal of revolutionizing the way knowledge and data are preserved and disseminated without censorship or temporal data management policies.

A.9 Summary

In conclusion, the variations between the different storage solutions are fundamentally nuanced. Table 2 may be used as an initial reference framework, where the comparison of solutions has been condensed into three main perspectives: data sovereignty, data mutability, and universal addressing.

In the context of data sovereignty, the degree of control a user has over their data is a critical variable. For instance, user control can be enhanced by employing networks like Gaia along with the Stacks blockchain, which connect data storage with proprietary storage solutions. In terms of data mutability, the ability to modify stored data after storage is important. Protocols such as IPFS excel at achieving content addressability and decentralized storage. However, they come with limitations concerning mutability, as changes to content require an update of the content's unique hash to reflect the changes made. Universal addressing refers to the capability to retrieve data from anywhere. Networks like Ceramic have implemented mechanisms to facilitate this accessibility by adopting a blockchain-agnostic framework and decentralized identifiers (DIDs), thereby enabling extensive data accessibility without being encumbered by infrastructural constraints.

Table 2: Initial reference framework comparing different solutions using the perspectives of data sovereignty, data mutability, and universal addressing.

Perspectives	Data Sovereignty	Data Mutability	Universal Addressing
IPFS	IPFS ensures data sovereignty through content-derived identifiers, although public access is inherent.	It is designed for immutable content, making file modifications challenging.	Content addressing enables universal addressing.
Filecoin	Through its connection to IPFS and an overlying incentive layer, Filecoin gives data creators a level of data sovereignty similar to IPFS.	Despite having an additional layer, the data stored on the network cannot be altered or modified, thus limiting the mutability of the data.	It is also using the mechanism of content addressing enabling universal addressing.
Swarm	Swarm provides a high-level of data sovereignty, as only the data owner has control over their data and can decide how it is used and shared.	It enables data mutability by allowing users to update and modify content at both the file and data level.	The use of the Ethereum Name Service (ENS) enables a universal addressing system that does not depend on storage location or data format.
Gaia	Gaia offers a high-level of data sovereignty, as users have control over the exact location and their access to data.	It facilitates fine-grained data mutability, allowing modifications at the data level without necessitating the publication of an entirely new version.	Due to the significant degree of data sovereignty and the individualized storage selection, the concept of universal addressing is unattainable.
Ceramic	Ceramic ensures data sovereignty by tying data to decentralized identifiers, allowing users to retain control over their information.	Data on Ceramic are mutable, allowing for updates and modifications to data streams without changing their unique identifiers.	Universal addressing is facilitated by Ceramic's cross-chain DID solution, which operates independently of the underlying blockchain protocol.
Storj	Storj provides users with data sovereignty through encrypted shards, ensuring that data is accessible only to those who have permission.	Data mutability is constrained since modifications require re-encryption and re-distribution of the file's shards.	It utilizes a decentralized network of storage nodes, but universal addressing is not its core focus; access is managed through encryption keys and permissions.
Arweave	Arweave grants data sovereignty by allowing permanent storage, where data owners can ensure that their data remain unchanged indefinitely.	Mutability is not a feature as it is designed for permanent, immutable data storage.	Universal addressing in Arweave is inherent, as each piece of data is stored with a unique transaction ID, accessible universally.