IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 26 July 2024, accepted 6 August 2024, date of publication 13 August 2024, date of current version 23 August 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3442968

RESEARCH ARTICLE

Empowering Dataspace 4.0: Unveiling Promise of Decentralized Data-Sharing

SAEED HAMOOD ALSAMHI[1,2,3], AMMAR HAWBANI[4], SANTOSH KUMAR[5], (Senior Member, IEEE), MOHAN TIMILSINA[1], MAJJED AL-QATF[1], RAFIQUL HAQUE[1], FARHAN M. A. NASHWAN[2], LIANG ZHAO[4], (Member, IEEE), AND EDWARD CURRY[1]

[1]Insight Centre for Data Analytics, University of Galway, Galway, H91 TK33 Ireland
[2]Department of Computer Science and Engineering, College of Informatics, Korea University, Seongbuk-gu, Seoul 02841, Republic of Korea
[3]Electrical Engineering Department, Faculty of Engineering, IBB University, Ibb, Yemen
[4]School of Computer Science, Shenyang Aerospace University, Shenyang 110136, China
[5]Department of CSE, IIIT Naya Raipur, Naya Raipur, Chhattisgarh 493661, India

Corresponding author: Saeed Hamood Alsamhi (saeed.alsamhi@insight-centre.org)

This work was supported by the Science Foundation Ireland under Grant SFI/12/RC/2289_P2.

ABSTRACT Recently, there has been more interest in Decentralized Data-Sharing (DDS) because of the introduction of Dataspace 4.0. DDS is becoming increasingly popular as a safe, open, and effective way for many parties to data-sharing. Unlike conventional, centralized methods, DDS has several benefits, such as better knowledge exchange, higher accessibility and interoperability, and data privacy and security. The paper covers DDS's advantages, including improved resilience, higher security, increased privacy, and improved interoperability. DDS gives people and organizations more control and ownership over their data while reducing the dangers of centralized data management. In this survey, we highlight promising technologies for DDS in Dataspace 4.0, including Federated Learning (FL), blockchain, decentralized file systems, semantic web and knowledge representation, and Peer-to-Peer (P2P) networks. We highlight the challenges, opportunities, and future directions of technology enabling further DDS advancement in Industry 4.0.

INDEX TERMS Industry 4.0, P2P network, dataspace, federated learning, dataspace 4.0, decentralized data-sharing, blockchain, decentralized file systems, semantic web.

I. INTRODUCTION

The Industrial Internet of Things (IIoT), big data, and cloud computing have all proliferated, leading to the emergence of Dataspace 4.0, a digital ecosystem that makes it simpler for stakeholders to connect and share massive amounts of data from many sources in a seamless manner [1], [2], [3]. "Dataspace 4.0" describes data management and sharing expected to facilitate data integration throughout various businesses and domains in Industry 4.0 [2], [4], [5], [6]. The capabilities of Dataspace 4.0 are anticipated to be significantly expanded with the introduction of 6G, opening up new possibilities for data-driven services and applications [2], [6], [7], [8]. Similarly, in Dataspace 4.0, data-sharing facilitates the establishment of networked data ecosystems where enterprises may safely share, access, and employ data to spur innovation and ensure a competitive edge in quickly changing markets. Dataspace 4.0's capabilities are expected to increase significantly, creating chances for data-driven services and applications in Industry 4.0 [9]. With the ability to analyze and access massive data for many reasons, including research, innovation, and decision-making, data-sharing has become a crucial part of modern society for organizations, governments, and individuals.

Data-sharing has several advantages that may affect people, businesses, and society. Data-sharing promotes collaboration between various stakeholders by enabling the sharing of data which supports the development of discoveries and innovations in addition to effective and efficient problem-solving. Data-sharing supports the development of

The associate editor coordinating the review of this manuscript and approving it for publication was Claudio Loconsole.

stakeholder trust, lowers corruption, promotes justice, and raises accountability and transparency [10], [11]. Additionally, data-sharing increases productivity and reduces waste by facilitating more effective resource use, which lowers costs and increases efficiency. Providing access to data that would not otherwise be available promotes innovation and results in discoveries, insights, goods, and services. By allowing people and organizations to submit data to open datasets and repositories, data-sharing may also advance the public good by promoting research and development in healthcare, the environment, and society.

However, there are several drawbacks to traditional centralized data-sharing systems, such as their restricted accessibility, privacy and data security concerns, interoperability problems, and reliance on single points of failure [12], [13], [14], [15], [16]. Centralized systems have facilitated data-sharing, but limitations hinder collaboration and innovation. Decentralized Data-Sharing (DDS) has emerged and empowered data-sharing collaborations and interactions effectively and efficiently overcoming the shortcomings of Centralized systems. In contrast to systems, DDS provides control and management of data, increased privacy, and security measures, as well as enhanced levels of accountability, accessibility, and transparency through the distribution of data across multiple nodes without the reliance on a central authority [17], [18], [19]. Table. 1 presents a comparison of centralized data sharing. DDS enables data sharing among parties in Dataspace 4.0 without the need for centralized control enhancing aspects, like data security, privacy, collaboration, and economic opportunities [20].

DSS-enabled technologies have shown promise in addressing data-sharing issues, which include blockchain [21], [22], Decentralized File Systems (DFS) and Semantic Web (SW), and Knowledge Representation (KR), Federated Learning (FL) [23], and Peer-to-peer (P2P) networks [24],. Therefore, the aforementioned technologies are empowered by DDS within the Dataspace 4.0 paradigm. When combined, the DSS-enabled technologies open the door to improved security, cooperation, and creativity in data exchange and management. Thus, According to the best of the authors' knowledge, studies have yet to address DDS, particularly about enabling technologies in supporting DDS in Dataspace4.0 domains and applications.

A. MOTIVATION AND CONTRIBUTIONS

Driven by the requirement to address issues related to data security, privacy, accessibility, and interoperability, this survey is to investigate the useful uses of DDS in the context of Dataspace 4.0. To understand how DSS-enabled technologies support Dataspace 4.0 domains and applications, we examine upcoming technologies such as blockchain, DFS, SW KR, FL, and P2P networks. Promising answers are provided by DDS models, which make it possible to create more safe, interoperable systems that effectively promote cooperation and information sharing. We provide a comprehensive survey to investigate the practical use of DDS technologies within the context of Dataspace 4.0. Ultimately, this article provides readers with a comprehensive grasp of the present situation regarding DDS for Dataspace 4.0 while highlighting the main prospects and obstacles for further study and advancement. The survey's findings offer insightful information on DDS needs and requirements. This information helps shape technologies to make data-sharing in the sector more effective, safe, and interoperable. The contributions of this study are summarised as follows:

1) We provide a comprehensive overview of DDS within the context of Dataspace 4.0, identifying and discussing the advantages, requirements for DSS, challenges, and potential of DDS.
2) This paper provides a novel framework for integrating FL, blockchain, decentralized file systems, semantic web, and P2P networks to create a cohesive and effective DDS environment. We describe the framework layers to achieve the DSS effectively and efficiently by highlighting the enabling technologies roles processing, and methods. Dataspace 4.0 is empowered by this multifaceted technology, which guarantees the many areas of data management, including privacy, security, accessibility, and interoperability.
3) We describe the enabled technologies that go into DDS, emphasizing the benefits, drawbacks, and possible uses of enabling Dataspace 4.0 for decentralized network management and upkeep, data security and privacy assurance, privacy and utility balance, impartial and equitable network participation, energy efficiency, and sustainability.
4) The proposed framework offers a structured framework for integrating DDS technologies in Dataspace 4.0. The framework intends to expand the capabilities of DDS, blockchain technology, and Industry 4.0 by combining important DDS technologies and tackling their issues with creative solutions. To improve data-sharing in the digital age and make it more safe, efficient, and interoperable, the findings provide insightful information for future study and development.
5) We provide insights into the future directions of DDS in Dataspace 4.0 to improve efficiency and accuracy, the standardization of DDS protocols, the development of governance models, and the focus on energy efficiency and sustainability.

B. RELATED SURVEYS

In the context of Dataspace 4.0, DDS emphasizes the method's increasing popularity and many advantages. The disadvantages of conventional centralized techniques are addressed by DDS, which is acknowledged for offering a safe, transparent, and effective means of data sharing for several parties [25], [26]. Improving data security and privacy is one of the main benefits of DDS, as highlighted in the surveys. DDS lessens the hazards of centralized data

TABLE 1. Comparison of centralized and DDS.

Items	Centralized Data-Sharing	DDS
Data Control	Central authority controls access and permissions for all data	Data owners control access and permissions for their data
Scalability	Limited scalability due to reliance on a single central server	High scalability due to distribution of data across multiple nodes
Security	Vulnerable to attacks and data breaches due to reliance on a single point of failure	More secure due to distribution of data and use of encryption
Interoperability	Limited interoperability with other systems and data sources	High interoperability through standard data formats and protocols
Flexibility and Agility	Limited flexibility and agility due to centralized decision-making	High flexibility and agility due to distributed decision-making and collaboration
Cost Efficiency	Higher cost due to maintenance and operation of centralized server	Lower cost due to the distribution of data and reduced infrastructure requirements

repositories, including cyberattack vulnerabilities and single points of failure, by decentralizing data management [5], [27]. DDS enhances data interoperability and accessibility, promoting improved information sharing across various institutions [28], [29].

Many surveys identify key enabling technologies for DDS in Dataspace 4.0. FL maintains confidentiality and privacy by enabling machine learning model training across decentralized data sources without sharing the data [30]. Given its reputation for security and transparency, blockchain technology plays a critical role in maintaining trust and data integrity in decentralized environments [31]. Distributed storage and file access are made possible by decentralized file systems like IPFS, which improve data availability and lessen the need for central servers [32]. Semantic web and knowledge representation approaches enhance data interoperability and increase data usability across many systems [2].

C. PAPER STRUCTURE

The rest of the paper is organized as follows and is shown in Figure.1. Section II introduces DSS, while the Dataspace 4.0 framework is provided in Section III. Section IV discusses the enabling technologies for DSS in Dataspace 4.0 including blockchain, FL, P2P network, DFS, and SW and KR. Section V discusses the comparative analysis of enabling technologies of DSS in Dataspace 4.0.

II. DECENTRALIZED DATA-SHARING

The framework consists of five layers, including data acquisition and preprocessing, data encryption and decryption, data-sharing and collaboration, data analysis and insights, and governance and ethics, as shown in Figure 3. To guarantee quality and consistency, raw data from various sources, including IoT and IIoT devices, is cleaned, formatted, normalized, and preprocessed at the first layer of data acquisition and preprocessing. This preliminary stage is crucial to transforming unprocessed data into a format that may be used for further examination. IoT and IIoT technologies are essential in this context because they make gathering data from sensors and devices in real-time more accessible, especially in industrial settings where optimization and monitoring are critical.

Distributing data across a network of autonomous participants instead of depending on a centralized body to oversee and regulate data access is known as DDS. Every participant in a DDS is in charge of keeping their copy of the data up to date and in good condition. Participants can exchange data using P2P networks or direct data-sharing with other participants. DSS minimizes the chance of unwanted access or data breaches and offers better security and privacy than centralized approaches, making security and privacy distinct. DDS gives people and businesses more control over their data, improving the safeguarding and transparency of intellectual property. DDS promotes increased network and platform compatibility, which lessens fragmentation and enhances data interchange. Additional advantages of decentralized sharing include trust and transparency, facilitating safe, verifiable data-sharing, and lessening disputes. Furthermore, by offering a platform for data-sharing and insights that result in discoveries, goods, and services, DDS across industries fosters creativity and cooperation.

Critical concepts of DDS include P2P networks, distributed ledgers, data ownership, and governance, security, and privacy. In P2P, DDS allows for direct communication and data-sharing between participants without relying on a centralized intermediary. This way offers benefits such as increased security, improved scalability, and reduced latency compared to centralized systems [33]. Depending on the system's specific requirements, P2P networks are designed in various ways, such as structured, unstructured, or hybrid. In distributed ledgers, DDS maintains a log of all modifications and transactions. Instead of being kept by a single entity, the ledger is shared among all network nodes, preventing fraud and manipulation and ensuring all users get an accurate picture of the data. Compared to centralized solutions, distributed ledgers can offer better security and transparency. Multiple individuals with varying degrees of ownership and control over the data are involved in DDS. Because of this, data ownership and governance are intricate and must be managed carefully. Without open governance, DDS is vulnerable to disagreements, data loss, and misuse. Therefore, a strong governance structure is essential to guaranteeing equitable and open data management procedures. To guard against data breaches and unwanted access to sensitive information, DDS is built with security and privacy [5]. To safeguard the data, DDS uses access

TABLE 2. Comparison of related surveys and enabling technologies.

Ref.	Highlighted	Enabled techn.	Benefits
[25](2021)	Benefits of DDS	FL	Enhanced data privacy and security
[26](2023)	Accessibility and interoperability in DDS	SW, knowledge representation	Improved data accessibility and interoperability
[27](2023)	Control and ownership in decentralized data management	DFS, P2P Networks	Greater control and ownership over data, reduced centralized risks
[5](2024)	Security benefits of DDS	Blockchain, FL	Higher security, minimized single points of failure
[28](2021)	FL for Privacy-Preserving ML	FL	Preserves privacy by localizing data
[29](2022)	Blockchain in DDS for ensuring data integrity	Blockchain	Secure and transparent data transactions
[30](2022)	DFS for data sharing	DFS (e.g., IPFS)	Enhanced data availability, reduced reliance on central servers
[31](2022)	Enhancing interoperability with SW	Semantic Web, Knowledge Representation	Better data usability across systems
[32](2020)	P2P networks in decentralized systems	P2P Networks	Direct data sharing, reduced need for intermediaries
[2](2023)	DDS for secure data sharing	Blockchain, FL, Dataspace	Identification of major implementation challenges
This survey	DDS and enabling technologies	Blockchain, FL, P2P, SW, DFS, Dataspace, Dataspace 4.0	Advantages, security, privacy, accessibility, and interoperability

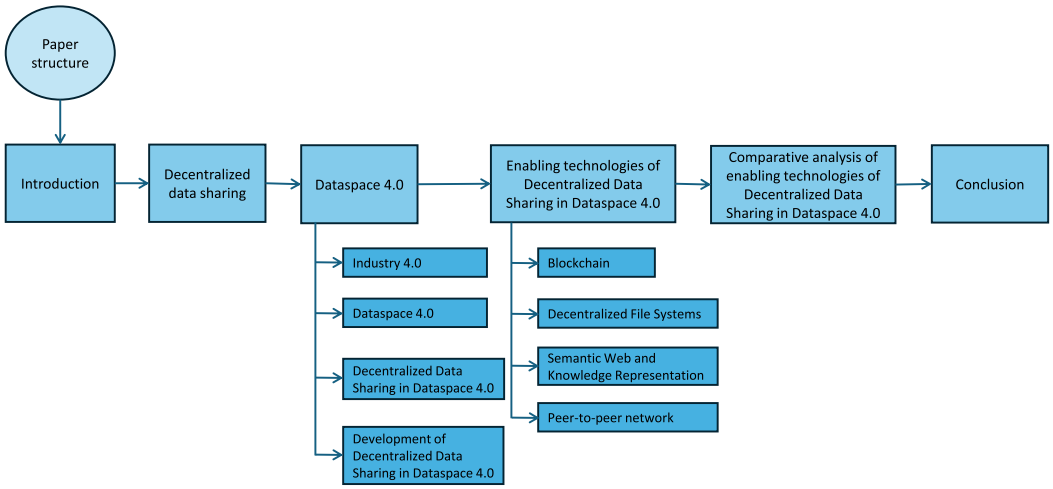

FIGURE 1. Paper structure.

restrictions, encryption, and other security measures [19], [34].

