JOURNAL OF SPATIAL INFORMATION SCIENCE
Number 27 (2023), pp. 1–26 doi:10.5311/JOSIS.2023.27.220

RESEARCH ARTICLE

Distributed spatial data sharing: a new model for data ownership and access control

Majid Hojati[1], Rob Feick[2], Steven Roberts[1], Carson Farmer[3], and Colin Robertson[1]

[1]Department of Geography and Environmental Studies, Wilfrid Laurier University, Waterloo, ON, Canada
[2]School of Planning, University of Waterloo, Waterloo, ON, Canada
[3]Textile.io, Victoria, BC, Canada

Received: April 27, 2022; returned: September 2, 2022; revised: November 26, 2022; accepted: June 2, 2023.

Abstract: With the advent of new technologies and broader participation in geospatial data production, new challenges emerge for spatial data sharing. Spatial data sharing practices are increasingly transacted through and, to varying degrees, controlled by a handful of privately controlled corporate services. Data production has evolved from being largely centralized, expert-oriented, and authoritative in nature to now also include hybrid data collection processes involving distributed assemblages of individuals who share and co-produce spatial data while interacting through centralized architectures and control regimes. These changes have resulted mainly from technological and social changes linked to the emergence of Web 2.0 and widely available Internet participation tools. Concerns about how spatial data access and sharing are controlled, particularly for sensitive or personally-identifying data, have increased interest in distributed file technologies that allow users to share resources independently of centralized platforms. This paper examines how spatial data sharing practices may move towards a more decentralized sharing ecosystem as technologies for a further distributed web mature. We identify this transition as increasingly hybridized forms of data ownership and access control concerns are coupled with new distributed systems (e.g., Web 3.0). We also discuss opportunities and barriers to distributed spatial data sharing, including possible benefits for big geographic data management and the need for protocols to share, integrate, and process spatial data shared on distributed networks.

Keywords: spatial data sharing, distributed data sharing, data ownership, blockchain, IPFS, SDI

1 Introduction

Daily, vast amounts of data are produced by individuals interacting with digital content and through automated and semi-automated sensors deployed across the environment. A growing portion of this data contains geographic information directly or indirectly embedded within it, vastly increasing the use-cases for geospatial data [74]. The advancements in information and communications technology, including the widespread adoption of GPS-based sensors, improvements in computational data processing, and satellite imagery, have resulted in new data sources, stakeholders [30], and methods of producing, using, and sharing spatial data. Many of these changes relate to the growth of user-generated geographic information through the widespread use of automated smart sensors, and a wider variety of georeferenced media. It is estimated that around 2% of tweets have explicit coordinates [70] while a more significant portion (55.6%) contain locational data that can be inferred from the content of tweets (e.g., see [21]). These changes have impacted spatial data sharing in several ways.

First, as Kotsev et al. [45] note with respect to the post-2020 path for INSPIRE [23], national and pan-national spatial data infrastructures (SDI) have been challenged to accommodate more heterogeneous data sources and diversity of licensing frameworks. Data integration, access, and rights management have long been challenging in spatial data sharing, given the need for users to conform to common standards, policies, and processes. However, the heightened role of individuals as data authors and the increasing part of private sector platforms in controlling geodata have added new dimensions to these challenges [11, 32, 50].

Second, the growing volumes of data documenting an individual's transactions, movements, and other aspects of behavior at high spatial-temporal resolutions have spurred concerns over data ownership, use, and control. Efforts to enable users to protect their geoprivacy, for example, while contributing personal movement data through COVID-19 exposure notification apps (e.g., see [44, 64]) are one example of this challenge. As [29, 35] note, this concern has been raised with other forms of actively and passively generated volunteered geographic information (VGI) for some time. Spatial data governance is less well defined within individual and group contexts, and data sharing is typically one-directional from user to platform. Data authors thereby lose control over where their data are stored, who can view them, or what purposes they can be used for. Platform licenses and terms of agreements play a key role in limiting contributors' rights, as illustrated by [69, 82]. An obvious example is Google Maps, where Google owns, controls access to, and resells data that users contribute both actively (e.g., editing places on the map) and passively (e.g., cell phone movement) [99].

While data production has shifted to a decentralized 'prosumer' model, the storage, management, and controlling access of such data have remained centralized in a logical, if not a physical, sense. This leads to a well-known power imbalance in the data economy and requires a form of social trust between users and the platforms they interact with [30]. To address these concerns, researchers have begun to explore data sharing approaches where data access and use control is distributed to individual data authors and owners. From a social perspective, this *distributed spatial data sharing (DSDS)* environment recognizes existing power imbalances between users and platforms in the data economy [30, 45] and the need to operationalize more dynamic forms of social trust between data authors, other users on a network, and platforms. Technologies to enable DSDS include many of the software and hardware architectures developed for storing and processing big data (e.g., digital earth platforms, data spaces, etc. [36, 50]). Distributed technology can be at the storage level, such as the InterPlanetary file system (IPFS), at the application level, such as decentralized applications (DApps), or at the process level, such as Apache Hadoop. In a DSDS approach, network entities share and control resources. The recent emergence of the distributed web (aka Web 3.0) is based primarily on peer-to-peer (P2P) networks, whereby nodes can communicate directly without intermediaries. This provides the capability to share data via distributed nodes rather than having centralized storage and control.

In summary, centralized data storage of user-contributed data is no longer sufficient to meet data sharing needs in the big data era [10]. In addition, current data storage approaches are vulnerable to data loss when data centres are damaged or hacked. Such platforms often also lack transparent policies about how data may be reused, leading to an 'economy' in the buying and selling of user-contributed data [97]. The trajectory from fully centralized spatial data storage, distribution, and manipulation (controlled by government and/or corporations) to distributed individual data sharing remains technically challenging. These problems are exacerbated for smaller organizations, as the collection and sharing of data over common existing platforms (such as geoportals, web-map services, etc.) entail specific legal and policy-related challenges [57]. In this paper, we identify several situations or use cases for which DSDS is most suited and identify GIScience research challenges in realizing DSDS as a solution to several current problems of spatial data sharing (SDS). The aims of this paper are as follows:

1. Review current models of SDS and their challenges;
2. Examine how DSDS can provide potential solutions for ownership and control challenges in current SDS models; and
3. Identify GIScience research challenges in the implementation of DSDS.

To move toward the above goals, we look first at the evolution of SDS and describe how some types of SDS are transitioning from a centralized data production, control, and processing model to a more distributed one.

2 SDS models

One way to conceptualize the changes taking place in how we share spatial data is through the shifting roles of individuals, institutions, governments, and enterprises (private sector), acting as data producers, data controllers, and data users [41]. Data producers are the individuals and entities that generate data. In this paper, data producers such as grassroot groups, indigenous peoples, and academic and non-profit groups are all classified as communities of interest. Data controllers are platform owners, data intermediaries, or license holders (see Figure 1). These roles overlap and interact fluidly [22]. For example, in the Keating et al. [41] model, the role of data providers is defined as stewardship of data, and data managers are responsible for quality assurance, metadata, data access control, and delivery. However, this classification is not always valid in cases such as VGI. Each SDS stakeholder, including individuals, governmental entities, enterprise entities, communities of interest, and hybrid participants, can be assigned to the above roles. We examine how large-scale shifts in the relative proportion of participations and roles have changed SDS's nature in recent decades, highlighting three main eras: centralized, user-centric, and distributed SDS eras. Classifying these eras can help us to identify how SDS practices and needs are changing and what types of data best suit DSDS.

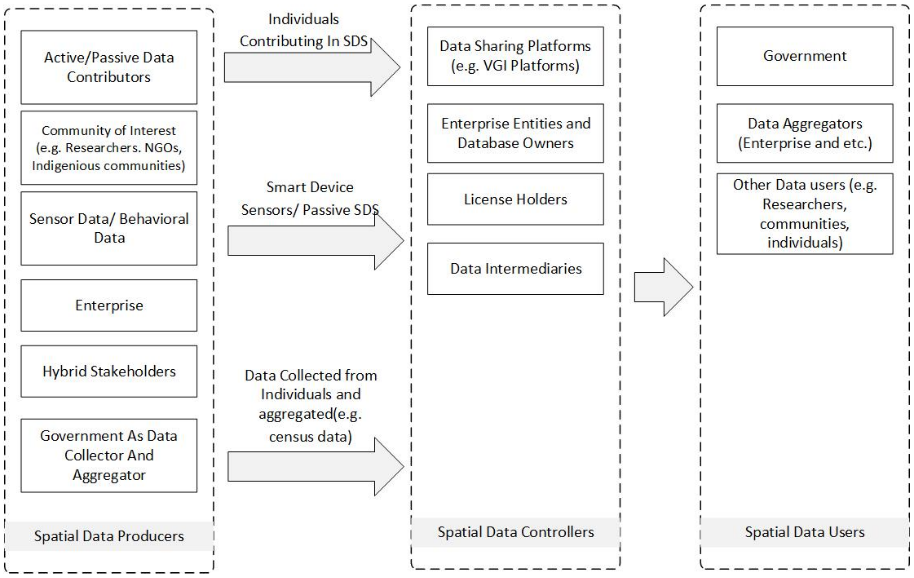
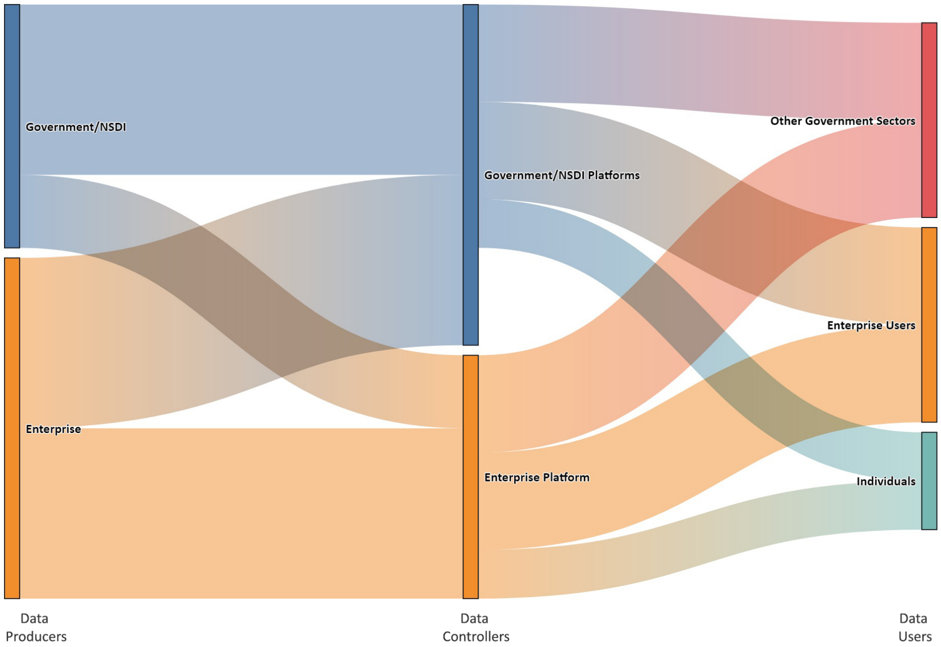