DDS does not require a central body to administer the data to be shared across several parties. Instead of being controlled by a single company or institution, the parties that possess the data store and manage it in a DDS.

Enhanced Security: By doing away with the requirement for a single central authority that might be the target of data breaches or cyberattacks, DDS can increase security [35]. Instead, information is dispersed among a network of nodes, increasing the difficulty with which malevolent actors may access all the data. DDS improved security by lowering the possibility of cyberattacks and data breaches [36]. For instance, a centralized system puts all of the data under one control or in one place, which makes it a target for hackers. In contrast, a decentralized system distributes data over a network of nodes, increasing the difficulty of data access for hackers [37]. DDS platforms use cryptography to protect the data. A private key is stored on each network node to encrypt and decode data. An extra degree of security is provided when shared data is encrypted using the recipient's public key, guaranteeing that only the intended recipient can access the data. The hacker manages to get access to the network, but the hacker will be unable to access the data without the private keys [38]. DDS uses consensus algorithms to guarantee the accuracy and immutability of data. Consensus methods provide every node in the network consensus on the data's current state and ensure that the network validates any alterations made to the data [39]. Thus, decentralised data-sharing delivers improved security by lowering the possibility of data breaches, securing the data with encryption, guaranteeing data correctness with consensus methods, and giving users more control over data access, as illustrated in Table. 3.

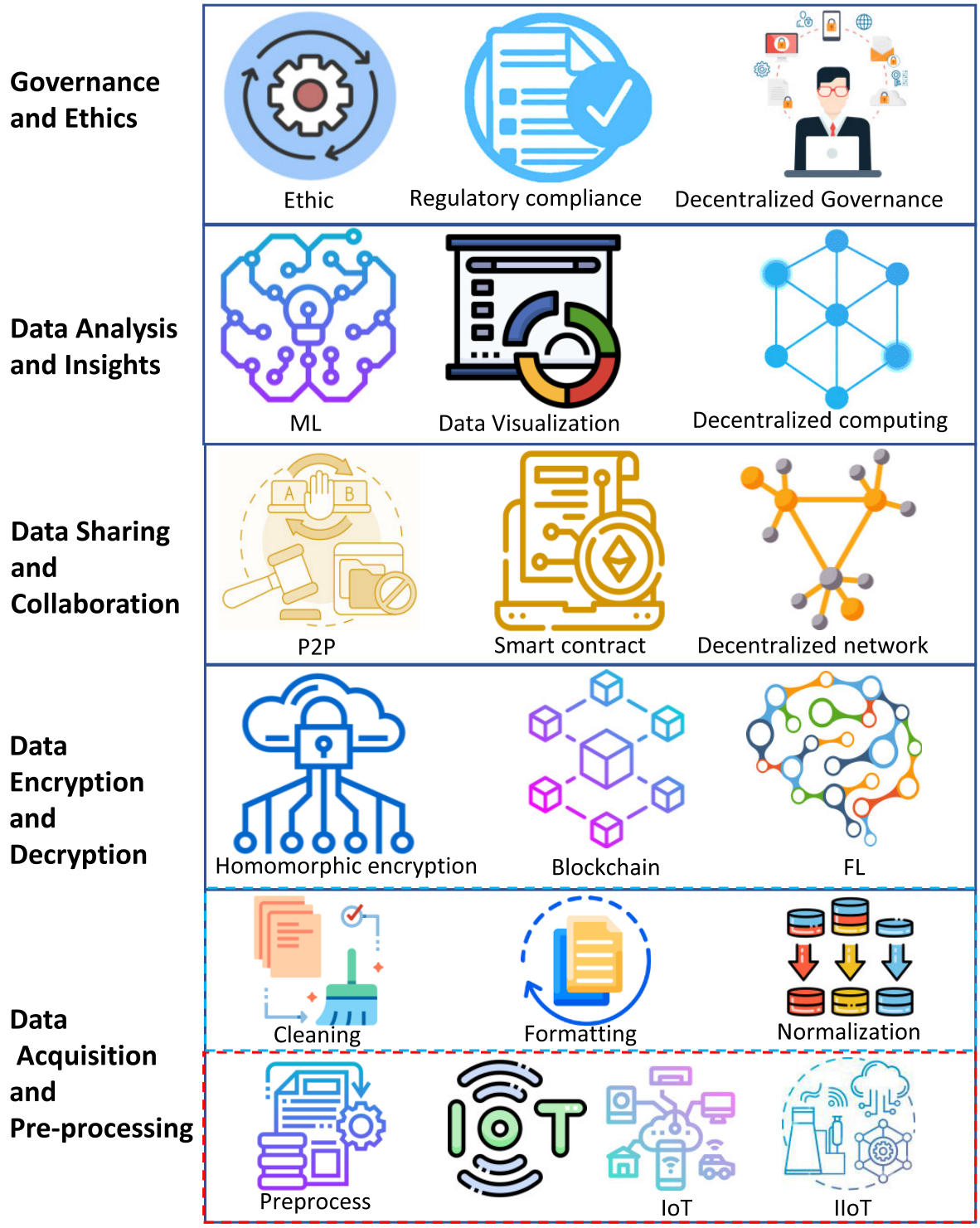

FIGURE 2. DDS framework.

Increased Privacy: DDS gives users control over who may access their data, which helps to safeguard their privacy. Instead of depending on a centralized body to oversee data access, people only provide access to those they can trust [40]. Each person or organization in a DDS controls the data dispersed over a network of nodes. DDS commonly employ cryptography to safeguard the confidentiality of data. A person or organization needs a private key to encrypt and decrypt data. This helps prevent unwanted data access, ensures that only the intended recipient can access it, and provides individuals with more privacy.

Improved Interoperability: By leveraging open protocols and standards for data-sharing across many applications and platforms, DDS may enhance interoperability across various systems and organizations. DDS enhances interoperability across systems and organizations by leveraging open protocols and standards to communicate data across many platforms and applications. DDS surpasses these restrictions by adopting open protocols and standards that allow data to be shared across many platforms and apps. A DDS enabled by blockchain technology enables safe and transparent data-sharing across many businesses. Through open protocols and standards facilitating data-sharing between many platforms and applications, DDS enhances interoperability and delays in data exchange, allowing different businesses to function more productively and successfully.

Greater Transparency: DDS improves transparency by enabling everyone to see and validate the shared data. DDS, therefore, promotes cooperation and helps to establish confidence between participants. Furthermore, by allowing everyone in a network to view and validate data, DDS can increase openness by producing a transparent and impenetrable record of data-sharing activities [41]. DDS can address transparency issues by recording transactions on a distributed ledger validated and approved by a node network. DDS promotes increased transparency by utilizing smart contracts to guarantee that data is utilized only for intended purposes and by producing a visible, tamper-proof record of data-sharing actions. As a result, data-sharing operations may become more trustworthy and accountable, improving productivity for both individuals and companies.

Improved Resilience: By guaranteeing that data is still accessible even if one or more network nodes fail, DDS increases resilience. The danger of data loss or outage is decreased since the data is dispersed throughout the network, allowing access from several places. DDS increases resilience by establishing a dispersed node network that keeps running even if nodes malfunction or are hacked. To mitigate the dangers, DDS establishes a dispersed node network that can function even if any of the malfunctions are hacked. Every node in the network keeps a copy of the data, and a consensus process involving many nodes verifies and approves transactions. Blockchain technology, for instance, establishes a DDS framework that boosts resilience. Blockchain technology records transactions approved and confirmed by a node network using a distributed ledger. Furthermore, even if specific nodes malfunction or are compromised, the system can still function since the data is prevalent among nodes.

Decentralized Integrated Governance: The seamless data-sharing within and across borders and sectors is a key requirement for gaining data sovereignty which is one of the foundational concepts of Dataspaces and the centre of the strategic goal of the European Union (EU). Nevertheless, cross-border and cross-sector data-sharing foster governance challenges which essentially creates a huge barrier for adopting Dataspaces. A well-defined data governance framework is a *sine quo non* for facilitating data-sharing within Dataspaces. The need for an overarching governance framework is imminent for federated Dataspaces. Governance is one of the critical instruments of standardized Dataspaces prescribed in the design concepts proposed by Open DEI[1] and the reference framework developed by the IDSA.[2]

TABLE 3. Summary of benefits of DDS.

Benefit	Description
Enhanced security	Nodes distributed across the network improve resilience, while encryption enhances data security.
Increased privacy	Encryption and smart contracts ensure that data is used for its intended purpose.
Improved interoperability	Decentralized standards and protocols enhance interoperability.
Improved efficiency	Streamlined processes and reduced costs through the elimination of intermediaries.

IDSA defined data governance principles as common standards due to the decentralized design of Dataspaces and hence the lack of a central supervisory authority. These principles are developed from user needs and specify the rights and duties for data management. Open DEI's governance model is composed of three building blocks, including *overarching cooperation agreement*, *service level agreement (SLA)*, and *continuity model*. These building blocks are needed to guide the responsibilities of different roles (*e.g.,* data owner, data provider, etc.), placing an SLA between stakeholders, and compliance with the rules specifying how data can be used by the different parties. To ensure data sovereignty in data-sharing and exchange, the governance framework is needed to enforce organizational and operational agreements within the Dataspace ecosystem. These agreements are critical requirements to support and allow data usage regulations and instill trust in the whole data ecosystem by acting as a trust anchor connecting the physical and digital worlds. Therefore, it is evident that the overall interoperability in Dataspaces largely depends on operational and organizational agreements which ensure that all parties must maintain and synchronize an acceptable interoperability strategy on an ongoing basis. Some sectors such as energy and healthcare need a tailored governance framework. However, cross-sectorial and cross-border governance frameworks are critically important for data sovereignty. The EU has developed Data Governance Act[3] which aims to facilitate data-sharing across sectors and EU countries to leverage the potential of data for the benefit of European citizens and businesses. The governance act is fully in line with EU values and principles and will bring significant benefits to EU citizens and companies.

The current regulatory frameworks focus on a centralized authority to govern data within Dataspaces. The centralized

[1]Open DEI: https://www.opendei.eu/
[2]IDSA: https://internationaldataspaces.org/
[3]Data Governance Act: https://digital-strategy.ec.europa.eu/en/policies/data-governance-act

regulatory frameworks might be challenging for cross-border Dataspaces, specifically for the ones that involve non-European participants. A decentralized governance- currently missing for Dataspaces- is required to address this challenge. The decentralized framework would enable to engage multiple independent governing authorities from different regimes. The governing authorities work concertedly to maintain the control policies stemming from local and global jurisdictions, organizational policies, and directives for operational, organizational, and data governance. Furthermore, Dataspaces need an integrated governance approach to facilitate intra and inter-Dataspace interoperability [42]. The current governance frameworks do not apply to intra and inter Dataspace interoperability. Therefore, the existing governance framework needs to be extended to support integration as well as decentralization.

Summary: Data collecting, preprocessing, encryption, data-sharing, analysis, and governance are the five levels that make up the framework. IoT and IIoT technologies help collect real-time data, especially in industries where raw data is cleaned and formatted before being analyzed. By allowing direct sharing across P2P networks and requiring users to keep copies of the data, while DDS guarantees security and anonymity. DDS promotes confidence and makes verified data exchange across sectors possible by improving security, privacy, and transparency. Distributed ledgers, P2P networks, and governance are vital ideas that guarantee Dataspace 4.0's improved security, privacy, interoperability, and resilience. For data to be shared seamlessly, integrated governance frameworks must ensure data sovereignty and interoperability inside and beyond industries and borders.

III. DATASPACE 4.0 FRAMEWORK

Data encryption and decryption deal with crucial components of data security. Methods such as homomorphic encryption allow for calculations on encrypted data, protecting privacy without sacrificing analytical power. Blockchain technology uses decentralized, immutable ledger systems to guarantee data integrity and transparency. FL significantly improves privacy by enabling machine learning model training on decentralized edge devices without exchanging sensitive data. P2P networks, smart contracts, and decentralized designs are ways the data-sharing and collaboration layer promotes decentralized data interchange and collaboration. Together, these layers create a robust framework that protects the privacy and security of data and encourages responsible data management and teamwork in a world where connectivity is growing.

A. INDUSTRY 4.0

Industry 4.0 refers to the process of integrating cutting-edge digital technologies into production and industry, including the IoT, big data, and AI [43], [44], [45]. Industrial 4.0 strongly emphasizes data-sharing throughout various industrial segments, sometimes in real-time. Industry 4.0 is beginning to emphasize the need for DDS [46]. Businesses may lower the risk of single points of failure, increase data integrity, and improve security by decentralizing data storage and access. Technologies like P2P networks and blockchain, which offer an irreversible, transparent, and safe record of transactions and spread data throughout a network instead of storing it centrally, facilitate DDS [47]. Furthermore, dispersing data among several nodes through decentralized data storage might improve data security by increasing the difficulty of a single cyberattack compromising the entire system [48], [49]. Additionally, DDS encourages more productive cooperation among various industry participants. A decentralized method allows each institution to manage who has access to its data, allowing trusted partners to share data securely [50]. However, standardization, interoperability, and data privacy considerations must be considered while adopting DDS in Industry 4.0 [51]. It will take ongoing study and advancement in DDS to overcome these obstacles.

Supply Chain Management: Every company that deals with product or service transportation must have a robust supply chain management system. Typically, several parties are involved, including distributors, retailers, suppliers, manufacturers, and customers. Data in supply chains is often compartmentalized and sluggish to flow, resulting in inefficiencies and a lack of transparency [52]. DDS platforms, on the other hand, greatly enhance this concept. Greater efficiency and transparency are made possible by the real-time data-sharing that all supply chain actors may do in a DDS [53]. This rapid access to data from production levels to inventory and demand to delivery status supports better decision-making, an enhanced reaction to changes or interruptions in the supply chain, and performance.

Manufacturing: DDS allows for more flexible and efficient operations, significantly affecting production. Take a network of 3D printers, for instance, that are connected via a DDS network yet can function independently. With this configuration, printers may work together in a coordinated manner on production activities by sharing design files and manufacturing instructions via the network [54]. Manufacturing processes may be made more flexible and efficient through DSS. For instance, by data-sharing amongst the printers, a decentralized network of 3D printers might be utilized to produce goods as needed. For example, design data may be distributed around the network of 3D printers in response to a specific product demand, and they can subsequently work together to build the necessary pieces. A highly flexible manufacturing process that can quickly adapt to shifting demands and eliminate the requirement for centralized finished goods storage is the outcome of DSS. The dangers connected to a single point of failure are also decreased by this decentralized strategy [55].

Energy Management: DDS can completely change the energy industry by enabling it to manage energy resources more sustainably and effectively. Consider, for example, a system of linked smart energy grids, each able to generate, store, and use power. These networks can communicate real-time energy generation, consumption, and storage data

TABLE 4. Benefits of DDS in industry 4.0.

Items	Centralized Data-Sharing	DDS
Supply Chain Management	Suffer from data silos and limited visibility across the supply chain.	Enables greater transparency and collaboration across the supply chain, reducing the risk of bottlenecks and improving efficiency.
Manufacturing	Centralized data-sharing systems can be vulnerable to cyber attacks, causing disruptions to production lines.	DDS allows for greater security and resilience against cyber threats, while also enabling collaboration and information sharing across the manufacturing process.
Healthcare	Centralized data-sharing may raise concerns about data privacy and patient confidentiality.	DDS can provide greater privacy and security, while also facilitating collaboration among healthcare professionals and patients.
Energy Management	Centralized data-sharing can be prone to system failures and blackouts due to the central point of failure.	DDS allows for greater flexibility and scalability, while also enabling more efficient use of energy resources.

TABLE 5. Opportunities for DDS in industry 4.0.

Opportunities	Description of DDS in Industry 4.0
Increased efficiency	DDS can streamline operations and reduce costs by eliminating data silos and enabling real-time collaboration.
Personalization	DDS can enable personalized products and services by providing access to real-time data from multiple sources.
Improved supply chain management	DDS can help optimize supply chain management by providing real-time visibility into inventory levels, production schedules, and shipping status.
Sustainability	DDS can support sustainability initiatives by enabling more efficient use of resources and reducing waste.

TABLE 6. Challenges of DDS in industry 4.0.

Direction	Description of DDS in Industry 4.0
Standardization	Developing standard data formats and communication protocols can achieve interoperability between different systems and facilitate DDS.
ML	AI and ML algorithms can help identify patterns and insights in large datasets, enabling more effective decision-making and improving operational efficiency.
Blockchain technology	Blockchain technology can provide a secure and transparent way to store and share data, enabling DDS in various contexts.
Advanced analytics	Advanced analytics tools can help extract insights from large and complex datasets, enabling organizations to make data-driven decisions and improve operational performance.

thanks to DDS [56]. Grids can better balance supply and demand by exchanging this data. For instance, a grid that generates excess energy can communicate this information with another grid that needs more energy, and another grid can use this surplus. Similarly, a grid with extra energy saved may share it with people in need. Energy costs are stabilized, waste is decreased, and resources are used efficiently thanks to this real-time data interchange. Furthermore, the grid's integration of renewable energy sources is facilitated by DDS. Grids may prepare for and adjust to variations in renewable energy output by exchanging meteorological data, such as wind speed or sunlight, to forecast the generation of renewable energy [57].

Industry 4.0, the fourth industrial revolution defined by the integration of cutting-edge technology to develop intelligent and autonomous systems, heavily relies on DSS. These systems run on data; thus, efficiently and securely transferring data is essential. A DDS architecture facilitates smooth data-sharing between various businesses and organizations, improving efficiency, transparency, and cooperation. To manage supply chains, forecast maintenance requirements, and improve production processes, for example, several equipment, devices, and systems can data-sharing in a smart factory environment [58], [59], [60]. Furthermore, new avenues for development and innovation may be opened by combining DDS with cutting-edge technologies like blockchain, artificial intelligence, and the Internet of Things. For instance, IoT devices may provide a wealth of real-time data that can be shared and used throughout the network, blockchain can offer safe and impenetrable data-sharing, and AI can analyze shared data to extract insights and automate operations [61].

The main advantages of DDS in Industry 4.0 are shown in Table.4. DDS can result in improved results and more innovation, promoting enhanced efficiency, transparency, and cooperation across several industries. The challenges and opportunities of DDS in Industry 4.0 are shown in Tables 5 and 6.

B. DATASPACE 4.0

A new paradigm for data management and sharing called Dataspace 4.0 is founded on decentralization, openness, and interoperability [1]. Dataspace 4.0 is based on a decentralized architecture in which data is stored and managed across a distributed network of participants [62], [63]. Therefore, Dataspace reduces the risk of data loss or tampering, increasing data management's flexibility and scalability. Dataspace 4.0 is designed to be open and accessible to various participants, including individuals, organizations, and machines [64]. The openness promotes collaboration and innovation and allows participants to share and access data more efficiently. Dataspace 4.0 supports interoperability between different data sources and platforms, allowing participants to share and integrate data from other sources easily [64]. Dataspace 4.0 is based on standardized protocols and technologies, ensuring consistency and reliability across different data sources and platforms [65].

Data-sharing has changed dramatically as new concepts and technologies have emerged. The next step in this progression is Dataspace 4.0, which promises a more efficient and integrated method of exchanging and managing data [2], [66]. Dataspace 4.0, with its focus on seamless data access and interoperability, can completely transform how businesses handle their data, resulting in increased

productivity and improved decision-making. The next stage of the industrial revolution, Dataspace 4.0, will further merge the digital and physical realms. DDS will be necessary to facilitate cooperation and creativity safely and effectively. To promote innovation and economic progress in the digital age, DDS will enable many stakeholders to share data while maintaining control over their data. This will ensure privacy, security, and confidence in the data exchange.

On the other hand, several challenges are associated with implementing Dataspace 4.0, including security and privacy, data governance, complexity, and interoperability. Data breaches and unwanted access to private information can occur when data is dispersed and exchanged across network members. Keeping data private and secure is a significant obstacle in the development of Dataspace 4.0. Several parties are involved in Dataspace 4.0, each with varying ownership and influence over the data. Establishing precise data governance standards and processes may become difficult as a result. Dataspace 4.0 uses cutting-edge technologies, including edge computing, semantic web technologies, and distributed ledgers. Dataspace 4.0 systems might be challenging to build and implement because of their technological complexity. While interoperability is a crucial feature of Dataspace 4.0, achieving interoperability between different platforms and technologies can be challenging, particularly given the wide range of data sources and platforms.

C. DDS IN DATASPACE 4.0

DDS paradigm distributes data across several independent nodes or devices instead of storing it in one central location. More control, security, privacy, and scalability are possible with DDS and fault tolerance. With the growing adoption of distributed ledger technologies like blockchain [67], DDS has recently acquired popularity. In Table.1 compares decentralized versus centralized data-sharing. Dataspace 4.0 depends on DDS, which makes it possible for systems and stakeholders to share data safely and effectively. The shortcomings of centralized data silos, which are frequently dispersed, ineffective, and difficult to combine, can be mitigated with the aid of DDS [68]. Organizations may better interact with partners, consumers, and stakeholders and realize the full potential of their data resources by implementing DDS.

DDS is a technique that allows for more security, privacy, and transparency by distributing data processing over several nodes. For Dataspace 4.0, DDS has several advantages. Participants can access a greater variety of platforms and data sources through DDS, potentially increasing the quantity and quality of data accessible for analysis and decision-making. By allowing users to access and validate data from many sources, DDS can raise the quality of the data by increasing its correctness and completeness. Through encryption, access restrictions, and other safeguards against illegal access and manipulation, DDS can improve data security and privacy. By allowing participants to create explicit standards and procedures for data management and sharing and fostering accountability and transparency throughout the process, DDS can enhance data governance.

However, there are several limitations to using DDS for Dataspace 4.0. The technical complexity of DDS might necessitate cutting-edge tools like edge computing, distributed ledgers, and semantic web technologies. DDS development and deployment may be complex due to their complexity. A network of participants is needed for DDS for data to be contributed and shared. This implies that the more people join the network, the more valuable it becomes, but it can also make it harder for smaller or less connected businesses to join. Data governance issues arise from DDS, such as determining who owns and controls the data and developing clear policies and procedures for sharing it. DDS can lead to difficulties with data governance, including figuring out who owns and controls the data and creating explicit guidelines and protocols for sharing it. The capacity to share and integrate data from many sources may be limited by DDS, which can lead to difficulties with interoperability across various platforms and technologies.

D. DEVELOPMENTS OF DDS IN DATASPACE 4.0

With several technical advancements, DDS has been the subject of intense research and development in recent years. It takes methods to make DDS safe and effective. Creating DDS protocols and platforms, such as P2P networks, distributed ledger technologies, and blockchain-based systems, has been a crucial field of study. With features like data encryption, access limits, and consensus processes to guarantee data integrity, the platforms and protocols let users share and access data safely and decentralizedly. Moreover, the advancements in edge computing and ML make decentralized data-sharing possible. By processing and analyzing data closer to the data source, edge computing lowers communication costs and latency and facilitates real-time decision-making. By creating models for learning from decentralized data sources, ML techniques allow users to use the collective wisdom of the network.

Another central area of research in DDS is semantic web technologies [69]. Semantic web technologies facilitate data integration and interoperability across many platforms and systems by enabling data to be represented and comprehended. Studies on the governance and incentive structures of DDS have also been conducted. Token economies, reputation systems, and other incentive mechanisms are among the strategies to incentivize users to submit data and preserve the network's integrity. Secure, effective, and cooperative data-sharing over decentralized networks is made possible by the most recent advancements and research in DDS for Dataspace 4.0. However, many issues still need to be resolved, including data governance, scalability, and interoperability, necessitating continued study and development.

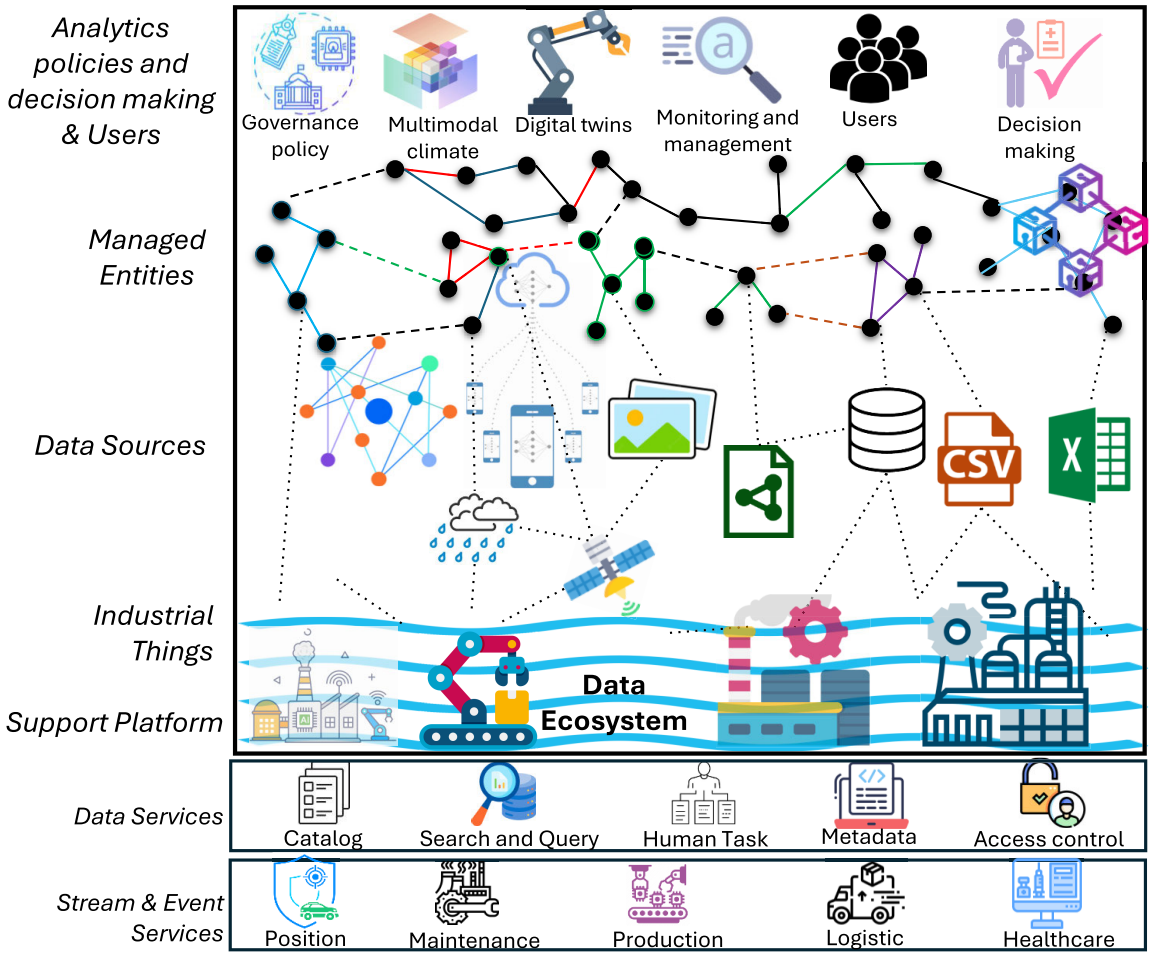

FIGURE 3. Empowering dataspace 4.0 with DSS technologies framework.

IV. ENABLING TECHNOLOGIES FOR DDS IN DATASPACE 4.0

DDS is a method in which several parties share control over the data instead of it being centralized in one place. Increased security, privacy, transparency, resilience, and interoperability are just a few advantages of this strategy. DDS is essential to tackling the difficulties of handling the enormous volumes of data produced by the IoT, AI, and other developing technologies in the framework of Dataspace 4.0, which refers to the data-sharing and management systems as shown in Figure 3. As a result, Dataspace 4.0 is distinguished by a distributed design that permits decentralized data processing and storing, fostering increased creativity and cooperation. DDS is possible with several technologies, including FL, P2P networks, and blockchain.