Figure 1: Participants in SDS as data producers, controllers, and users in the current data sharing environment.

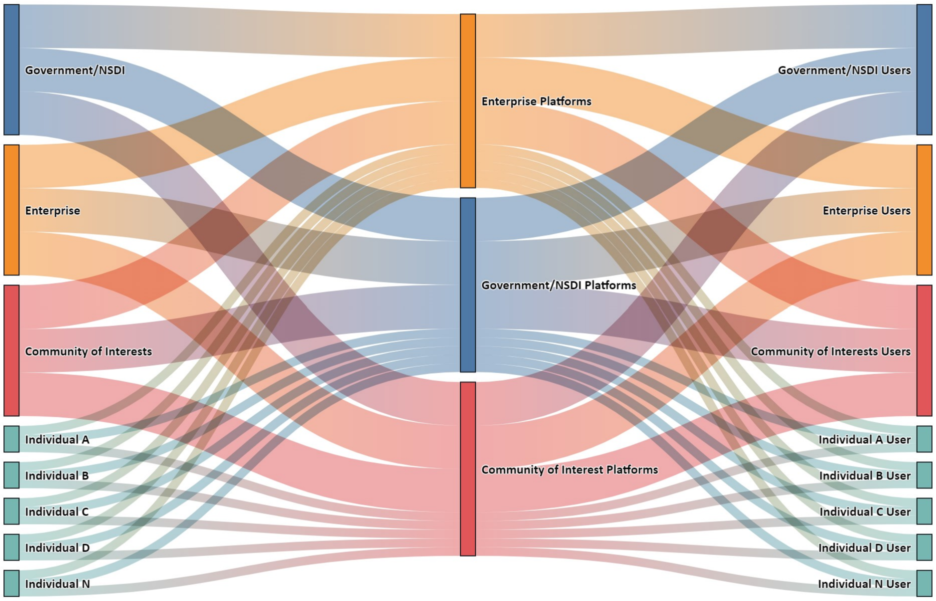
2.1 Centralized and user-centric SDS models

Two main approaches to SDS have developed over the past several decades are shown in Figure 2 with the main types of data flows between data producers, controllers, and users. Figure 2 (a) depicts the centralized model, where data sharing happens primarily between governmental bodies and corporations, who act both as data producers and as data controllers and mainly covers the SDS environment between the 1980s and mid-1990s [30]. In the centralized model, sensors and tools are only accessible to the government and enterprise sector, controllers and users are the same entities, and individuals only use a small portion of available data. In contrast, Figure 2 (b) portrays data flows in the current environment where individual users and communities of interest (e.g., NGOs, non-profit groups, universities with hybrid stakeholder model) play more substantial roles in authoring, controlling, and using more heterogeneous sources and forms of spatial data.

The increased use of digital data platforms over the past two decades [30] led to the currently dominant model of SDS that we refer to here as a *user-centric SDS model* [13, 19, 33]. The changing Internet environment, especially the dissemination of easy-to-use spatial data, decreased the relative importance of governments and third-party companies as data producers and individual users started to produce and use more spatial data [11]. The expanded role of citizens as data producers, facilitated by corporate-controlled platforms,



(a) Data Flow in centralized SDS model. The main participants in this model are enterprise and government.



(b) Data Flow in user-centric SDS models. In this model, individual-level data production increases. However, their rights over contributed data are controlled by the platforms.

Figure 2: Estimated data flow in centralized and user-centric data models. The vertical bars correspond to the data producer, controller, and user; the data flow is from left to right.

has increased the relative power of platform owners as arbiters of SDS. In this model, new players such as communities of interest (e.g., Birds Canada (https://www.birdscanada.org)) and individuals have emerged as important producers of spatial data (left of Figure 2 (b)). Data access policies vary between open (public) and paid in such groups and are limited by different types of licenses. Data controllers remain governments, corporations, and some communities of interest groups that own data sharing platforms (center of Figure 2 (b)), while users can be any of the mentioned groups (right of Figure 2 (b)). This model has hybrid platforms and licenses between enterprises and individuals (e.g., Google Maps reviews and photos) or university-maintained repositories in which users and project teams share data, define reuse terms, and document authorship in metadata. In the university-maintained repositories example, the rights favor the data producers. Still, in the enterprise platforms, the users need to acknowledge enterprise terms of service (TOS) to contribute. In this model, governments are only responsible for standardizing data sharing protocols and, in conjunction with corporations, for collecting large-scale and framework data [30]. Some types of user-collected data can have higher spatial and/or temporal granularity than government-produced data and thereby fill gaps and capture phenomena (e.g., public sentiment) not covered by traditional data collection practices (e.g., census data) [30]. From the perspective of data production, it is clear that the number of contributors has grown significantly [28, 61, 71]. Spatial data use has also changed from predominately expert-oriented to include more individual and community approaches (e.g., citizen science projects) and location-based services (see [66] for example). In the user-centric model, individuals have limited rights to control their data, and platforms' relative power has increased. As such, data ownership and data control topics have become a concern among users, communities, and even the platform owners (see [11]), especially with several high-profile data breaches of commercialized user data.

2.2 When centralized data sharing fails

Despite the current advancements in centralized spatial data sharing platforms, unsolved challenges remain. Šumarada [84] has identified intellectual property rights to protect producers' rights as one of the major legal issues related to geographic data sharing. Policies and licensing issues are changing as the actors in SDS are changing [45]. Take VGI tools as an example, Scassa [69], Cho [11] and Longhorn et al. [49], argue that VGI data hosted on a website is considered *compiled* for the purposes of copyright law. They consider data compilation in copyright law and claim that, for example, in Canada and the US, only original data is protected by the law, while in the EU, the protection might be broader as they have database protection laws. Some suggest that VGI platform owners can use license agreements as a fundamental tool to control intellectual property rights [38] and limit the rights of users who are contributing on VGI platforms in order to avoid any future legal claims by contributors [69]. Tracing the heritage of data in VGI platforms and hence the copyright ownership can be difficult [11]. Such difficulties are raised because *compiled* data might lose their copyright due to difficulties in tracing all previous owners of the data. Scassa [69] describes this problem as a *Wiki* effect, wherein multiple authors have contributed to a dataset. Granell and Ostermann [32] urge that there must be mechanisms to ensure the privacy of gathered data and that data governance must clearly identify who owns the data and define time limits for which to use them. Michener et al. [56] mention a lack of technical implementations for acknowledging data contributions and a dearth of easy-to-use tools for accessing data, their conversion, and analysis as some of the barriers to open data. Not having a clear owner associated with the data might impact on the user's willingness to contribute and raise trust-related concerns.

Ewis et al. [25] argue that trust is a motivational factor in data sharing behaviors and having control over one's data can increase that individual's willingness to share data with an organization, both now and in the future [81]. A lack of transparency and trust in centralized data infrastructures could be a key factor in preventing the true realization of participatory government models [73]. Similar to Ewis et al. [25], Wehn de Montalvo [90] use the term *perceived control* to represent an individual's willingness to impede or facilitate data sharing. The lack of transparency in many central platforms concerning how individuals' data are stored and shared with or sold to third parties reduces trust between contributors and platforms.

Some communities of interest, such as indigenous communities, illustrate other shortcomings of the current SDS context. Indigenous communities require sovereignty over data that document their territories, socio-cultural resources, and values, as well as their traditional lands. Such control varies from storage methods to the different access levels and usages [7]. *Data sovereignty* has been defined as the management of information in a way that aligns with the laws, practices, and customs of the nation-state in which it is located [16]. This concept has been extended in the context of SDS to self-governing groups or nations within states (e.g., indigenous communities). Indigenous data sovereignty involves data locality [16] and access control that is defined through distinct community laws, principles, and practices (see Find-able, Accessible, Inter-operable and usable (FAIR) [91] and Collective benefits, Authority to control, Responsibility and Ethics (CARE) [1] principles). Following these principles, information (including data collected voluntarily) must be available, accessible, and open to all community members, and data access is restricted for outsiders [82]. Such data sharing requirements preclude these communities from contributing to public platforms like OpenStreetMap due to its open data policies. As a result, interest is growing in tools and methods that will permit the data sovereignty concept to be extended further to individuals as well as communities. Sharing personal information or social media contributions is another example. In the current user-centric model, a range of public data processing practices, and regulations are used to control how personal information is shared and represented to users.

Sharing personal information or social media contributions is another example. In the current user-centric model, a range of public data processing practices, and regulations are used to control how personal information is shared and represented to users. Despite such regulations, Granell et al. [32], Alessi et al. [2] and Camenisch et al. [9] argue that individuals have limited control over their data and identify the need for easy-to-use tools to control access to personal data and data governance policies that clearly delineate who owns the data and the time limits for using them. Despite the advances these platforms provide, as Camenisch et al. [9] note, they do not guarantee that users' data will not be used in their internal studies and data processing.

By looking at the above example, it is clear that challenges in the current data sharing environment are related to the new participants in the current SDS environment. The current centralized platforms either provide tools and TOS to avoid legal confrontation with contributors or simply ignore individual or community-level data sharing needs. This issue can be solved by transferring the data controller role to the data producers. Table 1 shows a summary of SDS requirements which are not all supported in the centralized platforms.

Transparency	• Transparent decision-making procedures in participatory GIS (PGIS) processes by decentralized consensus • Immutable tracing the changes in the GI without middling third-party platforms • Transparency about GI which is used and stored in the platforms • Transparent technology and its internal mechanism
Ownership	• Avoiding the Wiki effect in the data contributions • Clear copyright and licensing at the feature level
Control	• Controlling who, when, and why contributed data should be used • Benefiting the data contributors using data markets

Table 1: SDS requirements which the current data sharing platforms are not able to fulfill.

2.3 Distributed SDS model

Social and economic life is increasingly facilitated through web-based data and services. In contrast, core issues of data privacy, ownership, and access remain concentrated among data controllers. There is a growing demand for a new data sharing and governance paradigm. In the context of SDS, for example, the recent focus in research has changed from SDIs to individuals' geoprivacy (e.g., [42]), rights (e.g., [23]), and other legal aspects of SDS (e.g., [5]). We believe that we are on the cusp of a new data sharing era which may take various forms. One particular form of SDS we want to highlight is that of Distributed Spatial Data Sharing (DSDS).

The critical shift in DSDS is that individuals are not only data producers and users, but they will also serve as data controllers (see Figure 3). Some recent initiatives related to data sharing (e.g., the right to be forgotten) emphasize the primary role of individuals. In this model, data ownership and access control are determined by the individuals, not by platforms, which provide only the infrastructure and tools. This transition is based on users' concerns regarding the behavioral data they have collected or extracted by using different platforms.

Distributed systems are defined by Tanenbaum et al. [78] as "transparently utilizing every kind of available resource on the network of the users and connecting the users on the network to the distributed resources transparently with support of openness and scalability." The resources can be data [43,48,95], processes [48,94] or knowledge [14]. A distributed ledger is a distributed database where data can only be appended or read [76]. Blockchains implement the structure and functionality of distributed ledgers so that each new information block is generated based on the previous block and appended as a new block on the end of the chain. These blocks of information can even be a program or series of smart contract calls. Smart contracts are a set of scenario-response procedural rules and logic on the blockchain that run when predetermined conditions are met [67, 88]. They, therefore, allow for secure transaction issuance without the need for third parties. For example, smart contracts can automatically move digital assets according to arbitrary pre-specified rules in a crypto-currency network [8]. The transactions that smart contracts issue are traceable and immutable [76].

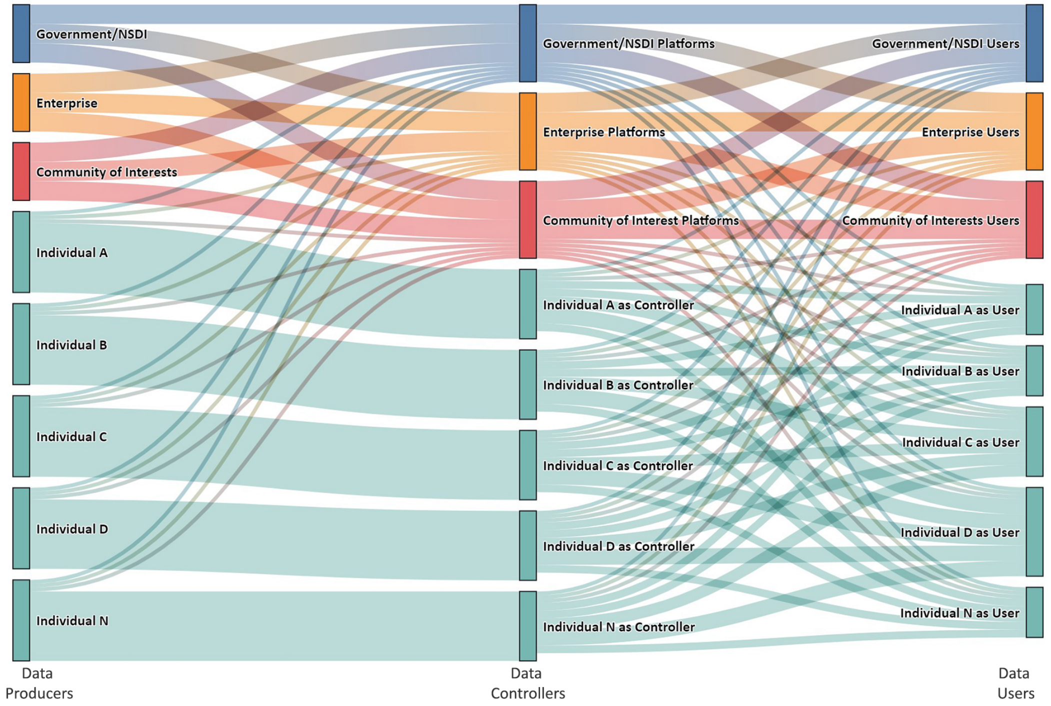
In a distributed data sharing method, the data workflow of data would be similar to those of traditional data sharing models (Figure 3 (a)). The prominent roles of data producers, controllers, and users remain intact. However, the role of each actor will change. In such an environment, any entity within the network can be a data producer, a data controller, and/or a data user. An individual, for example, can produce data, have control over the produced data, and also gain access to the data from other entities (Figure 3 (b)). The economic regime of data sharing will be more complex, for example, with the advent of data markets and the ability of users to control who, when, and where data are available. Individuals and community groups will be able to play a more important role in the data sharing environment.

In a DSDS, technical factors will be more impactful. With the emergence of Web 3.0, including peer-to-peer protocols and distributed systems, users are not dependent on a central platform for their data sharing needs and thus have a higher level of control over data. In addition, advocate users who are more interested in transparency in the platforms can access the documentation about the internal mechanism of the technology and examine it to make sure it is trustable.

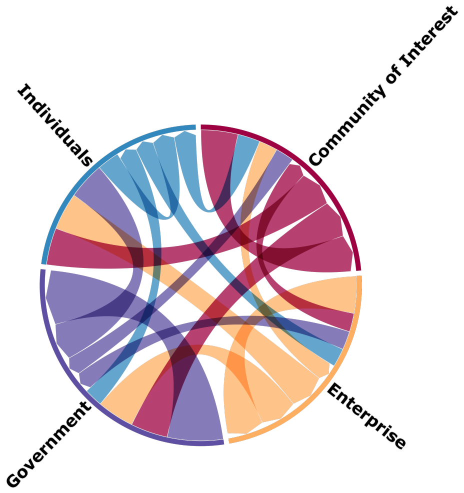
It is worth noting that distributed systems include different technologies such as distributed file storage, blockchains, and smart contracts. Such technologies can be integrated and are complementary, providing a distributed computing/sharing environment. For example, access control in distributed file-sharing systems can be done using smart contracts (e.g., [75]). In addition, because distributed ledgers are not meant to store large data, distributed file storage can be used [60]. The P2P paradigm is mostly popular for storing and sharing information in a totally decentralized manner [40]. Typically, a P2P file storage system is a distributed environment formed by autonomous peers that operate independently. Each peer stores a part of the available information and maintains links (indices) to other peers. In addition, we need to distinguish blockchain from crypto-currencies. The blockchain technology is the backbone of the crypto-currencies such as Bitcoin. For instance, Bitcoin is a financial use case of the blockchain, and its mechanism (such as its proof-of-work and energy consumption concerns) differs from other Web 3.0 applications.