A. BLOCKCHAIN

Blockchain is a well-known distributed and DDS solution with solid security and openness characteristics, functions as a ledger, with data divided into sections and linked in a chronological and unchangeable order. A network of nodes employs a consensus process to authenticate and approve a sequence of transactions stored in each block, as shown in Figure 4. Blockchain is a dependable solution for data recording and sharing because of the procedure that guarantees the ledger's correctness and timeliness [70], [71], [72].

Several things need to be considered to develop a blockchain-based system for DSS. First and foremost, the system must be highly scalable and performant to handle massive amounts of data and transactions. Secondly, the integrity and confidentiality of the data depend on strong security measures like encryption and tamper-proofing. Lastly, the system has to be adaptable to accommodate a range of use cases and applications, including digital identity verification, supply chain management, and financial transactions [73]. A vital role in the development of data-sharing systems like Dataspace 4.0 is played by blockchain technology because of its fundamental requirements and potential, which open the door to the creation of highly transparent, safe, and effective DDS platforms [74], [75].

Blockchain technology offers several advantages for DSS. By offering a safe and impenetrable method of storing and exchanging data, blockchain improves the security of DSS. Cryptographic methods safeguard data on blockchains, making it nearly hard to change or corrupt the data without being

discovered. Furthermore, proof-of-work and proof-of-stake, two sophisticated consensus techniques, may be used by blockchain-based systems to guarantee that only authorized users can see and alter the data. Blockchain enhances privacy in DDS by allowing data owners to retain control over their data and selectively offer access to authorized parties. Blockchain-based systems encrypt and decode data using cryptographic methods like public and private keys, which makes it harder for unauthorized parties to access confidential information. Data encryption and decryption are managed using cryptographic techniques like public and private keys, making it very difficult for unauthorized people or entities to access sensitive information [74]. Consequently, blockchain technology's intrinsic security and privacy features highlight its applicability and significance in creating data-sharing frameworks for cutting-edge technologies like Dataspace 4.0.

With a standard data format and cryptographic techniques, blockchain technology provides a unified foundation that enables secure and efficient system-to-system interaction [76], [77]. This dramatically improves interoperability in DSS. Blockchain-based technologies preserve data integrity while enabling safe data interchange across enterprises and systems. Another essential component of blockchain is transparency, which guarantees that all users have an extensive and shared understanding of the data and its history. Using a distributed ledger to store data allows for transparency and provides an immutable, transparent record of all transactions and data modifications [78]. The feature encourages responsibility and compliance while enhancing confidence amongst the involved parties. For DDS to work, blockchain technology's durability is essential. Data availability is ensured despite system failures or network outages because of its decentralized and distributed structure consisting of several nodes. This resilience is strengthened by sophisticated consensus processes, which lessen the system's vulnerability to hostile assaults and data breaches [79].

The inventive potential of blockchain technology is a crucial factor in the evolution of 3DS4.0. The main characteristics of blockchain, such as its transparency and security, make it a valuable platform for exchanging and storing sensitive data [76]. Blockchain can securely store and distribute data by generating tamper-proof, decentralized ledgers, which reduces the possibility of unwanted access or manipulation. One domain in which blockchain technology has demonstrated immense use is supply chain management, a crucial facet of Industry 4.0. The safe and transparent tracking of products via blockchain across the supply chain substantially lowers the risk of fraud and guarantees regulatory compliance [80]. Additionally, because blockchain offers real-time tracking of commodities and eliminates the need for intermediaries, it is far more efficient and reduces costs [81]. Medical data security and privacy are critical for patient confidentiality in the healthcare sector. By offering a private and secure platform for handling and storing health data, blockchain technology allays these worries [82]. Additionally, blockchain technology makes it easier for healthcare professionals to share medical data. Collaboration and patient outcomes can be improved through data-sharing [83].

Blockchain technology has been widely applied in many sectors and use cases, such as voting systems, supply chain management, digital identity verification, and financial services. Blockchain enhances efficiency, accountability, and transparency by offering a transparent and safe platform for data recording and sharing across all situations [76]. With its enticing features like strong security, transparency, and decentralization, blockchain is a promising technology for DDS that may be used for various purposes [78]. To fully realize blockchain's potential for DDS,. However, certain limitations and challenges must still be addressed despite the potential benefits [84]. These include scalability issues, the high energy consumption of some consensus mechanisms, and the requirement for legal and regulatory frameworks for blockchain applications. The challenges, prospects, and future directions of blockchain-based DDS are outlined in Table 7 and Figure.5.

B. DECENTRALIZED FILE SYSTEM

A decentralized file system (DFS) is a type of file system that works without the need for a central server, distributing files among several network nodes or devices. DFS is a revolutionary data management and storage approach that disperses files over several network nodes or devices instead of relying on a single central server. This distributed approach has several advantages over traditional centralized file systems, such as increased data availability, fault tolerance, and resistance to censorship and single points of failure. Decentralized file systems can provide people and organizations more control over their data as technology develops and promotes an open and robust internet infrastructure. Decentralized file systems are appropriate for DSS. One network and protocol aiming to build a distributed file system that enables users to share and access data decentralized is the InterPlanetary File System (IPFS). It uses a content-addressable system, which guarantees that identical files are only saved once and allows for deduplication by identifying files based on their content rather than their location.

DFS stores files across several network nodes by dividing them into smaller chunks. This distribution ensures redundancy and fault tolerance as data loss does not occur from the loss of a single node. Data replication systems, which store several copies of the same file on different nodes, are widely used by DFSs. Data availability is ensured by this redundancy, even in node failures. Cryptographic techniques are employed to guarantee data security and integrity. Files are regularly checked and encrypted to guard against tampering and unwanted access. Content addressing, which recognizes and retrieves files based on their content rather than their location, is used by several DFS. This technique makes content distribution more effective by doing away with the need for preset file paths. P2P architecture,

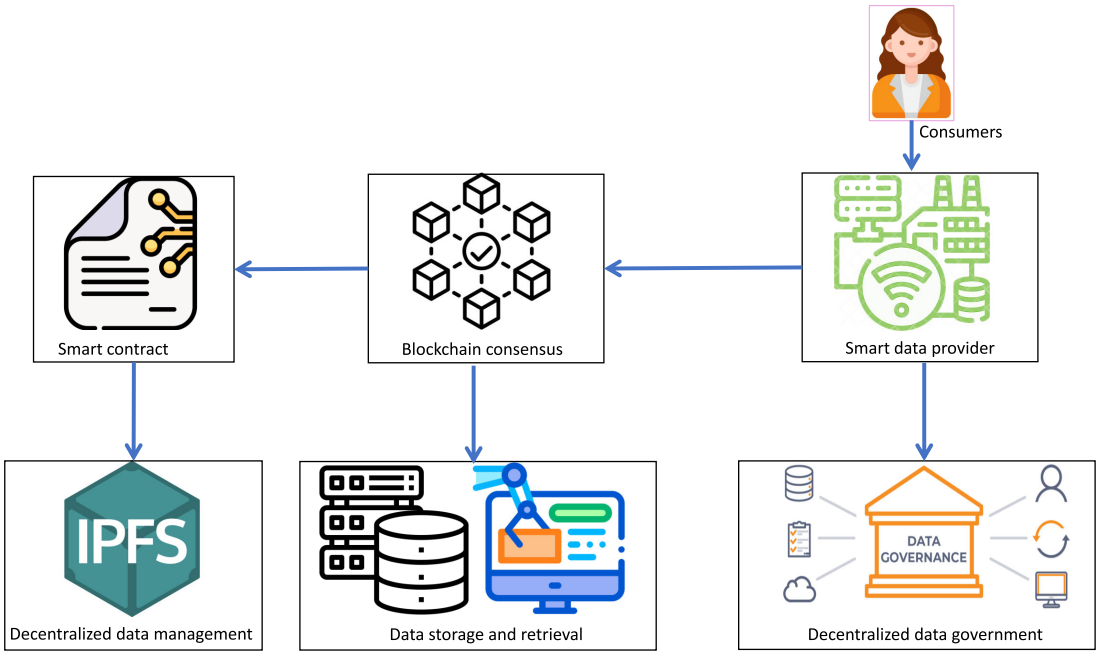

FIGURE 4. DDS of dataspace 4.0 using blockchain.

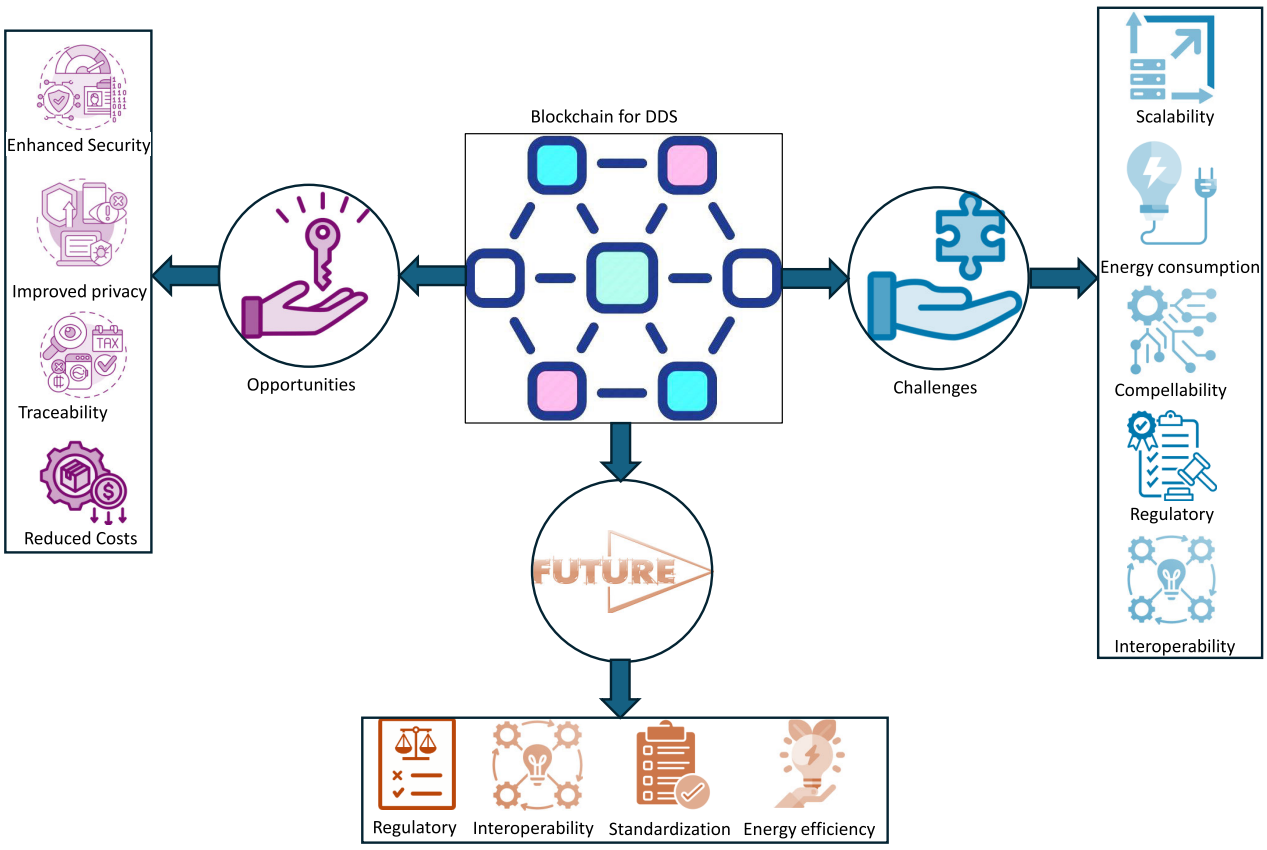

FIGURE 5. Blockchain for DDS challenges, opportunities and future directions.

which allows network nodes to interact and share data directly, is a common feature of DFSs. This makes a central server unnecessary and enables direct data communication between nodes.

TABLE 7. Challenges and opportunities of DDS using blockchain.

Items	Issue	Description
Challenges	Scalability	As the number of users and transactions on the blockchain network increases, the processing time and costs also increase, leading to slow and congested networks. This can limit the usability and effectiveness of blockchain for large-scale data-sharing applications.
	High energy consumption	In blockchain mining and validation, the consensus mechanism used in blockchain networks requires many nodes to validate transactions, which requires significant computational power and energy. High energy consumption can make blockchain less environmentally sustainable than other decentralized technologies and limit scalability.
	Complexity	Blockchain technology is still relatively new and complex, making it challenging to implement and use effectively, leading to a lack of adoption and hindering the growth of blockchain-based DDS.
	Regulatory	As blockchain-based data-sharing systems often operate in a decentralized and borderless environment, these systems can pose regulatory challenges for governments and organizations. Therefore, adopting blockchain-based solutions for DDS leads to uncertainty and hesitation.
	Standardization and interoperability	Standardization and interoperability between different blockchain networks can make integrating and data-sharing between different systems challenging. This challenge can limit the potential of blockchain for DDS and hinder collaboration between different organizations and networks.
Opportunities	Enhanced Security	Blockchain's distributed ledger technology can provide enhanced security by ensuring data immutability, transparency, and accountability.
	Improved Privacy	DDS solutions can help to improve privacy by enabling users to control their data and only share it with authorized parties.
	Traceability	Blockchain can provide a traceable audit trail for data, improving supply chain efficiency and enabling better data usage tracking.
	Reduced Costs	DDS can reduce operational costs by eliminating the need for intermediaries and enabling direct P2P transactions.
Future Directions	Interoperability	Efforts are being made to develop interoperability solutions between blockchain networks to facilitate seamless data-sharing across multiple systems.
	Standardization	Standardization of blockchain protocols can improve the scalability and interoperability of DDS solutions.
	Energy efficiency	Research and development are being conducted to create more energy-efficient blockchain networks to reduce environmental impact and operational costs.
	Regulatory frameworks	Developing regulatory frameworks can create a conducive environment for blockchain technology and facilitate the adoption of DDS solutions.

TABLE 8. Merits of decentralized file systems.

Advantages	Descriptions
Data Availability and Reliability	By distributing files across multiple nodes, decentralized file systems ensure high data availability, even if some nodes go offline or become inaccessible. This redundancy increases reliability and reduces the risk of data loss.
Censorship Resistance	Decentralized file systems are resistant to censorship since there is no central authority controlling access to the files. This makes it difficult for any entity to block or remove specific files or content.
Scalability	DFSs can scale efficiently by adding more nodes to the network. As the number of nodes increases, the storage capacity and overall performance of the system can improve.
Cost-Effectiveness	Since decentralized file systems utilize resources from participating nodes in the network, they can reduce the need for dedicated infrastructure and costly data centers. This can result in lower storage and maintenance costs.
Data Privacy	DFSs can offer enhanced data privacy since files are often encrypted and stored across multiple nodes. This reduces the risk of unauthorized access or data breaches.