Data access control on distributed platforms is often handled using public/private key management methods and smart contracts (e.g., [20, 37, 75]). In such platforms the data owner can distribute secret keys to users and encrypt shared data based on different access policies [89]. Such an ability in distributed systems enables owners to encrypt data and share it with users directly. In contrast, in centralized systems, a third party is responsible for sharing collected spatial data. In centralized systems, users have to trust the cloud and the application providers—users have no choice but to rely on the security and availability of the application providers, to accept their policies, and to adhere to their licenses [72]. Such methods have been used in distributed personal data stores to provide control over sharing personal data (including location) with different platforms (e.g., see [2]). Users who accept a third-party license agreement risk losing their intellectual property rights.

Figures 2 and 3 illustrate the complexities of each SDS context. In Figure 3 (a), the data sharing between the data controllers and users is more complex because individual data producers act as data controllers. New approaches for SDS are required for ownership and access control since current SDS platforms do not support these needs. In Figure 2 (b), the main focus is on individuals who are data contributors and users in the data sharing platforms but are not able to assume the data controller role for their contributions (middle column).



(a) Each actor in DSDS can be a data producer, a data controller, and a user. The complex relations in the right side of the graph indicates individual level access control.



(b) Data flow in DSDS model.

Figure 3: The flow of the data in DSDS model and the main participants.

3 DSDS capabilities

Distributed systems and smart contracts have the potential to shift the ways data are used and controlled in society [6]. Ølnes et al. [65] have identified 16 benefits of distributed systems and smart contracts in governmental bodies in five classes: strategic, economic, informational, technological, and organizational impacts. The benefits of such systems are attributed primarily to other technologies like encryption methods. In addition, some of these benefits are not limited to the distributed technology, but to social and cultural benefits as well. However, looking at the extension of the existing technologies the aforementioned benefits have not yet been put into practice.

As distributed systems are being developed and used by different applications, it becomes possible to categorize the advantages of distributed systems as follows:

Decentralized consensus Decentralized consensus was used by Bolin [6] as one of the key features of public ledgers enabling trust-less transactions. Data can only be stored via group consensus in the distributed networks, consequently becoming more transparent and verifiable [6, 24]. Such features of distributed networks prohibit data tampering—the data cannot be altered on the network since it is stored individually on the nodes on the network. The data are not exclusively maintained by a single individual or entity, but rather, are available to everyone, and the state of the data are decided upon via a consensus protocol [26].

Immutability Immutability is another feature of distributed systems. Immutability is the central reason why participants trust distributed systems and ledgers [55,77]. Immutability means that records and data cannot be altered and changed after they have been added to the network [34]. Distributed file-sharing systems usually preserve a unique HASH address for each unique file. Content-based addressing (in the case of the IPFS) provides the unique feature of tracking data in the network using its content HASH address. Content-based addressing can be used as a key component in conceptualizing a distributed network's file sharing as a database shared over nodes of the network. Immutable spatial data can be used in cadastral applications, for instance, used for a rental transaction or registration of land plots as spatial objects [17, 54]. The combination of content-based addressing with geohash algorithms—which encodes a geographic location into a unique short string of letters and digits [58]—or locality sensitivity hashing (see [4, 85]) might even provide a means of providing spatial data queries over a distributed network.

Append-only nature The ability to record the state of phenomena in an immutable format and to store such records as a chain of events provides a suitable means by which the state of the data at any historical block may be queried. In a distributed ledger each new record is appended to the previous record and the entire chain of records can be verified by the network nodes. This feature can help future platforms manage the information and life-cycle of spatial data. Such models can also be used for spatial-temporal analysis by providing the capability of querying the chain of transformations over each spatial entity.

No single point of failure From a technical perspective, decentralized storage solves the problem of the single point of failure in traditional storage systems and cloud-based systems [89]. Shafagh et al. [72] compare conventional cloud-based solutions for IoT data storage and sharing and argue that they are able to (1) provide secure data storage, and (2) provide IoT-compatible data streams, but are unable to provide decentralized access control management. Smart-contract-based applications can provide all of the above requirements for cloud-based data sharing [72].

One of the challenges to most of the current SDIs that are developed in remote areas, like the Arctic SDI, is user connectivity [59]. The accessibility of the Internet to users and user access to the central servers of SDI in remote areas is mentioned in most of the related technical documents. In the first pilot of the Arctic SDI project (www.arctic-sdi.org), challenges such as data integration, limited telecommute resource/bandwidth in the North, and end-users' concerns about the data policies are highlighted [51]. DApps and distributed data sharing methods are not dependent on central servers and data can be accessed from existing nodes on the network, potentially addressing communication issues with the 'outside' by enabling local area access within remote areas.

Scalability Distributed systems are scalable by nature. Considering the amount of geographic data which are being produced every day from many sources and at different resolutions and in different environments such as digital Earth, the scalability of data sharing, management, and processing is a necessity. There have been many studies on distributed geographic processing but the scalability of geographic data sharing has not been studied before (e.g., [87]). Looking at the scalability of existing GIS architectures, it is claimed that the existing GIS architectures are not able to fulfill the scalability challenges raised by the enormous number of users, data, and heterogeneous data sources [93, 97]. Thus, a new Internet GIS architecture that can scale up to accommodate these needs is essential. A fully distributed P2P spatial architecture may fulfill the scalability needs of new architectures [47]. They can scale well because it is possible to effectively partition access across the network. No one individual needs to store all data for that data to be available to the whole network.

4 DSDS as a solution

Considering spatial data sources, Cuno et al. [15] classified different data sources for an urban data space. They categorize the source of data into 8 classes: (1) official institutional, (2) enterprise, (3) research, (4) personal, (5) behavioral, (6) freely available, (7) commercially available, and (8) internally available. Looking at this classification scheme, it is obvious that some of the classes have clear data ownership, (1), (2), (4), (8), while some of the classes have arguable ownership, (3), (5), (7). For example, personal data are subject to data production regulations (especially in the EU and North America). Third parties require the data subject's consent to store and process the data. In addition, physical persons have the right to inspect the data and to initiate its removal from third parties' central servers. Freely available data can be discussed under open data sharing policies. For the purpose of this discussion, we have used this data source classification. However, we acknowledge that this classification is not complete and there are overlaps between such classes. We also added a new data source called (9) hybrid data which are datasets that are combined from any of the classes of data from (1) to (8).

In order to use DSDS to share each category of the above data a different approach is required. In categories such as (1), (2), and (8) depending on the policies of the data owner they might opt to use DSDS. However, to share data that includes personal information, or collective information DSDS can be used. Take personal data, for example. Users will be able to share their identity, location, and other related information in distributed social media. They are able to store and retrieve their location information with different resolutions on the P2P networks and encrypt them with their own private keys. Only contacts that are allowed to access a certain level of privacy can access the specific geodata resolution (See [35] for storage and retrieval of user location with dynamic geoprivacy from distributed networks). Figure 4 (c), (d) shows a simplified architecture of personal data storage using distributed systems.

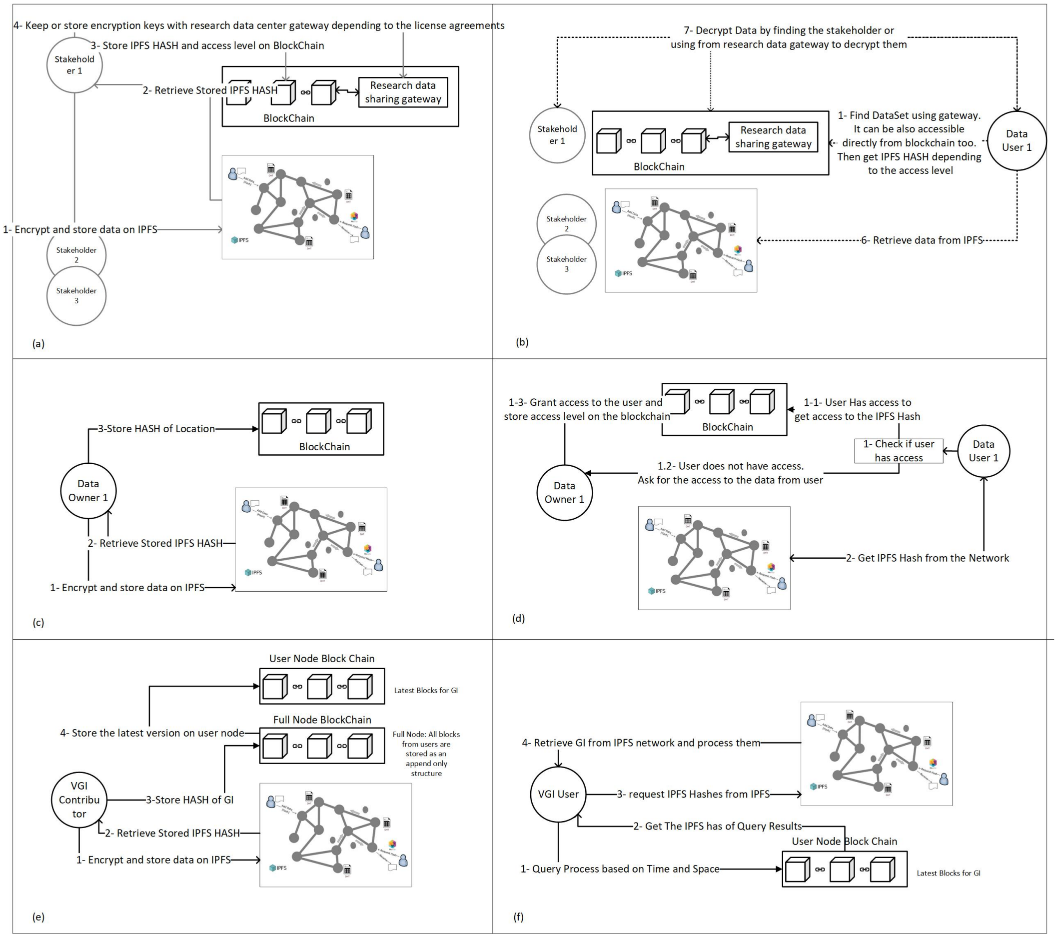


Figure 4: The proposed architecture of a decentralized model of data storage for (a), (b) Research data, (c), (d) Personal data, and (e), (f) VGI platform. The left column is the process of sharing data from a data owner's perspective and the right column is the process of sharing data from a data user's perspective.

Research data is another category that usually includes shared data ownership between different stakeholders under various license terms. This information can be another good example of using the DSDS model. However, the complexity of the models can increase as the licensing terms get more complicated. Figure 4 (a), (b) shows the architecture for storing such information. As the figure shows each stakeholder in the research data encrypts data and stores them on the IPFS network (Steps 1 and 2). The IPFS Hash will be later stored on a blockchain alongside the access level and related metadata (Step 3). A research data gateway will be responsible to index existing datasets on the blockchain and providing a method to query data from the blockchain (b). A user who wants to access the shared data needs to first find the IPFS Hash of the dataset. This can be done using a research data gateway or querying the blockchain directly. In this step, the user's access level can be examined at the blockchain level. Once they get a hold of the IPFS hash, they need to decrypt the data from the dataset. The decryption can be done either by directly inquiring about the data from stakeholders, requesting a blockchain, or a research gateway. However, more in-depth access control methods are required for such architecture.

A distributed ledger structures a series of transactions/blocks which can then be used to trace any changes made to the data. An example of such a use case is traceable logs for health data management [10, 47, 63]. Similar mechanisms, like [18, 97] or [68], can be used for VGI platforms to provide an immutable trace of the collected data since the rights to the collected data can be preserved for the collector. An example of copyright preservation for shared spatial data using IPFS and blockchain can be found in work done by [97]. The benefits of the DSDS for VGI platforms are as follows: (1) provides an immutable trace of the changes in GI and, as a result, clear data ownership and avoids the wiki effect, and (2) provides individual-level control of the data which results in a more applicable VGI platform for small communities. There are examples of VGI projects (see [98]) that use a distributed network of nodes to share sensor data. An example of such project architecture can be found in Figure 4 (e), (f). Such approaches use two sets of blockchains, a full blockchain to retain a history of the sensor data and a user blockchain to retain only the latest version of sensor data. Spatio-temporal data can be stored on IPFS, and the query process can be done using spatial indices. A smart contract can handle querying data from IPFS and returning query results to the user node.

In the above examples, smart contracts have been used as a set of functions that are responsible for performing the business logic of data sharing. They can be used as a small program to either control access or share the IPFS hashes or provide other logic to allow for adding new blocks to the chain. The current versions of popular programming languages for smart contracts, such as Solidity, do not support native geometry objects. In order to use geographic data in smart contracts, we can use discrete global grid systems (DGGS) based geohashes to perform geographic analysis inside these applications without the need to transfer/process high-resolution geographic information. In addition, it is possible to perform algebraic functions in smart contracts, which makes it possible to use regular geometry objects.

Table 2 shows a summary of the different types of datasets and potential approaches to sharing them using DSDS.

Type of the data	Sharing approach	Addressed challenges
Official institutional	Centralized authority	
Enterprise	Centralized authority	
Commercially available	Centralized authority	
Internally available	Centralized authority	
Research including PGIS and VGI	Store on IPFS and use blockchain for access control and licensing of different stakeholders	Access control Transparency for PGIS Projects, Monetization of the contribution for VGI
Personal and communities of interest data	Store on IPFS and blockchain with decryption	Personal data access control and transparency at the community level
Behavioral	Once the data product producer accesses the personal data then can generate a new product and that product can go under other types of data	Trust and access control
Freely available	Can be shared on IPFS without any decryption, However, it needs data access gateways	Transparency
Hybrid data	Complex ownership. This sort of data depends on the origin of each data set and their related policies and licenses. It is possible to store on IPFS and use blockchain for access control and licensing of different stakeholders.	Access control and transparency

Table 2: Different types of datasets and potential approaches to share them using DSDS.

5 GIScience research in DSDS

Distributed systems provide a transparent decision-making process by facilitating coordination and trust, and by addressing the corruption inherent to decision-making in different organizations [92]. For example, see the work done by Farnaghi et al. [27] in which they used smart contracts for a transparent and participatory site selection process. They and other authors show that smart contracts can increase the openness, transparency, and accountability of participatory planning processes by democratizing data access and keeping transaction histories on every node [3,65,79]. Similar work has been done in [46,52,53] who have used distributed ledgers in PGIS projects. Having a transparent technology guarantee the traceability and immutability of the data. Once the internal mechanism of technology is transparent it will be able to provide a trustworthy backbone to the PGIs and VGI projects.

There are also technical limitations in the distributed data sharing platforms which require more research. Spatio-temporal query processing is one of the issues which needs to be addressed. Currently, many P2P networks, such as IPFS are used as storage entities. Sharing geographic information at the feature level needs a method to address data retrieval, querying, and transformation over P2P networks. This requires the development of methods for indexing multidimensional data on distributed networks that are capable of addressing range and k-nearest-neighbour queries while also capturing the locality and directionality of the multidimensional space (e.g., see [39, 40]). P2P multidimensional query processing refers to the execution of advanced query operators over multidimensional data stored in a distributed system [83].

Another part of the technical aspects of the GIScience research of DSDS is the need to develop standards and protocols to share geographic information using distributed systems. These protocols can vary from the low-level object definition to the higher level of distributed web APIs. In DSDS the shared information is in a very heterogeneous format. Geographic information representations of phenomena in DSDS need to be handled using data models that are capable of representing complex geometric, topological, and semantic elements (e.g, see [12, 86] for other representations of geographic information) and be able to provide a level of the data masking to be able to use it to address privacy-related concerns. At the higher level, there can be standards to look at the entire P2P network as one entity and perform spatial processing using the nodes and receive processed data instead of just receiving other raw data itself. However, distributed data processing has been studied on traditional P2P networks for many years but with the current advantages, there might be a need to revisit previous approaches.

Moving toward a distributed model of spatial data sharing requires the ability to share different data from different sources and standards with digital earth (DE) platforms using a distributed approach. The use of emerging data models and technologies in DE such as DGGS means the sharing of spatial data between different providers can be facilitated. Figure 5 shows the potential relations between Digital Earth, SDS, and SDI. SDS is the connecting bridge between DE and SDI. However, SDIs are often seen as contributors to the vision of DE [62]. Despite their common components, an SDI acts mainly as a data collection coordinator and can also provide guidelines for communications between different data owners and data customers. In addition, the role of DGGS as a data model to share and integrate data from different sources can be seen as a distributed data model too. In a DGGS, each cell represents a portion of the earth's surface which can have different identities over time.

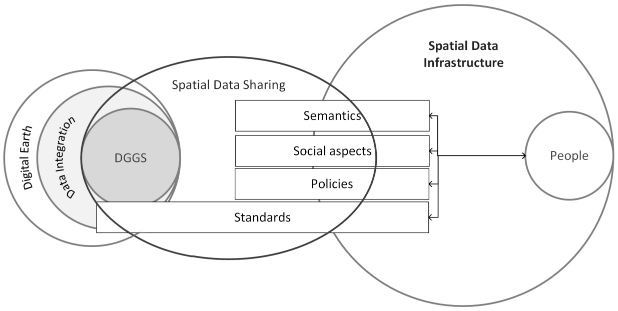


Figure 5: The interrelation of spatial data sharing, Digital Earth frameworks, and SDI.

Another challenge in DSDS is access control and data locality on distributed networks. Since data are stored on the user's devices researchers must develop protocols to provide access control based on the user's criteria (such as the user's friendship status or distance in the social network context). Distributed geoprivacy is another concept that needs more study. Once users are sharing their location with another node, they should be able to have a dynamic geoprivacy setting. Blurring geographic data in the context of distributed networks can be done by using other nodes in the network or by simply sharing data over the network with lower precision. It is worth noting that the *Right to be forgotten* is an ongoing debate and is not currently supported in the latest versions of IPFS and blockchains.

In order to provide a comprehensive data sharing platform, it is also required to have distributed license management systems and distributed data gateways/ marketplaces are required. License management systems can provide the ability for users to assign different licensing to the shared data and don't need to comply with the data sharing platform's license. These systems can store key-value pairs on the blockchains where the key is the IPFS hash of the stored information and the value is the licensing method attached to that particular dataset. Since IPFS hashes can be deconstructed into the feature level, the licensing can also be at the feature level. Due to the limitations of search engines on the IPFS network, gateways are required to access the datasets. These gateways can also act as marketplaces where users are able to monetize their contributions on different platforms. For example, [18] has used content-based addressing methods and proposed a global data market with the goal of making contributors the stockholders in the dataset(s) they create. On their platform, the ownership of the data initially belongs to the nodes who created the data. The owner is then free to transfer the ownership of the data to others. In addition, they have also provided a mechanism whereby the data owner is rewarded when their data set is used. Another similar data marketplace platform for urban applications was also developed by [68]. Such platforms clearly define and enforce information ownership [96]. Similar initiatives also help to increase collaborative projects, leading to improved data sharing, improved data quality, and improved decision-making [31].