C. SEMANTIC WEB AND KNOWLEDGE REPRESENTATION

Semantic Web (SW) and Knowledge Representation (KR) refers to technologies allowing the sharing and integration of data and knowledge on the web. These technologies provide a standardized way of representing and describing data, making it possible to automate the processing and integration of information from multiple sources. The Semantic Web is an expansion of the World Wide Web that allows machines to share, connect, and understand data. It seeks to give information semantic meaning, allowing computers to perceive and process it more intelligently. Knowledge Representation, on the other hand, is the act of encoding and organising knowledge in a form that computing systems can use. This can be particularly useful in DDS for Dataspace 4.0, as it can help to ensure that data is accurately understood and interpreted by all participants.

Knowledge representation is the basic building block of Semantic Web. It plays critical roles predominantly defining data semantics and intelligent reasoning. Several technologies and standards have been created to aid in the implementation of the Semantic Web. The Resource Description Framework (RDF) provides a framework for defining online resources and their interactions. Web Ontology Language (OWL) enables the building of ontologies that specify ideas, connections, and restrictions. OWL is built around a description logic, which is a sub-language of first-order predicate logic that uses only unary and binary predicates and a limited usage of quantifiers and is structured in such a way that logical deductive reasoning over the language is possible [85]. Even after the standard was published, the community debated whether description logics were the best paradigm option, with rule-based languages a strong rival [86]. Although the debate was finally resolved, the Rule Interchange Format RIF [85], which was subsequently adopted as a rule-based W3C standard, received little traction. SPARQL is a query language for retrieving information from RDF databases. These technologies collaborate to enable the representation, integration, and querying of semantic data.

Linked Data [87] is a major driver for Semantic Web applications and persist as such until the early 2010s. What is commonly associated with the phrase ''Linked Data'' is that it comprises of a (now fairly extensive) group of RDF graphs that are ''linked'' in the sense that many IRI identifiers

TABLE 9. Challenges and opportunities of DDS using DFS.

Items	Issue	Description
Challenges	Performance	When opposed to centralized systems, the decentralized structure of file systems might impose extra delay and slower data transfer speeds. When retrieving files, data may be fetched from many nodes, which can have an influence on speed, especially for bigger files.
	Data Consistency	Maintaining good data integrity across a decentralized file system can be difficult. When several nodes store and replicate data, there may be delays in updating or conflicts that must be resolved.
	Storage and Bandwidth Requirements	To store and distribute files, decentralized file systems sometimes necessitate large storage and bandwidth resources from participating nodes. For devices with limited resources or slower network connections, this might be a hindrance.
	Reliability	The network's dependability and availability can have an influence on the overall performance and accessibility of decentralized file systems. Data availability and retrieval can be impacted if a large number of nodes go unavailable or the network encounters disturbances.
Opportunities	Data Ownership and Control	User control and data ownership are increased through decentralized file systems. Users may store and manage their files directly on the network, without the need for centralized organizations, giving them complete control over data-sharing and access permissions.
	Censorship Resistance	Due to their distributed structure, decentralized file systems make it challenging for any one body to restrict or control material. This might promote more freedom of expression and stop data censorship.
	Fault Tolerance and Resilience	Decentralized file systems distribute data across multiple nodes, ensuring redundancy and fault tolerance. If a node goes offline or fails, the data can still be accessed from other nodes, providing resilience to the system.
	Collaboration and Data-sharing	File sharing and collaborative workflows are made easier by decentralized file systems. By sending the distinct content address or cryptographic hash, users may effortlessly share files with others. This makes cooperation easier and does away with the need for complicated authorization systems or central servers.
Future Directions	High Scalability and Performance	Research is needed to improve the scalability and performance of decentralized file systems, such as optimizing data storage and retrieval algorithms, reducing latency, and exploring techniques to handle large-scale deployments and high-volume data transfer efficiently.
	Data Integrity	Consistency models should be developed to ensure strong data consistency, incentive mechanisms should be explored to encourage cooperation, and privacy and security mechanisms should be enhanced.
	Interoperability	Interoperability and standardization should be explored to facilitate seamless integration across different systems and networks.
	Dynamic Network Environment	The system should be adaptable to dynamic network conditions, ensuring reliability and performance in changing environments.
	Integration	The integration between decentralized file systems and blockchain technologies should be explored, leveraging their complementary features.
	Sustainability	Distributed file systems consume more energy than centralized systems as they include more resources (e.g., computing nodes). Energy efficiency should be explored to mitigate environmental impact, considering factors such as implementation, number of nodes, algorithms used, and consensus mechanisms.
	Ease of Use	Usability and user experience should be improved by developing user-friendly interfaces, data discovery mechanisms, and intuitive tools.

in the graphs also exist in other, and sometimes multiple, graphs. In some ways, the collection of all these connected RDF graphs may be thought of as one massive RDF graph.

Google launched the Knowledge Graph in 2012 [88]. knowledge graph technology has found a prominent role in business, including key information technology companies other than Google, such as Microsoft, IBM, Facebook, and eBay. However, given the history of Semantic Web technologies, particularly linked data and ontologies, it appears that knowledge graph is mostly a new framing of ideas that originated in the Semantic Web sector, with, of course, some noteworthy variations in emphasis. Knowledge graphs are often seen as considerably more internally consistent and carefully managed objects. As a result, the usefulness of ''external links,'' i.e., to external graphs without strict quality control, is questioned25, while the quality of content and/or the underlying schema is emphasized.

In the context of DDS for Dataspace 4.0, the use of semantic web technologies can help address the challenges of data integration and interoperability, making it easier for participants to share and use data from multiple sources. This can enable more accurate and comprehensive data analysis and decision-making and support the development of new applications and services.

The Semantic Web and Knowledge Representation have a bright future with plenty of room for growth. One approach is to combine ML techniques with semantic data to improve automated reasoning and knowledge extraction. Extending the usage of ontologies and vocabularies across domains and sectors can help to improve interoperability. Furthermore, solving scalability issues and encouraging the use of Semantic Web technologies will be critical in achieving the full potential of these techniques.

Future research directions in this area also include the development of more advanced ontologies and knowledge representation techniques and exploring new approaches to integrating semantic web technologies with other advanced technologies such as blockchain, edge intelligence, and FL. Additionally, the research could focus on developing new applications and services that leverage semantic web technologies to support DDS in Dataspace 4.0.

D. FEDERATED LEARNING

With FL, multiple parties may jointly train a shared ML model while protecting the confidentiality and privacy of their data [89]. FL is a revolutionary approach to ML. By storing data locally on each device or network, FL participants do away with the necessity to transmit private information to a central repository. A significant development in privacy-preserving machine learning, FL processing is essential in many industries where data security is vital. More on how the FL procedure operates: After training an ML

model on their local data, each participant transmits the model updates—the variations in the model parameters—to a central server. The server combines all participants' updates [90]. Depending on the configuration, the aggregation procedure may entail applying a more sophisticated algorithm or calculating an average or median. After collecting changes, the central server delivers each participant's new global model (updated with the aggregated parameters).

Until the global model meets specific stopping requirements or reaches an acceptable degree of accuracy, the steps of local training, updating the server, and receiving the updated model from the server are repeated several times. Many individuals can participate in the ML process without disclosing their raw data because of FL's unique structure. A more accurate global model results from the procedure being performed consistently thanks to the common ML model that all participants utilize [91], [92]. However, for FL to work well, there has to be a vital communication infrastructure. This infrastructure is required to effectively share the ML models between participants and the central server. A weakened learning process and slower model updates might result from inadequate infrastructure or poor connectivity.

For example, the decentralized infrastructure layer has three nodes (Nodes 1, 2, and 3; as shown in Figure 6). The FL Framework layer represents an instance of the FL framework unique to each node. Through the decentralized architecture, the nodes interact with one another to work together on training a MLmodel with their local data while protecting the privacy and security of the data utilizing FL methods.

By storing sensitive data locally, FL significantly improves data security and privacy and reduces the possibility of data breaches during transmission [90]. Potential attackers cannot readily access sensitive data since the raw data remains on the local device. They would have to hack several devices to reconstitute the entire dataset, which is significantly more complex than just focusing on one centralized server. Furthermore, FL's design suggests that the model updates—which usually involve aggregated and abstracted data—are the only ones delivered to the central server. The updates further preserve data privacy by avoiding disclosing the raw data generated from [91]. FL guarantees user privacy by refusing to share user data with a central server. The information stays local to each user's device, so third parties—including the organizations running the learning process—cannot directly access it. To strengthen privacy even more, FL uses methods like differential privacy [93]. Differential privacy makes it challenging to re-identify people based on the shared information by adding statistical noise to the data or the model updates. FL makes it easier for disparate entities to collaborate and creates a single model without requiring direct data transfer. The FL function is handy in industries like healthcare, banking, and telecommunications, where direct data interchange is restricted by stringent data protection requirements [94].

To ensure that no sensitive data is disclosed throughout the model-building process, transparency is also essential for fostering trust between cooperating parties [95]. FL improves the system's robustness in several ways, contributing to its resilience. For example, FL enables efficient data use even in settings with spotty network access [96], [97], [98]. This is achieved by having most of the computation done locally on the edge devices, requiring only sporadic network connectivity to transfer aggregated model changes. Furthermore, FL is built to withstand data corruption and device malfunctions with resilience. Due to its reliance on several devices that carry out local calculations, the FL process is not significantly disrupted if a device becomes unavailable or its data is damaged. Due to this redundancy, FL models are far more reliable and continue to work even in the worst situations [91]. The combination of blockchain and FL plays a vital role in improving security and privacy [5], [6], [99], [100]. However, the challenges and opportunities shows in Table. 10.

E. P2P NETWORKS

DDS is made possible via P2P networks with a decentralized network design. Every node in a P2P network is equal and in charge of sending and receiving data. The robustness of P2P networks to single points of failure is a plus for DDS. The availability of the data is unaffected by the failure of a single node because there is no central server. Because P2P networks eliminate the need for middlemen and enable direct data-sharing between users, they significantly increase privacy and security [33]. P2P networks protect data in transit and guarantee that only authorized users can access it using encryption and authentication procedures. P2P networks can allow consumers more control over their data regarding privacy. For instance, P2P networks enable users to store and data-sharing directly and decentralized with other users instead of storing it on centralized servers managed by a single body. As a result, users have control over what information they disclose and with whom. They can also ensure that their information is only accessed or utilized with their permission.

P2P networks benefit significantly from being resilient to single points of failure. The lack of a centralized server allows the network to function even if a node fails, guaranteeing data availability [101]. P2P networks place much emphasis on security. These systems ensure that only authorized users may access specific data using encryption and authentication mechanisms to safeguard data in transit and verify user identities. This decentralization further improves data security by doing away with the need for middlemen. P2P networks provide consumers more control over their data in terms of privacy. Users can keep data on their devices and distribute it directly over P2P networks instead of having it stored on centralized servers under the control of a single party. Users can choose what information they disclose and with whom,

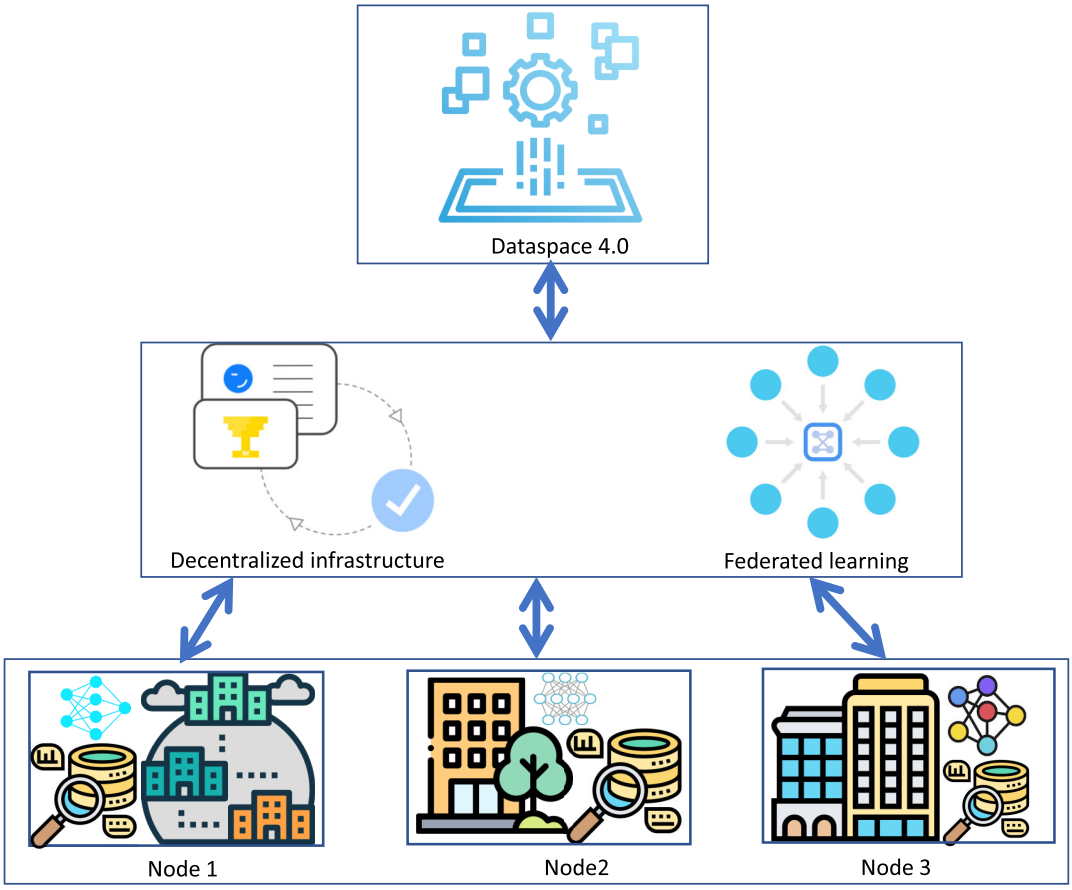

FIGURE 6. DDS of dataspace 4.0 using FL.

TABLE 10. Challenges and opportunities of DDS using FL.