A first requirement to make an accessible DSDS is the technical aspect of the availability of tools for P2P networks. IPFS is one of the more accessible P2P networks, and the community behind it provides many tools and SDKs to connect to the network and work with it. The recent native Internet browser support for IPFS helps to solve this gap between developers and users. Smart contracts also lack native support for geographic analysis which requires users to develop alternative methods such as geohash to work with geographic information inside the smart contracts. Another requirement is the education of user communities about how these technologies work. This can help in two ways. First, it helps developers and eases the process of integration and sharing geographic data on the P2P networks and blockchains. Second, it helps with the transparency of the technology which may contribute to building trust in the technology. Table 3 shows a summary of the requirements to have an accessible DSDS.

6 Conclusions

Sharing distributed geographic information is becoming an effort visible across different scales—from global collaborations such as digital earth and global earth observation systems to small, individual-scale—over time [94]. Data ownership and access control are two of the key aspects driving changes in the data sharing environment. Data ownership as the right to control data can be considered as a reward for the data collectors in SDS [80]. Ownership of the data can be compensation to the data providers for sharing it and as a result, incentives can improve data quality by providing responsibility and liability for data collectors [11, 82]. We believe that in the current era of big geographic data collection,

Aspect	Requirements
Social	- Transparency in the technology should exist for the users (advocates) who need to learn about their data and how their data are being stored or encrypted. Such a level of transparency can help with building up trust between small communities and enterprises or governments.
Technical	- More multi-platform SDKs to communicate with P2P networks. - Research on query processing of spatio-temporal data on distributed networks - More research on access control methods on distributed networks - Developing Geo-enabled smart contracts

Table 3: The requirements to have an accessible DSDS.

distributed data sharing and processing is necessary and there must be protocols in place to share, integrate, and process distributed data.

In this work, we addressed the current status of data ownership and access in different SDS models and demonstrated how a transition to DSDS addresses some of the existing challenges. The novelty of this paper comes from the fact that the current SDS is changing and we need to identify the challenges in current SDS practices and technologies and look forward to the socio-technological advantages of DSDS in addressing these GIScience challenges. Distributed systems provide scalability, no singular point of failure, ownership, trust, and transparency. The storage, distribution, and manipulation of spatial data are changing from a fully centralized approach (i.e., controlled by the government and/or corporations) towards a distributed, individual spatial data sharing approach, though this approach remains technically challenging. Distributed technology is in its early stages, and it requires development of other tools/methods and algorithms in order to handle, share and query geographic information. Once developed, it will be possible to contrast DSDS against other data systems, and thereby evaluate the practical benefit of such systems. A distributed data sharing platform not only needs a standard to share data between different users, but it also requires a data model that can integrate different spatial data from different sources and various accuracy levels. Such standards must be in alignment with multiple sharing platforms and frameworks.

A distributed framework for spatial data sharing is not the only solution to the aforementioned challenges, but it provides a user-centred approach to addressing some of them. As mentioned previously, technology is only one piece of the data ownership and control puzzle, albeit a critical element. It should be noted that distributed data sharing is not a universal solution. Many types of authoritative or proprietary data sets (e.g., geodetic control, cadastral boundaries, municipal addresses) will continue to require centralized access and authoring control for social, competitive, or legal reasons. In this light, DSDS and centralized systems are complementary and together enable more fine-grained and flexible data governance architectures. For example, in the digital earth context, a mix of distributed and centralized systems can be used depending on the type of data being shared (e.g., a DEM v.s. social media activity, etc.).

To sum up, we define DSDS as a new data sharing model in which individuals are active in all dimensions of data sharing: as producers, controllers, and users. The data are stored, controlled, and maintained by the data producer. The ownership and license of the data can be transferred to other users without data intermediaries. Immutable spatial data are stored on a distributed file-sharing network. At the same time, the access rights, state transitions, and version history are managed by smart contracts and stored on a blockchain.

Acknowledgments

The authors thank the Global Water Futures research program for funding this work under the Developing Big Data and Decision Support Systems theme. Thanks also to the anonymous reviewers for their thorough review and helpful suggestions for improving the publication of our research.

References

[1] International Indigenous Data Sovereignty IG. Implementing the care principles: The care-full process, 2022.

[2] Alessi, M., Camillo, A., Giangreco, E., Matera, M., Pino, S., and Storelli, D. Make users own their data: A decentralized personal data store prototype based on ethereum and IPFS. In *2018 3rd International Conference on Smart and Sustainable Technologies (SpliTech)* (2018), pp. 1–7.

[3] Atzori, M. Blockchain technology and decentralized governance: Is the state still necessary? *Social Science Research Network 1*, ID 2709713 (2015). doi:10.2139/ssrn.2709713.

[4] Bahmani, B., Goel, A., and Shinde, R. Efficient distributed locality sensitive hashing. In *Proceedings of the 21st ACM international conference on Information and knowledge management* (2012), pp. 2174–2178.

[5] Bhushan, G., and Margam, M. Access versus ownership of information and DESIDOC's balancing act. *DESIDOC Journal of Library & Information Technology 36* (2016), 320–324. doi:10.14429/djlit.36.5.10323.

[6] Bolin, K. Decentralized public ledger systems and securities law: New applications of blockchain technology and the revitalization of sections 11 and 12(a)(2) of the securities act of 1933. *Washington University Law Review 95*, 4 (2018), 955–980.

[7] Briggs, C., Burfurd, I., Duckham, M., Guntarik, O., Kerr, D., McMillan, M., and Saldias, D. S. M. Bridging the geospatial gap: Data about space and indigenous knowledge of place. *Geography Compass 4* (2020), e12542. doi:10.1111/gec3.12542.

[8] Buterin, V. Ethereum white paper: A next generation smart contract & decentralized application platform. https://github.com/ethereum/wiki/wiki/White-Paper, 2013.

[9] Camenisch, J., Leenes, R., and Sommer, D. *Digital Privacy: PRIME - Privacy and Identity Management for Europe*. Springer, 2011.

[10] CHEN, W., MU, Y., LIANG, X., AND GAO, Y. Medical data sharing model based on blockchain. *Journal of Physics: Conference Series 1267* (2019), 012014. doi:10.1088/1742-6596/1267/1/012014.

[11] CHO, G. Some legal concerns with the use of crowd-sourced geospatial information. *IOP Conference Series: Earth and Environmental Science 20* (2014), 012040. doi:10.1088/1755-1315/20/1/012040.

[12] COHN, A. G. A hierarchical representation of qualitative shape based on connection and convexity. In *Spatial information theory a theoretical basis for GIS* (1995), A. U. Frank and W. Kuhn, Eds., Springer Berlin Heidelberg, pp. 311–326.

[13] COLEMAN, D. J., RAJABIFARD, A., AND KOLODZIEJ, K. W. Expanding the SDI environment: comparing current spatial data infrastructure with emerging indoor location-based services. *International Journal of Digital Earth 9*, 6 (2016), 629–647. doi:10.1080/17538947.2015.1119207.

[14] CRAGLIA, M., BIE, K. D., JACKSON, D., PESARESI, M., REMETEY-FÜLÖPP, G., WANG, C., ANNONI, A., BIAN, L., CAMPBELL, F., EHLERS, M., GENDEREN, J. V., GOODCHILD, M., GUO, H., LEWIS, A., SIMPSON, R., SKIDMORE, A., AND WOODGATE, P. Digital earth 2020: towards the vision for the next decade. *International Journal of Digital Earth 5*, 1 (2012), 4–21. doi:10.1080/17538947.2011.638500.

[15] CUNO, S., BRUNS, L., TCHOLTCHEV, N., LÄMMEL, P., AND SCHIEFERDECKER, I. Data governance and sovereignty in urban data spaces based on standardized ICT reference architectures. *Data 4*, 1 (2019), 16. doi:10.3390/data4010016.

[16] DALY, A., MANN, M., AND DEVITT, S. K. *Good Data*. Lulu.com, 2019.

[17] DANIEL, D., AND IFEJIKA SPERANZA, C. The role of blockchain in documenting land users' rights: The canonical case of farmers in the vernacular land market. *Frontiers in Blockchain 3* (2020). doi:10.3389/fbloc.2020.00019.

[18] DAO, D., ALISTARH, D., MUSAT, C., AND ZHANG, C. Databright: Towards a global exchange for decentralized data ownership and trusted computation. *arXiv:1802.04780 [cs]* (2018).

[19] DE KLEIJN, M., VAN MANEN, N., KOLEN, J., AND SCHOLTEN, H. J. Towards a user-centric SDI framework for historical and heritage european landscape research. *Int. J. Spatial Data Infrastructures Res. 9* (2014), 1–35.

[20] DI FRANCESCO MAESA, D., MORI, P., AND RICCI, L. Blockchain based access control. In *Distributed Applications and Interoperable Systems* (2017), L. Y. Chen and H. P. Reiser, Eds., Lecture Notes in Computer Science, Springer International Publishing, pp. 206–220. doi:10.1007/978-3-319-59665-5_15.

[21] DRAKONAKIS, K., ILIA, P., IOANNIDIS, S., AND POLAKIS, J. Please forget where I was last summer: The privacy risks of public location (meta)data, 2019.

[22] ELWOOD, S. Grassroots groups as stakeholders in spatial data infrastructures: challenges and opportunities for local data development and sharing. *International Journal of Geographical Information Science 22*, 1 (2008), 71–90. doi:10.1080/13658810701348971.

[23] EU. Directive 2007/2/EC of the European Parliament and of the Council of 14 march 2007 establishing an infrastructure for spatial information in the European community (INSPIRE). *Official Journal of the European Union OJ L*, 32007L0002 (2007).

[24] EVANS, D. S. *Economic Aspects of Bitcoin and Other Decentralized Public-Ledger Currency Platforms*. Social Science Research Network, 2014.

[25] EWIS, E.-S., BREGT, A., AND CROMPVOETS, J. Spatial data sharing: A cross cultural conceptual model. *Research and Theory in Advancing Spatial Data Infrastructure Concepts* (2006).

[26] FAIRFIELD, J. A. T. Smart contracts, bitcoin bots, and consumer protection. *Washington and Lee University School of Law* (2014), 17.

[27] FARNAGHI, M., AND MANSOURIAN, A. Blockchain, an enabling technology for transparent and accountable decentralized public participatory GIS. *Cities 105* (2020), 102850. doi:10.1016/j.cities.2020.102850.

[28] GOETZ, M., AND ZIPF, A. *The Evolution of Geo-Crowdsourcing: Bringing Volunteered Geographic Information to the Third Dimension*. Springer, Dordrecht, 2013, pp. 139–159. doi:10.1007/978-94-007-4587-2_9.

[29] GOODCHILD, M. F. Citizens as sensors: the world of volunteered geography. *GeoJournal 69*, 4 (2007), 211–221. doi:10.1007/s10708-007-9111-y.

[30] GOODCHILD, M. F., FU, P., AND RICH, P. Sharing geographic information: An assessment of the geospatial one-stop. *Annals of the Association of American Geographers 97*, 2 (2007), 250–266. doi:10.1111/j.1467-8306.2007.00534.x.

[31] GRAINGER, A. Citizen observatories and the new earth observation science. *Remote Sensing 9*, 2 (2017), 153. doi:10.3390/rs9020153.

[32] GRANELL, C., AND OSTERMANN, F. O. Beyond data collection: Objectives and methods of research using vgi and geo-social media for disaster management. *Computers, Environment and Urban Systems 59* (2016), 231–243. doi:10.1016/j.compenvurbsys.2016.01.006.

[33] HENNIG, S., AND BELGUI, M. User-centric SDI: Addressing Users Requirements in Third- Generation SDI. The Example of Nature-SDIplus. *Geoforum Perspektiv* (Sept. 2013), Vol 10 No 20 (2011): Spatially Enabled Society. doi:10.5278/OJS.PERSK..V10I20.448.

[34] HOFMANN, F., WURSTER, S., RON, E., AND BOHMECKE-SCHWAFERT, M. The immutability concept of blockchains and benefits of early standardization. In *2017 ITU Kaleidoscope: Challenges for a Data-Driven Society (ITU K)* (2017), IEEE, pp. 1–8. doi:10.23919/ITU-WT.2017.8247004.

[35] HOJATI, M., FARMER, C., FEICK, R., AND ROBERTSON, C. Decentralized geoprivacy: leveraging social trust on the distributed web. *International Journal of Geographical Information Science* (2021), 1–27. doi:10.1080/13658816.2021.1931236.

[36] JAIN, P., GYANCHANDANI, M., AND KHARE, N. Big data privacy: a technological perspective and review. *Journal of Big Data 3*, 1 (2016), 25. doi:10.1186/s40537-016-0059-y.

[37] JEMEL, M., AND SERHROUCHNI, A. Decentralized access control mechanism with temporal dimension based on blockchain. In *2017 IEEE 14th International Conference on e-Business Engineering (ICEBE)* (2017), pp. 177–182. doi:10.1109/ICEBE.2017.35.

[38] JUDGE, E., AND SCASSA, T. Intellectual property and the licensing of canadian government geospatial data: An examination of geoconnections' recommendations for best practices and template licences. *The Canadian Geographer 54* (2010). doi:10.1111/j.1541-0064.2010.00308.x.

[39] KANTERE, V., AND SELLIS, T. Handling spatial data in distributed environments. In *Proceedings of the 15th Annual ACM International Symposium on Advances in Geographic Information Systems* (2007), GIS '07, Association for Computing Machinery. doi:10.1145/1341012.1341017.

[40] KANTERE, V., SKIADOPOULOS, S., AND SELLIS, T. Storing and indexing spatial data in p2p systems. *IEEE Transactions on Knowledge and Data Engineering 21*, 2 (2009), 287–300. doi:10.1109/TKDE.2008.139.

[41] KEATING, G., RICH, P., AND WITKOWSKI, M. Challenges for enterprise GIS. *URISA Journal* (2020).

[42] KESSLER, C., AND MCKENZIE, G. A geoprivacy manifesto. *Transactions in GIS 22*, 1 (2018), 3–19. doi:10.1111/tgis.12305.

[43] KHANEGHAH, E. M., MIRTAHERI, S. L., SHARIFI, M., AND BIDGOLI, B. M. Modeling and analysis of access transparency and scalability in p2p distributed systems. *International Journal of Communication Systems 27*, 10 (2014), 2190–2214. doi:10.1002/dac.2467.

[44] KLAR, R., AND LANZERATH, D. The ethics of COVID-19 tracking apps – challenges and voluntariness. *Research Ethics 16*, 3-4 (2020), 1–9. doi:10.1177/1747016120943622.

[45] KOTSEV, A., MINGHINI, M., TOMAS, R., CETL, V., AND LUTZ, M. From spatial data infrastructures to data spaces–a technological perspective on the evolution of European SDIs. *ISPRS International Journal of Geo-Information 9*, 3 (2020), 176. doi:10.3390/ijgi9030176.

[46] LASKOWSKI, M. A blockchain-enabled participatory decision support framework. In *Social, Cultural, and Behavioral Modeling* (2017), D. Lee, Y.-R. Lin, N. Osgood, and R. Thomson, Eds., Lecture Notes in Computer Science, Springer International Publishing, pp. 329–334. doi:10.1007/978-3-319-60240-0_40.

[47] LIANG, X., SHETTY, S., ZHAO, J., BOWDEN, D., LI, D., AND LIU, J. Towards decentralized accountability and self-sovereignty in healthcare systems. In *Information and Communications Security* (2018), S. Qing, C. Mitchell, L. Chen, and D. Liu, Eds., Lecture Notes in Computer Science, Springer International Publishing, pp. 387–398. doi:10.1007/978-3-319-89500-0_34.

[48] LIU, S., QU, Q., CHEN, L., AND NI, L. M. SMC: A practical schema for privacy-preserved data sharing over distributed data streams. *IEEE Transactions on Big Data 1*, 2 (2015), 68–81. doi:10.1109/TBDATA.2015.2498156.

[49] LONGHORN, R. A. *1 "The Impact of Data Access Policies on Regional Spatial Data Infrastructure"*.

[50] MAHDAVI-AMIRI, A., ALDERSON, T., AND SAMAVATI, F. A survey of digital earth. *Computers & Graphics 53* (2015), 95–117. doi:10.1016/j.cag.2015.08.005.

[51] MALOLEY, M., WILSON, C., AND RIOPEL, S. Enabling access to arctic location based information. In *Arctic Data Committee Meeting* (7 2017), Arctic Data Committee Meeting, Arctic Data Committee Meeting.

[52] MARSAL-LLACUNA, M., AND OLIVER-RIERA, M. The standards revolution: Who will first put this new kid on the blockchain? In *2017 ITU Kaleidoscope: Challenges for a Data-Driven Society (ITU K)* (2017), pp. 1–7. doi:10.23919/ITU-WT.2017.8246988.

[53] MARSAL-LLACUNA, M.-L., AND SEGAL, M. E. The intelligenter method (i) for making "smarter" city projects and plans. *Cities 55* (2016), 127–138. doi:10.1016/j.cities.2016.02.006.

[54] MARTYN, A. The concept of land plot as a combination of smart contracts: A vision for creating blockchain cadastre. *Balt. Surv. 8* (2018), 68–73.

[55] MEMON, M., HUSSAIN, S. S., BAJWA, U. A., AND IKHLAS, A. Blockchain beyond bitcoin: Blockchain technology challenges and real-world applications. In *2018 International Conference on Computing, Electronics Communications Engineering (iCCECE)* (2018), pp. 29–34. doi:10.1109/iCCECOME.2018.8658518.

[56] MICHENER, W., VIEGLAIS, D., VISION, T., KUNZE, J., CRUSE, P., AND JANÉE, G. DataONE: Data observation network for Earth–preserving data and enabling innovation in the biological and environmental sciences. *D-Lib Magazine 17*, 1/2 (2011), 12.

[57] MOHAMMADI, H., RAJABIFARD, A., AND WILLIAMSON, I. P. Enabling spatial data sharing through multi-source spatial data integration. *GSDI* (2009), 20.

[58] MORTON, G. M. *A Computer Oriented Geodetic Data Base; and a New Technique in File Sequencing*. 1966.

[59] NATURAL RESOURCES CANADA. The spatial data infrastructure (SDI) manual for the arctic, 2018.