Items	Issue	Description
Challenges	Computational Resources	Limited computational resources on some devices can restrict participants' ability to train complex machine-learning models.
	Data Heterogeneity	Different participants may have varying data types and quality, which makes it challenging to build an accurate and representative shared machine-learning model.
	Privacy Concerns	FL does not eliminate the risk of exposing sensitive data. Updated model parameters sent to the central server may still contain some information about local data, raising concerns about data privacy and security.
	Communication Overhead	FL requires frequent communication between participants and the central server, which can lead to high communication overhead and latency.
	Lack of Standardization	Insufficient standardization and interoperability between different FL systems make integrating and sharing models challenging.
	Model Convergence	FL requires careful tuning of hyperparameters to ensure model convergence.
Opportunities	Collaboration	FL allows multiple organizations to collaborate and train models on decentralized data without sharing raw data, enabling them to leverage each other's expertise.
	Privacy Preservation	FL enables the training of models while preserving data privacy, making it ideal for industries where data privacy is paramount.
	Resource Utilization	FL enables the efficient use of resources across multiple organizations, reducing the need for each organization to have dedicated infrastructure and personnel.
Future Trends	FL for Heterogeneous Data	Future research can focus on developing techniques to handle heterogeneous data effectively in FL.
	Privacy and Security Enhancements	Research can focus on developing new methods and protocols to enhance the security and privacy of FL models.
	Communication Optimization	Future research can explore ways to optimize communication overhead in FL, such as through compression techniques or edge devices.
	FL for Dynamic Environments	Future work can explore using FL in dynamic environments, where data distribution may change over time.

adding a degree of permission and preventing unauthorized access to or use of their data [102].

The power of P2P networks is to promote interoperability among various applications and systems. P2P networks facilitate data-sharing and communication across systems, even ones that may otherwise be incompatible or isolated, by relying on open technologies and standard protocols. P2P networks facilitate cooperation between stakeholders

TABLE 11. Challenges and opportunities of DDS using P2P networks.

Items	Issue	Description
Challenges	Limited Scalability	P2P networks can struggle to scale up to handle large amounts of data and users.
	Lack of Standardization	Different P2P protocols may complicate efforts to establish standards for interoperability.
	Security Risks	P2P networks can be vulnerable to security threats, including denial-of-service attacks, distributed attacks, and Sybil attacks.
	Dependence on Network Connectivity	P2P networks rely on connectivity between nodes, which can be affected by unreliable or slow network connections.
Opportunities	Enhanced Privacy	P2P networks allow users to share data directly without intermediaries, thus reducing the risk of data breaches and enhancing user privacy.
	Improved Data Access	P2P networks facilitate data sharing and access by providing a distributed system for storing and retrieving data.
	Decentralized Control	P2P networks offer a decentralized approach to data sharing, enabling users to control their data and its sharing mechanisms.
	Cost Savings	By leveraging P2P networks, organizations can reduce costs associated with centralized data storage and maintenance.
Future Trends	Integration with Other Technologies	P2P networks can be combined with technologies such as blockchain and federated learning (FL) to enhance the robustness of decentralized data-sharing systems.
	Standardization Efforts	To improve interoperability, efforts can be made to standardize P2P protocols and establish best practices for deployment.
	Scalability	Continued research is needed to improve the scalability of P2P networks, making them suitable for large-scale data-sharing applications.
	Improved Security Measures	To mitigate security risks, P2P networks can integrate additional security measures, such as encryption, to protect shared data and improve their suitability for sensitive applications.

TABLE 12. Comparative analysis of enabling technologies for DDS in dataspace 4.0.

Technology	Functionality	Strengths	Limitations	Applications	Security & Privacy	Scalability
FL	Collaborative training of ML models without sharing raw data	Enhances data privacy by keeping data local. Improves model accuracy through collaboration	Requires complex coordination and communication. Vulnerable to data heterogeneity issues	Healthcare, financial services, and other data-sensitive industries	Strong privacy preservation through local data. Requires secure aggregation techniques	Scalable concerning data size. Limited by communication overhead
Blockchain	Decentralized ledger for secure and transparent transactions	Ensures data integrity and immutability. Enhances transparency and trust	High energy consumption. Scalability issues with increasing data volume	Supply chain management, digital identity verification, and financial transactions	High security through cryptographic techniques. Potential privacy concerns with public ledgers	Scalability challenges with transaction processing speed and data size
DFS	Secure and resilient data storage and sharing	Reduces risk of data loss and breaches. Offers high availability and fault tolerance	Performance can be affected by network latency. Complex implementation and management	Financial services, cloud storage, and content distribution	Secure data storage. Risk of unauthorized access if not properly managed	Scalable for storage. May face latency issues
SW and KR	Enhances data interoperability and knowledge sharing through standardized data formats	Improves data integration from diverse sources. Enables sophisticated querying and reasoning	Requires extensive ontology development. Can be computationally intensive	Academic research, enterprise knowledge management, and smart cities	Privacy depends on data and ontology design. Security through access controls	Scalable with efficient data management. Computationally intensive
P2P Networks	Decentralized communication and resource sharing among peers	Reduces reliance on central servers. Enhances network resilience and efficiency	Security risks from malicious peers. Difficult to manage and maintain	IoT networks, decentralized communication platforms, and content sharing	Inherent risks from peer interactions. Security measures needed to protect data integrity	Scalable with increased peers. Requires robust protocols to manage peer interactions

and organizations by removing these data barriers [103]. Another crucial component of P2P networks is transparency. It becomes more difficult for any one node to change or alter the data without being discovered as it is dispersed over several nodes within the network. P2P networks can offer more transparency and accountability than typical client-server systems, where the data may be controlled by a single organization [104]. Furthermore, P2P networks are resilient because of their decentralized structure. Decentralized P2P networks allow data to be saved and retrieved from several places to share and store data, even if one node fails or goes down. The approach enhances the system's resilience by guaranteeing that crucial data is still available even during network disturbances or outages [105]. The challenges and opportunities of P2P are shown in Table. 11.

To meet the requirements for DSS, P2P networks are essential. P2P networks offer a distributed network design in which data-sharing and receiving are the responsibilities of every node with an equal status. The decentralized method has several benefits that meet the needs of DDS. P2P networks reduce single points of failure, which improves security. In conventional centralized systems, P2P can result in a massive data breach if a central server is compromised. P2P networks, on the other hand, spread data over several nodes, lowering the possibility of unwanted access. Data may be authenticated and encrypted by any node in the

network, guaranteeing that only authorized users can access it [106].

P2P networks are spread, which increases security by making it harder for bad actors to access all the data. Second, by giving users more control over their data, P2P networks support privacy. Users can directly data-sharing in a decentralized way with other trusted users rather than depending on a central authority to govern access to the data [107], [108]. Decentralized gives people the freedom to decide what information to share and with whom, protecting their security and privacy. Additionally, P2P networks use open standards and protocols to increase interoperability. Systems and applications that may be incompatible or compartmentalized can communicate and share data more easily thanks to these networks [109]. P2P networks facilitate smooth communication between stakeholders and organizations by removing obstacles to data, which boosts productivity and effectiveness.

In addition, P2P networks provide more transparency than conventional client-server topologies. When dispersed throughout several nodes, it becomes more difficult for a single node to change or alter the data covertly. Among network participants, transparency fosters responsibility and trust [110]. P2P networks also improve resilience by guaranteeing data availability during network outages or node failures. Data availability is not affected by the failure of a single node since it is stored and accessible from several places [111]. Maintaining continuous data-sharing activities and reducing the risks associated with single points of failure requires resilience.

V. COMPARATIVE ANALYSIS OF ENABLING TECHNOLOGIES IN DATASPACE 4.0

Within the framework of Dataspace 4.0, we present a comparison of the leading technologies associated with DDS. In this section, we emphasize each technology's features, advantages, disadvantages, and theoretical components, including interoperability, security, privacy, and scalability, as shown in the Table. 12. The comparison provides fresh perspectives on these technologies' real-world uses and future directions and synthesizes the body of available knowledge. In addition to improving knowledge of each technology's capabilities, we direct the choice and integration of the best technologies for Dataspace 4.0 use cases.

VI. CONCLUSION

In the rapidly evolving digital transformation landscape, DDS is a significant factor in solving data security, privacy, accessibility, and interoperability concerns. The results of the survey demonstrate that DDS, mainly when utilized in Dataspace 4.0, offers several benefits over traditional centralized data management solutions. Through improved data security and privacy, as well as increased accessibility, interoperability, and knowledge exchange, DDS empowers individuals and businesses to have more ownership and control over their data. The leading DDS-supporting technologies—FL, blockchain, decentralized file systems, semantic web and knowledge representation, and P2P networks—have been described together with their benefits, drawbacks, and potential applications. When combined, these technologies drive the innovations needed to implement DDS in Dataspace 4.0. We have also identified critical challenges that must be addressed, like maintaining decentralized networks, ensuring data security and privacy, balancing privacy with utility in FL, and addressing sustainability and energy efficiency. In Dataspace 4.0, DDS has a bright future but still needs ongoing innovation and development. Building DDS protocols and robust governance frameworks, inventing new algorithms and techniques, boosting accuracy and efficiency, and emphasizing sustainability and energy efficiency are all crucial challenges. In addition to providing a comprehensive overview of DDS's current state, the poll draws attention to the critical potential and difficulties that lie ahead for improving data-sharing in the digital era regarding compatibility, efficiency, and safety.

REFERENCES

[1] *Dataspace 4.0*. Accessed: Jun. 23, 2023. [Online]. Available: https://manufacturingdataspace-csa.eu/
[2] S. H. Alsamhi, E. Curry, A. Hawbani, S. Kumar, U. U. Hassan, and N. S. Rajput, "Dataspace in the sky: A novel decentralized framework to secure drones data sharing in B5G for industry 4.0 toward industry 5.0," Tech. Rep., 2023.
[3] R. Tallat, A. Hawbani, X. Wang, A. Al-Dubai, L. Zhao, Z. Liu, G. Min, A. Y. Zomaya, and S. Hamood Alsamhi, "Navigating industry 5.0: A survey of key enabling technologies, trends, challenges, and opportunities," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 2, pp. 1080–1126, 2nd Quart., 2024.
[4] S. Peng, N. Kumar, S. H. Alsamhi, Q. He, and L. Zhao, "Securing IoT data: FDUP-RDIC—A fully decentralized approach for privacy-preserving and efficient data integrity," *IEEE Internet Things J.*, early access, Jun. 11, 2024, doi: 10.1109/JIOT.2024.3412224.
[5] S. H. Alsamhi, R. Myrzashova, A. Hawbani, S. Kumar, S. Srivastava, L. Zhao, X. Wei, M. Guizan, and E. Curry, "Federated learning meets blockchain in decentralized data sharing: Healthcare use case," *IEEE Internet Things J.*, vol. 11, no. 11, pp. 19602–19615, Jun. 2024.
[6] R. Myrzashova, S. H. Alsamhi, A. Hawbani, E. Curry, M. Guizani, and X. Wei, "Safeguarding patient data-sharing: Blockchain-enabled federated learning in medical diagnostics," *IEEE Trans. Sustain. Comput.*, early access, Jun. 4, 2024, doi: 10.1109/TSUSC.2024.3409329.
[7] M. A. Elaziz, M. A. A. Al-Qaness, A. Dahou, S. H. Alsamhi, L. Abualigah, R. A. Ibrahim, and A. A. Ewees, "Evolution toward intelligent communications: Impact of deep learning applications on the future of 6G technology," *WIREs Data Mining Knowl. Discovery*, vol. 14, no. 1, Jan. 2024, Art. no. e1521.
[8] M. H. Alsamh, A. Hawbani, S. Kumar, and S. H. Alsamhi, "Multisensory metaverse-6G: A new paradigm of commerce and education," *IEEE Access*, vol. 12, pp. 75657–75677, 2024.
[9] S. Scerri and S. Augustin, "Industrial data space-digital sovereignty over data," in *Proc. Digitising Eur. Ind. WG2 Meeting*, vol. 8, Brussels, Belgium, 2016, pp. 1–40, doi: 10.13140/RG.2.1.2673.0649.
[10] S. Alansari, "A blockchain-based approach for secure, transparent and accountable personal data sharing," Ph.D. thesis, Dept. Fac. Eng., Sci. Math., School Electron. Comput. Sci., Univ. Southampton, Southampton, U.K., 2020.
[11] I. Jao, F. Kombe, S. Mwalukore, S. Bull, M. Parker, D. Kamuya, S. Molyneux, and V. Marsh, "Research Stakeholders' views on benefits and challenges for public health research data sharing in Kenya: The importance of trust and social relations," *PLoS ONE*, vol. 10, no. 9, Sep. 2015, Art. no. e0135545.
[12] T. White, E. Blok, and V. D. Calhoun, "Data sharing and privacy issues in neuroimaging research: Opportunities, obstacles, challenges, and monsters under the bed," *Hum. Brain Mapping*, vol. 43, no. 1, pp. 278–291, Jan. 2022.

[13] J. Sen, "Security and privacy issues in cloud computing," Innov. Labs, Tata Consultancy Services Ltd., Kolkata, India, 2009, doi: 10.4018/978-1-4666-4514-1.ch001.

[14] A. Torab-Miandoab, T. Samad-Soltani, A. Jodati, and P. Rezaei-Hachesu, "Interoperability of heterogeneous health information systems: A systematic literature review," *BMC Med. Informat. Decis. Making*, vol. 23, no. 1, p. 18, Jan. 2023.

[15] M. A. Uddin, A. Stranieri, I. Gondal, and V. Balasubramanian, "A survey on the adoption of blockchain in IoT: Challenges and solutions," *Blockchain, Res. Appl.*, vol. 2, no. 2, Jun. 2021, Art. no. 100006.

[16] Z. Ullah, B. Raza, H. Shah, S. Khan, and A. Waheed, "Towards blockchain-based secure storage and trusted data sharing scheme for IoT environment," *IEEE Access*, vol. 10, pp. 36978–36994, 2022.

[17] V. Neumann, G. Davidge, M. Harding, J. Cunningham, N. Davies, S. Devaney, G. Leeming, S. Holm, and J. Ainsworth, "Examining public views on decentralised health data sharing," *PLoS ONE*, vol. 18, no. 3, Mar. 2023, Art. no. e0282257.

[18] L. T. Nguyen, L. Duc Nguyen, T. Hoang, D. Bandara, Q. Wang, Q. Lu, X. Xu, L. Zhu, P. Popovski, and S. Chen, "Blockchain-empowered trustworthy data sharing: Fundamentals, applications, and challenges," 2023, *arXiv:2303.06546*.

[19] S. Athanere and R. Thakur, "Blockchain based hierarchical semi-decentralized approach using IPFS for secure and efficient data sharing," *J. King Saud Univ., Comput. Inf. Sci.*, vol. 34, no. 4, pp. 1523–1534, Apr. 2022.

[20] E. Curry, T. Tuikka, A. Metzger, S. Zillner, N. Bertels, C. Ducuing, D. D. Carbonare, S. Gusmeroli, S. Scerri, I. L. de Vallejo, and A. G. Robles, "Data sharing spaces: The BDVA perspective," in *Designing Data Spaces: The Ecosystem Approach To Competitive Advantage*. Cham, Switzerland: Springer, 2022, pp. 365–382.

[21] H. Niavis, N. Papadis, V. Reddy, H. Rao, and L. Tassiulas, "A blockchain-based decentralized data sharing infrastructure for off-grid networking," in *Proc. IEEE Int. Conf. Blockchain Cryptocurrency (ICBC)*, May 2020, pp. 1–5.

[22] C. F. L. Hickman, H. Alshubbar, J. Chambost, C. Jacques, C.-A. Pena, A. Drakeley, and T. Freour, "Data sharing: Using blockchain and decentralized data technologies to unlock the potential of artificial intelligence: What can assisted reproduction learn from other areas of medicine?" *Fertility Sterility*, vol. 114, no. 5, pp. 927–933, Nov. 2020.

[23] V. P. Chellapandi, A. Upadhyay, A. Hashemi, and S. H. Zak, "On the convergence of decentralized federated learning under imperfect information sharing," 2023, *arXiv:2303.10695*.

[24] S. Peng, W. Bao, H. Liu, X. Xiao, J. Shang, L. Han, S. Wang, X. Xie, and Y. Xu, "A peer-to-peer file storage and sharing system based on consortium blockchain," *Future Gener. Comput. Syst.*, vol. 141, pp. 197–204, Apr. 2023.

[25] J. Scheibner, J. L. Raisaro, J. R. Troncoso-Pastoriza, M. Ienca, J. Fellay, E. Vayena, and J.-P. Hubaux, "Revolutionizing medical data sharing using advanced privacy-enhancing technologies: Technical, legal, and ethical synthesis," *J. Med. Internet Res.*, vol. 23, no. 2, Feb. 2021, Art. no. e25120.

[26] A. Pliatsios, K. Kotis, and C. Goumopoulos, "A systematic review on semantic interoperability in the IoE-enabled smart cities," *Internet Things*, vol. 22, Jul. 2023, Art. no. 100754.

[27] M. I. Khalid, I. Ehsan, A. K. Al-Ani, J. Iqbal, S. Hussain, S. S. Ullah, and Nayab, "A comprehensive survey on blockchain-based decentralized storage networks," *IEEE Access*, vol. 11, pp. 10995–11015, 2023, doi: 10.1109/ACCESS.2023.3240237.

[28] C. Briggs, Z. Fan, and P. Andras, "A review of privacy-preserving federated learning for the Internet-of-Things," in *Federated Learning Systems: Towards Next-Generation AI*, vol. 965. Cham, Switzerland: Springer, 2021, pp. 21–50, doi: 10.1007/978-3-030-70604-3_2.

[29] Q. Wei, B. Li, W. Chang, Z. Jia, Z. Shen, and Z. Shao, "A survey of blockchain data management systems," *ACM Trans. Embedded Comput. Syst.*, vol. 21, no. 3, pp. 1–28, May 2022.

[30] T. V. Doan, Y. Psaras, J. Ott, and V. Bajpai, "Toward decentralized cloud storage with IPFS: Opportunities, challenges, and future considerations," *IEEE Internet Comput.*, vol. 26, no. 6, pp. 7–15, Nov. 2022.

[31] A. Rejeb, J. G. Keogh, W. Martindale, D. Dooley, E. Smart, S. Simske, S. F. Wamba, J. G. Breslin, K. Y. Bandara, S. Thakur, K. Liu, B. Crowley, S. Desaraju, A. Ospina, and H. Bradau, "Charting past, present, and future research in the semantic web and interoperability," *Future Internet*, vol. 14, no. 6, p. 161, May 2022.

[32] N. Masinde and K. Graffi, "Peer-to-peer-based social networks: A comprehensive survey," *Social Netw. Comput. Sci.*, vol. 1, no. 5, p. 299, Sep. 2020.

[33] G. Ding and B. Bhargava, "Peer-to-peer file-sharing over mobile ad hoc networks," in *Proc. 2nd IEEE Annu. Conf. Pervasive Comput. Commun. Workshops*, Mar. 2004, pp. 104–108.

[34] Y. Ye, L. Zhang, W. You, and Y. Mu, "Secure decentralized access control policy for data sharing in smart grid," in *Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, May 2021, pp. 1–6.

[35] M. Sultana, A. Hossain, F. Laila, K. A. Taher, and M. N. Islam, "Towards developing a secure medical image sharing system based on zero trust principles and blockchain technology," *BMC Med. Informat. Decis. Making*, vol. 20, no. 1, pp. 1–10, Dec. 2020.

[36] Y. Lu, X. Huang, Y. Dai, S. Maharjan, and Y. Zhang, "Blockchain and federated learning for privacy-preserved data sharing in industrial IoT," *IEEE Trans. Ind. Informat.*, vol. 16, no. 6, pp. 4177–4186, Jun. 2020.

[37] M. Stolpe, "The Internet of Things: Opportunities and challenges for distributed data analysis," *ACM SIGKDD Explor. Newslett.*, vol. 18, no. 1, pp. 15–34, Aug. 2016.

[38] R. E. Endeley, "End-to-end encryption in messaging services and national security—Case of WhatsApp messenger," *J. Inf. Secur.*, vol. 9, no. 1, pp. 95–99, 2018.

[39] B. Lashkari and P. Musilek, "A comprehensive review of blockchain consensus mechanisms," *IEEE Access*, vol. 9, pp. 43620–43652, 2021.

[40] A. Salehi Shahraki, C. Rudolph, and M. Grobler, "Attribute-based data access control for multi-authority system," in *Proc. IEEE 19th Int. Conf. Trust, Secur. Privacy Comput. Commun. (TrustCom)*, Dec. 2020, pp. 1834–1841.

[41] G. Habib, S. Sharma, S. Ibrahim, I. Ahmad, S. Qureshi, and M. Ishfaq, "Blockchain technology: Benefits, challenges, applications, and integration of blockchain technology with cloud computing," *Future Internet*, vol. 14, no. 11, p. 341, Nov. 2022.

[42] B. Otto and M. Jarke, "Designing a multi-sided data platform: Findings from the international data spaces case," *Electron. Markets*, vol. 29, no. 4, pp. 561–580, Dec. 2019.

[43] S. H. Alsamhi, A. A. F. Saif, E. Curry, S. Kumar, and A. Hawbani, "Autonomous multi-robot collaboration in virtual environments to perform tasks in industry 4.0," in *Proc. 2nd Int. Conf. Emerg. Smart Technol. Appl. (eSmarTA)*, Oct. 2022, pp. 1–7.

[44] M. Soori, B. Arezoo, and R. Dastres, "Internet of Things for smart factories in industry 4.0, a review," *Internet Things Cyber-Phys. Syst.*, vol. 3, pp. 192–204, Jan. 2023.

[45] S. H. Alsamhi, A. V. Shvetsov, S. Kumar, J. Hassan, M. A. Alhartomi, S. V. Shvetsova, R. Sahal, and A. Hawbani, "Computing in the sky: A survey on intelligent ubiquitous computing for UAV-assisted 6G networks and industry 4.0/5.0," *Drones*, vol. 6, no. 7, p. 177, Jul. 2022.

[46] R. Y. Zhong, S. T. Newman, G. Q. Huang, and S. Lan, "Big data for supply chain management in the service and manufacturing sectors: Challenges, opportunities, and future perspectives," *Comput. Ind. Eng.*, vol. 101, pp. 572–591, Nov. 2016.

[47] T. D. Nguyen, M. Chiesa, and M. Canini, "Decentralized consistent updates in SDN," in *Proc. Symp. SDN Res.*, Apr. 2017, pp. 21–33.

[48] I. Lee and K. Lee, "The Internet of Things (IoT): Applications, investments, and challenges for enterprises," *Bus. Horizons*, vol. 58, no. 4, pp. 431–440, Jul. 2015.

[49] I. Arshad, K. R. R. Bharatwaj, S. H. Alsamhi, and E. Curry, "EHRCoI4: A novel framework for enhancing human-robot collaboration in industry 4.0," in *Proc. 3rd Int. Conf. Emerg. Smart Technol. Appl. (eSmarTA)*, Oct. 2023, pp. 1–6.

[50] R. Roman, J. Zhou, and J. Lopez, "On the features and challenges of security and privacy in distributed Internet of Things," *Comput. Netw.*, vol. 57, no. 10, pp. 2266–2279, Jul. 2013.

[51] M. Hermann, T. Pentek, and B. Otto, "Design principles for industrie 4.0 scenarios," in *Proc. 49th Hawaii Int. Conf. Syst. Sci. (HICSS)*, Jan. 2016, pp. 3928–3937.

[52] M. Christopher, *Logistics & Supply Chain Management*. London, U.K.: Pearson, 2016.

[53] S. Kamble, A. Gunasekaran, and H. Arha, "Understanding the blockchain technology adoption in supply chains-Indian context," *Int. J. Prod. Res.*, vol. 57, no. 7, pp. 2009–2033, Apr. 2019.

[54] H. Vahabi, F. Laoutid, M. Mehrpouya, M. R. Saeb, and P. Dubois, "Flame retardant polymer materials: An update and the future for 3D printing developments," *Mater. Sci. Eng. R, Rep.*, vol. 144, Apr. 2021, Art. no. 100604.