[60] NAZ, M., AL-ZAHRANI, F. A., KHALID, R., JAVAID, N., QAMAR, A. M., AFZAL, M. K., AND SHAFIQ, M. A secure data sharing platform using blockchain and interplanetary file system. *Sustainability 11* (2019), 7054. doi:10.3390/su11247054.

[61] NEIS, P., ZIELSTRA, D., AND ZIPF, A. The street network evolution of crowdsourced maps: Openstreetmap in germany 2007-2011. *Future Internet 4* (2012), 1–21. doi:10.3390/fi4010001.

[62] NOUCHER, M., GOURMELON, F., GAUTREAU, P., GEORIS-CREUSEVEAU, J., MAULPOIX, A., PIERSON, J., PINÃĹDE, N., PISSOAT, O., AND ROUAN, M. Spatial data sharing: A pilot study of French SDIs. *ISPRS International Journal of Geo-Information 6*, 44 (2017), 99. doi:10.3390/ijgi6040099.

[63] NUGENT, T., PETRONI, F., SMITH, B. W., AND LEIDNER, J. L. An on-chain method for automatic entitlement management using blockchain smart contracts. In *Business Information Systems Workshops* (2019), W. Abramowicz and R. Corchuelo, Eds., Lecture Notes in Business Information Processing, Springer International Publishing, pp. 255–266. doi:10.1007/978-3-030-36691-9_22.

[64] OLBRECHTS, A. Statement by the edpb chair on the processing of personal data in the context of the covid-19 outbreak. text. european data protection board-european data protection board, 2020.

[65] ØLNES, S., UBACHT, J., AND JANSSEN, M. Blockchain in government: Benefits and implications of distributed ledger technology for information sharing. *Government Information Quarterly 34*, 3 (2017), 355–364. doi:10.1016/j.giq.2017.09.007.

[66] PARKER, C., MAY, A., AND MITCHELL, V. The role of vgi and pgi in supporting outdoor activities. *Applied ergonomics 44* (2012). doi:10.1016/j.apergo.2012.04.013.

[67] PRAUSE, G. Smart contracts for smart supply chains. *IFAC-PapersOnLine 52*, 13 (2019), 2501âĂŞ2506. doi:10.1016/j.ifacol.2019.11.582.

[68] RAMACHANDRAN, G. S., RADHAKRISHNAN, R., AND KRISHNAMACHARI, B. Towards a decentralized data marketplace for smart cities. In *2018 IEEE International Smart Cities Conference (ISC2)* (2018), pp. 1–8. doi:10.1109/ISC2.2018.8656952.

[69] SCASSA, T. Legal issues with volunteered geographic information. *The Canadian Geographer 57*, 1 (2013), 1–10. doi:10.1111/j.1541-0064.2012.00444.x.

[70] SCHLOSSER, S., TONINELLI, D., AND CAMELETTI, M. Comparing methods to collect and geolocate tweets in great britain. *Journal of Open Innovation: Technology, Market, and Complexity 7*, 1 (2021). doi:10.3390/joitmc7010044.

[71] SEHRA, S., SINGH, J., AND RAI, H. Assessment of openstreetmap data - a review. *International Journal of Computer Applications 76* (2013). doi:10.5120/13331-0888.

[72] SHAFAGH, H., BURKHALTER, L., HITHNAWI, A., AND DUQUENNOY, S. Towards blockchain-based auditable storage and sharing of IoT data. *arXiv:1705.08230 [cs]* (2017).

[73] SHEN, C., AND PENA-MORA, F. Blockchain for cities–a systematic literature review. *IEEE Access 6* (2018), 76787–76819. doi:10.1109/ACCESS.2018.2880744.

[74] STEFANIDIS, A., CROOKS, A., AND RADZIKOWSKI, J. Harvesting ambient geospatial information from social media feeds. *GeoJournal 78*, 2 (2013), 319–338.

[75] STEICHEN, M., FIZ, B., NORVILL, R., SHBAIR, W., AND STATE, R. Blockchain-based, decentralized access control for IPFS. In *2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE*

Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData) (2018), IEEE, pp. 1499–1506. 10.1109/Cybermatics_2018.2018.00253.

[76] SUNYAEV, A. *Distributed Ledger Technology*. Springer International Publishing, 2020, pp. 265–299. doi:10.1007/978-3-030-34957-8_9.

[77] SWAN, M. *Blockchain: Blueprint for a New Economy*. "O'Reilly Media, Inc.", 2015.

[78] TANENBAUM, A. S., AND STEEN, M. V. *Distributed Systems: Principles and Paradigms*, 2 edition ed. Pearson Prentice Hall, 2016.

[79] UNDERWOOD, S. Blockchain beyond bitcoin. *Communications of the ACM 59*, 11 (2016), 15–17. doi:10.1145/2994581.

[80] VAN ALSTYNE, M., BRYNJOLFSSON, E., AND MADNICK, S. Why not one big database? principles for data ownership. *Decision Support Systems 15*, 4 (1995), 267–284. doi:10.1016/0167-9236(94)00042-4.

[81] VERDUCCI, S., AND SCHRÖER, A. *Social Trust*. Springer US, 2010, pp. 1453–1458. doi:10.1007/978-0-387-93996-4_68.

[82] VERPLANKE, J., MCCALL, M. K., UBERHUAGA, C., RAMBALDI, G., AND HAKLAY, M. A shared perspective for PGIS and VGI. *The Cartographic Journal 53*, 4 (2016), 308–317. doi:10.1080/00087041.2016.1227552.

[83] VLACHOU, A., DOULKERIDIS, C., NØRVÅG, K., AND KOTIDIS, Y. Peer-to-peer query processing over multidimensional data.

[84] ŠUMARADA, R. Legal issues regarding spatial data. *FIG XXII International Congress* (2002), 10.

[85] WADHWA, S., AND GUPTA, P. Distributed locality sensitivity hashing. In *2010 7th IEEE Consumer Communications and Networking Conference* (2010), pp. 1–4. doi:10.1109/CCNC.2010.5421655.

[86] WANG, F., AND HALL, G. B. Fuzzy representation of geographical boundaries in gis. *International Journal of Geographical Information Systems 10*, 5 (1996), 573–590. doi:10.1080/02693799608902098.

[87] WANG, S., AND LIU, Y. Teragrid giscience gateway: Bridging cyberinfrastructure and giscience. *International Journal of Geographical Information Science 23*, 5 (2009), 631–656. doi:10.1080/13658810902754977.

[88] WANG, S., YUAN, Y., WANG, X., LI, J., QIN, R., AND WANG, F.-Y. An overview of smart contract: Architecture, applications, and future trends. In *2018 IEEE Intelligent Vehicles Symposium (IV)* (2018), pp. 108–113. doi:10.1109/IVS.2018.8500488.

[89] WANG, S., ZHANG, Y., AND ZHANG, Y. A blockchain-based framework for data sharing with fine-grained access control in decentralized storage systems. *IEEE Access 6* (2018), 38437–38450. doi:10.1109/ACCESS.2018.2851611.

[90] WEHN, U. The distribution of spatial data: Data sharing and mediated cooperation. In *Inside the Communication Revolution: Evolving Patterns of Social and Technical Interaction*. Oxford University Press, 2002, pp. 186–205. doi:10.1093/acprof:oso/9780198296553.003.0009.

[91] WILKINSON, M. D., DUMONTIER, M., AALBERSBERG, I. J., APPLETON, G., AXTON, M., BAAK, A., BLOMBERG, N., BOITEN, J.-W., DA SILVA SANTOS, L. B., BOURNE, P. E., ET AL. The fair guiding principles for scientific data management and stewardship. *Scientific data 3*, 1 (2016), 1–9.

[92] WRIGHT, A., AND DE FILIPPI, P. Decentralized blockchain technology and the rise of lex cryptographia. *Social Science Research Network 1*, ID 2580664 (2015). doi:10.2139/ssrn.2580664.

[93] YANG, C., LI, W., XIE, J., AND ZHOU, B. Distributed geospatial information processing: sharing distributed geospatial resources to support digital earth. *International Journal of Digital Earth 1*, 3 (2008), 259–278. doi:10.1080/17538940802037954.

[94] YANG, C., AND RASKIN, R. Introduction to distributed geographic information processing research. *International Journal of Geographical Information Science 23*, 5 (2009), 553–560. doi:10.1080/13658810902733682.

[95] YANG, C. P., AND TAO, C. V. *Distributed Geospatial Information Service*. Springer, 2006, pp. 103–120. doi:10.1007/3-540-31305-2_5.

[96] YERMACK, D. Corporate governance and blockchains. *Review of Finance 21*, 1 (2017), 7–31. doi:10.1093/rof/rfw074.

[97] ZHANG, Y., TANG, Z., HUANG, J., DING, Y., HE, H., XIA, X., AND LI, C. A decentralized model for spatial data digital rights management. *ISPRS International Journal of Geo-Information 9*, 2 (2020), 84. doi:10.3390/ijgi9020084.

[98] ZIA, M. B-drive: A blockchain based distributed iot network for smart urban transportation. *Blockchain: Research and Applications 2*, 4 (2021), 100033.

[99] ZOOK, M. A., AND GRAHAM, M. The creative reconstruction of the internet google and the privatization of cyberspace and DigiPlace. *Geoforum 38*, 6 (2007), 1322–1343. doi:10.1016/j.geoforum.2007.05.004.