[55] C. Weller, R. Kleer, and F. T. Piller, "Economic implications of 3D printing: Market structure models in light of additive manufacturing revisited," *Int. J. Prod. Econ.*, vol. 164, pp. 43–56, Jun. 2015.
[56] E. Mengelkamp, J. Gärttner, K. Rock, S. Kessler, L. Orsini, and C. Weinhardt, "Designing microgrid energy markets: A case study: The Brooklyn microgrid," *Appl. Energy*, vol. 210, pp. 870–880, Jan. 2018.
[57] S. N. H. Shah, "IoT enabled smart grid integration with edge computing method," in *Proc. Int. Conf. Commun., Comput. Digit. Syst. (C-CODE)*, May 2023, pp. 1–6.
[58] R. Sahal, S. H. Alsamhi, J. G. Breslin, K. N. Brown, and M. I. Ali, "Digital twins collaboration for automatic erratic operational data detection in industry 4.0," *Appl. Sci.*, vol. 11, no. 7, p. 3186, Apr. 2021.
[59] R. Y. Zhong, X. Xu, E. Klotz, and S. T. Newman, "Intelligent manufacturing in the context of industry 4.0: A review," *Engineering*, vol. 3, no. 5, pp. 616–630, Oct. 2017.
[60] R. Sahal, S. H. Alsamhi, and K. N. Brown, "Personal digital twin: A close look into the present and a step towards the future of personalised healthcare industry," *Sensors*, vol. 22, no. 15, p. 5918, Aug. 2022.
[61] D. Ivanov, A. Dolgui, and B. Sokolov, "The impact of digital technology and industry 4.0 on the ripple effect and supply chain risk analytics," *Int. J. Prod. Res.*, vol. 57, no. 3, pp. 829–846, Feb. 2019.
[62] G. M. Sang, L. Xu, P. de Vrieze, Y. Bai, and F. Pan, "Predictive maintenance in industry 4.0," in *Proc. 10th Int. Conf. Inf. Syst. Technol.*, 2020, pp. 1–11.
[63] M. Timilsina, S. Alsamhi, R. Haque, C. Judge, and E. Curry, "Knowledge graphs, clinical trials, dataspace, and AI: Uniting for progressive healthcare innovation," in *Proc. IEEE Int. Conf. Big Data (BigData)*, Dec. 2023, pp. 4997–5006.
[64] M. Cavallo, M. Dholakia, M. Havlena, K. Ocheltree, and M. Podlaseck, "Dataspace: A reconfigurable hybrid reality environment for collaborative information analysis," in *Proc. IEEE Conf. Virtual Reality 3D User Interfaces (VR)*, Mar. 2019, pp. 145–153.
[65] E. Curry, *Real-time Linked Dataspaces: Enabling Data Ecosystems for Intelligent Systems*. Springer, 2020.
[66] T. Czvetkó and J. Abonyi, "Data sharing in industry 4.0—AutomationML, B2MML and international data spaces-based solutions," *J. Ind. Inf. Integr.*, vol. 33, Jun. 2023, Art. no. 100438.
[67] K. E. Knutsen, Q. Liang, N. Karandikar, I. H. B. Ibrahim, X. G. T. Tong, and J. J. H. Tam, "Containerized immutable maritime data sharing utilizing distributed ledger technologies," *J. Phys., Conf. Ser.*, vol. 2311, no. 1, Jul. 2022, Art. no. 012006.
[68] "Data sharing in the age of deep learning," *Nat. Biotechnol.*, vol. 41, p. 433, 2023, doi: 10.1038/s41587-023-01770-3.
[69] S. Kirrane, M. Sabou, J. D. Fernández, F. Osborne, C. Robin, P. Buitelaar, E. Motta, and A. Polleres, "A decade of semantic web research through the lenses of a mixed methods approach," *Semantic Web*, vol. 11, no. 6, pp. 979–1005, Oct. 2020.
[70] M. Pilkington, "Blockchain technology: Principles and applications," in *Research Handbook on Digital Transformations*. Cheltenham, U. K.: Edward Elgar Publishing, 2016.
[71] S. H. Alsamhi and B. Lee, "Blockchain-empowered multi-robot collaboration to fight COVID-19 and future pandemics," *IEEE Access*, vol. 9, pp. 44173–44197, 2021.
[72] S. H. Alsamhi, B. Lee, M. Guizani, N. Kumar, Y. Qiao, and X. Liu, "Blockchain for decentralized multi-drone to combat COVID-19 and future pandemics: Framework and proposed solutions," *Trans. Emerg. Telecommun. Technol.*, vol. 32, no. 9, Sep. 2021, Art. no. e4255.
[73] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," *Decentralized Bus. Rev.*, p. 21260, 2008, doi: 10.2139/ssrn.3440802.
[74] H. Xu, P. V. Klaine, O. Onireti, B. Cao, M. Imran, and L. Zhang, "Blockchain-enabled resource management and sharing for 6G communications," *Digit. Commun. Netw.*, vol. 6, no. 3, pp. 261–269, Aug. 2020.
[75] R. Sahal, S. H. Alsamhi, K. N. Brown, D. O'Shea, C. McCarthy, and G. Guizani, "Blockchain-empowered digital twins collaboration: Smart transportation use case," *Machines*, vol. 9, no. 9, p. 193, 2021.
[76] N. Radziwill, "Blockchain revolution: How the technology behind Bitcoin is changing money, business, and the world," *Quality Manag. J.*, vol. 25, no. 1, pp. 64–65, 2018.
[77] S. H. Alsamhi, A. V. Shvetsov, S. V. Shvetsova, A. Hawbani, M. Guizani, M. A. Alhartomi, and O. Ma, "Blockchain-empowered security and energy efficiency of drone swarm consensus for environment exploration," *IEEE Trans. Green Commun. Netw.*, vol. 7, no. 1, pp. 328–338, Mar. 2023.
[78] W. Mougayar, *The Business Blockchain: Promise, Practice, and Application of the Next Internet Technology*. Hoboken, NJ, USA: Wiley, 2016.
[79] M. Swan, *Blockchain: Blueprint for a New Economy*. Sebastopol, CA, USA: O'Reilly Media, 2015.
[80] S. Saberi, M. Kouhizadeh, J. Sarkis, and L. Shen, "Blockchain technology and its relationships to sustainable supply chain management," *Int. J. Prod. Res.*, vol. 57, no. 7, pp. 2117–2135, Apr. 2019.
[81] N. Kshetri, "1 blockchain's roles in meeting key supply chain management objectives," *Int. J. Inf. Manage.*, vol. 39, pp. 80–89, Apr. 2018.
[82] Q. Xia, E. Sifah, A. Smahi, S. Amofa, and X. Zhang, "BBDS: Blockchain-based data sharing for electronic medical records in cloud environments," *Information*, vol. 8, no. 2, p. 44, Apr. 2017.
[83] X. Yue, H. Wang, D. Jin, M. Li, and W. Jiang, "Healthcare data gateways: Found healthcare intelligence on blockchain with novel privacy risk control," *J. Med. Syst.*, vol. 40, no. 10, pp. 1–8, Oct. 2016.
[84] M. Crosby, P. Pattanayak, S. Verma, and V. Kalyanaraman, "Blockchain technology: Beyond Bitcoin," *Appl. Innov.*, vol. 2, nos. 6–10, p. 71, 2016.
[85] L. Morgenstern, C. Welty, H. Boley, and G. Hallmark, "RIF primer," in *Proc. W3C Recommendation*, vol. 22, 2010, p. 190.
[86] P. F. Patel-Schneider and I. Horrocks, "Position paper: A comparison of two modelling paradigms in the semantic web," in *Proc. 15th Int. Conf. World Wide Web*, May 2006, pp. 3–12.
[87] C. Bizer, T. Heath, and T. Berners-Lee, "Linked data: The story so far," in *Semantic Services, Interoperability and Web Applications: Emerging Concepts*. Hershey, PA, USA: IGI Global, 2011, pp. 205–227.
[88] D. Fensel, U. Şimşek, K. Angele, E. Huaman, E. Kärle, O. Panasiuk, I. Toma, J. Umbrich, and A. Wahler, "Introduction: What is a knowledge graph?" in *Knowledge Graphs: Methodology, Tools and Selected Use Cases*. Cham, Switzerland: Springer, 2020, pp. 1–10, doi: 10.1007/978-3-030-37439-6_1.
[89] J. Konečný, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, "Federated learning: Strategies for improving communication efficiency," 2016, *arXiv:1610.05492*.
[90] K. Bonawitz, H. Eichner, W. Grieskamp, D. Huba, A. Ingerman, V. Ivanov, C. Kiddon, J. Konečný, S. Mazzocchi, B. McMahan, T. Van Overveldt, D. Petrou, D. Ramage, and J. Roselander, "Towards federated learning at scale: System design," in *Proc. Mach. Learn. Syst.*, vol. 1, 2019, pp. 374–388.
[91] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Federated optimization in heterogeneous networks," in *Proc. Mach. Learn. Syst.*, vol. 2, 2020, pp. 429–450.
[92] S. H. Alsamhi, A. V. Shvetsov, A. Hawbani, S. V. Shvetsova, S. Kumar, and L. Zhao, "Survey on federated learning enabling indoor navigation for industry 4.0 in B5G," *Future Gener. Comput. Syst.*, vol. 148, pp. 250–265, Nov. 2023.
[93] C. Dwork and A. Roth, "The algorithmic foundations of differential privacy," *Found. Trends Theor. Comput. Sci.*, vol. 9, nos. 3–4, pp. 211–407, 2014.
[94] N. Rieke et al., "The future of digital health with federated learning," *NPJ Digit. Med.*, vol. 3, no. 1, p. 119, 2020.
[95] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," *ACM Trans. Intell. Syst. Technol.*, vol. 10, no. 2, pp. 1–19, 2019.
[96] A. V. Shvetsov, S. H. Alsamhi, A. Hawbani, S. Kumar, S. Srivastava, S. Agarwal, N. S. Rajput, A. A. Alammari, and F. M. A. Nashwan, "Federated learning meets intelligence reflection surface in drones for enabling 6G networks: Challenges and opportunities," *IEEE Access*, vol. 11, pp. 130860–130887, 2023.
[97] T. Nishio and R. Yonetani, "Client selection for federated learning with heterogeneous resources in mobile edge," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–7.
[98] S. H. Alsamhi, A. Hawbani, A. V. Shvetsov, and S. Kumar, "Advancing pandemic preparedness in healthcare 5.0: A survey of federated learning applications," *Adv. Hum.-Comput. Interact.*, vol. 2023, no. 1, 2023, Art. no. 9992393.
[99] S. H. Alsamhi, F. A. Almalki, F. Afghah, A. Hawbani, A. V. Shvetsov, B. Lee, and H. Song, "Drones' edge intelligence over smart environments in B5G: Blockchain and federated learning synergy," *IEEE Trans. Green Commun. Netw.*, vol. 6, no. 1, pp. 295–312, Mar. 2022.
[100] R. Myrzashova, S. H. Alsamhi, A. V. Shvetsov, A. Hawbani, and X. Wei, "Blockchain meets federated learning in healthcare: A systematic review with challenges and opportunities," *IEEE Internet Things J.*, vol. 10, no. 16, pp. 14418–14437, Aug. 2023.

[101] N. Andrade, M. Mowbray, A. Lima, G. Wagner, and M. Ripeanu, "Influences on cooperation in BitTorrent communities," in *Proc. ACM SIGCOMM Workshop Econ. Peer-to-Peer Syst.*, 2005, p. 111.
[102] S. Rieche, K. Wehrle, M. Fouquet, H. Niedermayer, L. Petrak, and G. Carle, "Peer-to-peer-based infrastructure support for massively multiplayer online games," in *Proc. 4th IEEE Consum. Commun. Netw. Conf.* Princeton, NJ, USA: Citeseer, Jan. 2007, pp. 763–767.
[103] Q. Lv, P. Cao, E. Cohen, K. Li, and S. Shenker, "Search and replication in unstructured peer-to-peer networks," in *Proc. 16th Int. Conf. Supercomput.*, 2002, pp. 84–95.
[104] J. Ioannidis and S. M. Bellovin, "Implementing pushback: Router-based defense against DDoS attacks," in *Proc. NDSS*, 2002, pp. 1–8.
[105] M. Ripeanu and I. Foster, "Mapping the Gnutella network: Macroscopic properties of large-scale peer-to-peer systems," in *Proc. 1st Int. Workshop Peer-to-Peer Syst.*, Cambridge, MA, USA. Springer-Verlag, Mar. 2002, pp. 85–93.
[106] Q. Huang, Y. Yang, and L. Wang, "Secure data access control with ciphertext update and computation outsourcing in fog computing for Internet of Things," *IEEE Access*, vol. 5, pp. 12941–12950, 2017.
[107] H. Haggi and W. Sun, "Multi-round double auction-enabled peer-to-peer energy exchange in active distribution networks," *IEEE Trans. Smart Grid*, vol. 12, no. 5, pp. 4403–4414, Sep. 2021.
[108] C. Zuo, R. Li, and Z. Lu, "A novel approach to improve the security of P2P file-sharing systems," *Int. J. Commun., Netw. Syst. Sci.*, vol. 2, no. 3, pp. 229–236, 2009.
[109] I. Abdullahi Yari, T. Dehling, F. Kluge, J. Geck, A. Sunyaev, and B. Eskofier, "Security engineering of patient-centered health care information systems in peer-to-peer environments: Systematic review," *J. Med. Internet Res.*, vol. 23, no. 11, Nov. 2021, Art. no. e24460.
[110] J. Koskela, J. K. Nurminen, and A. Gurtov, "A secure peer-to-peer application framework," Dept. Comput. Sci. Eng., Aalto Univ., Finland, 2015. [Online]. Available: https://urn.fi/URN:ISBN:978-952-60-6064-4
[111] J. Risson and T. Moors, "Survey of research towards robust peer-to-peer networks: Search methods," *Comput. Netw.*, vol. 50, pp. 3485–3521, 2006, doi: 10.1016/j.comnet.2006.02.001.

SAEED HAMOOD ALSAMHI received the M.Tech. degree in communication systems and the Ph.D. degree from the Department of Electronics Engineering, Indian Institute of Technology (Banaras Hindu University)—IIT (BHU), Varanasi, India, in 2012 and 2015, respectively. In 2009, he was a Lecturer Assistant with the Engineering Faculty, IBB University. Afterward, he held a postdoctoral research position with the School of Aerospace Engineering, Tsinghua University, Beijing, China. Since 2019, he has been an Assistant Professor with Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen. In 2020, he was a MSCA Smart 4.0 Fellow with Athlone Institute of Technology, Athlone, Ireland. Currently, he is a Senior Research Fellow with the Insight Centre for Data Analytics, University of Galway, Ireland; an Adjunct Professor with the Department of Computer Science and Engineering, College of Informatics, Korea University, Seongbuk-gu, Seoul, Republic of Korea; and an Assistant Professor with the Faculty of Engineering, IBB University, Ibb, Yemen. He has published more than 180 articles in high-reputation journals in IEEE, Elsevier, Springer, Wiley, and MDPI publishers. His research interests include green and semantic communication, the green Internet of Things, QoE, QoS, cybersecurity, multi-robot collaboration, blockchain technology, peatland and wastewater into energy, federated learning, and space technologies (high altitude platforms, drones, and tethered balloon technologies).

AMMAR HAWBANI received the B.S., M.S., and Ph.D. degrees in computer software and theory from the University of Science and Technology of China (USTC), Hefei, China, in 2009, 2012, and 2016, respectively. From 2016 to 2019, he worked as Postdoctoral Researcher with the School of Computer Science and Technology, USTC. He is a Full Professor at the School of Computer Science, Shenyang Aerospace University, Shenyabg, China. His research interests include the IoT, WSNs, WBANs, WMNs, VANETs, and SDN.

SANTOSH KUMAR (Senior Member, IEEE) is an Assistant Professor in Computer Science and Engineering with IIIT Naya Raipur, Chhattisgarh, India. Prior to joining IIIT-NR, he was a Ph.D. scholar with the Department of Computer Science and Engineering, IIT (B.H.U.), Varanasi, Uttar Pradesh, India. He has published over 14 journal (SCI-index journal) and five conference papers (including 2 tier-1 international conferences (ACM multimedia-2016 and Mobisys-2016); filed one patent, and 14 book chapters in edited books (Springer publication and IGI publication). He is an active member of the Computer Society and the Association for Computing Machinery.

MOHAN TIMILSINA received the Ph.D. degree in computer science from the Data Science Institute, University of Galway, Ireland, in 2020. He is currently a Senior Postdoctoral Researcher with the Data Science Institute, University of Galway. His research interests include applied machine learning, bioinformatics, graph mining, and information retrieval from tabular and networked data.

MAJJED AL-QATF received the B.S. degree in network technology and computer security from Sana'a University, Sana'a, Yemen, in 2013, the M.S. degree in computer science and technology from Central South University, Changsha, China, in 2019, and the Ph.D. degree in computer science and technology from the University of Science and Technology of China (USTC), Hefei, China, in 2023. Currently, he is a Postdoctoral Researcher with the Data Science Institute, University of Galway, Galway, Ireland. His research interests include AI, CV, NLP, image captioning, large language models, and data space and knowledge graphs for data sharing.

RAFIQUL HAQUE received the Ph.D. degree in computer science and information systems from the University of Limerick. Currently, he holds a Research Fellow with Insight, the SFI Center for Data Analytics, University of Galway. Prior to joining Insight, he was a Senior Postdoctoral Researcher with University College Cork and a Postdoctoral Researcher with Claude Bernard University Lyon 1 (UCBL 1) and Université de Versailles Saint-Quentin-en-Yvelines (UVSQ). He co-founded two startups, Cognitus and Intelligencia, Paris, and was the Vice President of both companies. Over the past 15 years, he has been involved in innovation, development, entrepreneurship, team leadership, teaching, and sales. He has led numerous projects funded by European Union, Enterprise Ireland, SFI, and ANR. His current research interests include interoperability in data spaces, governance of data spaces, and secure AI and data analytics. He has collaborated with various European institutions. Additionally, he worked in the IT industry for seven years, handling product management, project leadership, and coordination.

FARHAN M. A. NASHWAN received the Master of Science degree in electrical engineering from Electronic and Communication Engineering, Facultiy of Engineering, Al-Mustansiriyah University, Baghdad, Iraq, and the Ph.D. degree in electronics and electrical communications engineering from Cairo University, Egypt, in 2014. He is currently an Assistant Professor at Electrical Engineering Department and Dean of faculty of engineering, IBB University, Ibb, Yemen. His current research interests are in the fields of clustering, object tracking, pattern recognition, and computer vision.

LIANG ZHAO (Member, IEEE) received the Ph.D. degree from the School of Computing, Edinburgh Napier University, in 2011. He is currently a Professor with Shenyang Aerospace University, China. Before joining Shenyang Aerospace University, he was an Associate Senior Researcher with Hitachi (China) Research and Development Corporation, from 2012 to 2014. His research interests include ITS, VANET, WMN, and SDN. He is also a JSPS Invitational Fellow, in 2023, and a Visiting Professor with the University of Electro-Communications, Japan. He was listed as Top 2% of Scientists in the world by Standford University (2022 and 2023).

EDWARD CURRY is currently the Established Professor in data science and the Director of the Insight SFI Research Centre for Data Analytics, University of Galway. He has made substantial contributions to semantic technologies, incremental data management, event processing middleware, software engineering, and distributed systems and information systems. He combines strong theoretical results with high-impact practical applications. The excellence and impact of his research have been acknowledged by numerous awards, including best paper awards and the University of Galway President's Award for Societal Impact, in 2017. His team's technology enables intelligent systems for smart environments in collaboration with several industrial partners. He is an organizer and the Programme Co-Chair of major international conferences, including CIKM 2020, ECML 2018, IEEE Big Data Congress, and European Big Data Value Forum. He is also the Co-Founder and the elected Vice President of the Big Data Value Association, an industry-led European big data community, has built consensus on a joint European big data research and innovation agenda, and influenced European data innovation policy to deliver on the agenda.