Survey on Digital Sovereignty and Identity: From Digitization to Digitalization

KHENG LEONG TAN and CHI-HUNG CHI, Digital Trust Centre, Nanyang Technological University, Singapore
KWOK-YAN LAM, School of Computer Science and Engineering and Digital Trust Centre, Nanyang Technological University, Singapore

Through digital transformation, lots of personal data are captured, but individuals often do not have ownership or control over them. This results in the emerging Web 3.0, where people demand data sovereignty. There are actually two conceptually related terms, data sovereignty and digital sovereignty. This paper first explains these two concepts in terms of their points of focus, guiding principles, laws and regulations requirements, and then analyses the requirements and technical challenges of their implementation. To understand the emerging trend shift in digital sovereignty towards individuals taking control of security and privacy preserving over their own digital assets, this paper conducts a systematic review and analysis on Self-Sovereign Identity (SSI), which is a user-centric decentralized model and autonomy for an individual to self-determine the access and use of one's identity and credentials. The review covers existing SSI solutions and points out that an efficient key management system, the scalability and interoperability of the solution, and a well-established standard are some of the challenges for SSI deployment. Finally, the paper concludes with open issues about digital identity, including dynamic attributes, persona, and attribute ownership, that challenge the current reference architecture of SSI as well as its implementation.

CCS Concepts: • **Security and privacy → Social network security and privacy** • **Social and professional topics → Privacy policies**; **Government technology policy**; **Cultural characteristics** • **General and reference → Surveys and overviews**;

Additional Key Words and Phrases: Self-sovereign identity, privacy, identity management, digital identity, digital sovereignty

ACM Reference format:
Kheng Leong Tan, Chi-Hung Chi, and Kwok-Yan Lam. 2023. Survey on Digital Sovereignty and Identity: From Digitization to Digitalization. *ACM Comput. Surv.* 56, 3, Article 61 (October 2023), 36 pages.
https://doi.org/10.1145/3616400

This research is supported by the National Research Foundation, Singapore and Infocomm Media Development Authority under its Trust Tech Funding Initiative and Strategic Capability Research Centres Funding Initiative. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of National Research Foundation, Singapore and Infocomm Media Development Authority.
Authors' addresses: K. L. Tan (Corresponding author) and C.-H. Chi, Digital Trust Centre, Nanyang Technological University, 50 Nanyang Drive, Research Techno Plaza, X-Frontiers Block, #03-01, Singapore 637553; e-mails: {khengleong, chihung.chi}@ntu.edu.sg; K.-Y. Lam, School of Computer Science and Engineering and Digital Trust Centre, Nanyang Technological University, 50 Nanyang Ave, Block N4 #02a, Singapore 639798; e-mail: kwokyan.lam@ntu.edu.sg.

0360-0300/2023/10-ART61
https://doi.org/10.1145/3616400

1 INTRODUCTION

For decades, cyberspace has already become the operational environment for technology innovations such as smart cities, industry 4.0, and FinTech. Advances in emerging technologies, including 5G telecommunication, **Internet of Things (IoT)**, data analytics, and **artificial intelligence (AI)**, have resulted in the acceleration of digital transformation with the pervasive adoption of digitalization in the economy and society. *Digitization* refers to the process of encoding analog data into machine-readable digital formats for storage, transmission, and information processing. With the movement toward big data, the amount of data generated and collected in the digitization process has increased exponentially. Very often, these data are people's personal data that are closely related to their online activities and behaviors; they are identifiable via people's online identity. This leads to *digitalization*, which is an eventual technological trend that leverages technologies to transform the digitized data into some quantifiable knowledge that can be used for decision making.

With cyberspace being interwoven into the fabric of people's daily lives, a lot of personal data are being captured inside. These data often cover individual personal information in the form of **personally identifiable information (PII)**, online profiles, activities, and behaviors. In other words, it is an individual's digital shadow or identity relating to her/his life in cyberspace. Currently, these data are often in the hand of big third-party technology companies such as Google and Amazon, who use their online social and e-commerce platforms to collect these user data for big data analytics and Artificial Intelligence (AI) related activities such as business decision making and personalized recommendations. From the data owners' perspective, consumers do not have much "content right" (or sovereignty over content), which is defined as the capability of a data owner to decide the destiny of his/her data. They might not be able to define how their data privacy should be preserved, nor to decide how their data should be used, shared, and interpreted.

Even worse, the problem of data sovereignty or consumer content rights is much more complicated than it appears. With advances in IoT, data analytics, and AI, people's awareness and demand for privacy preserving and sovereignty on their own data increase. At the same time, people's mindset on data protection is also shifting, from securing data to preserving privacy of data [1] and finally sovereignty [2] over data. This evolution of user requirements is rooted from the people's willingness to open up and share their data actively, as is shown in social network for end-users and data sharing network from the government (e.g., data.gov.xx, where xx is the abbreviation of a country such as ca, sg, and au), and the need of collaboration among different data owners for intelligence and value co-creation. This kind of mindset shifting is also reflected in the evolution of the web, from Web 1.0, Web 2.0, to Web 3.0.

Web 3.0 (Web3) [147] is defined as the third generation of the evolution of web technologies; it is the foundational layer for how the internet is used, providing website and application services. Web 3.0 is still evolving and being defined, and as such, there is not a canonical, universally accepted definition. What is clear, though, is that Web 3.0 will have a strong emphasis on decentralized applications and make extensive use of blockchain-based technologies. As a matter of fact, the three fundamental characters of Web 3.0 that people agree on are decentralization, blockchain, and token economics. One unique characteristic of Web 3.0 over the previous generations is that while Web 1.0 focuses on "read" only, Web 2.0, on "read/write", and Web 3.0 on "read/write/own". In other words, data and digital sovereignty is a distinct feature for Web 3.0.

With the omnipresent digital literacy in smart nations, digital sovereignty is getting increasing attention and demand from the perspective of people's ownership, custody, and control over their own digital assets to the inalienable user data rights of exercising self-determination on their data

destiny. It is observed that such a shift of empowering an individual to have complete control over their digital data, including her/his discretion on what to share and with whom, called the digital sovereignty, is often built on top of the concept of digital identity [3]. A *digital identity* of a person is a set of validated digital attributes and credentials of the person in the digital world, like a person's identity in the real world [4–7]. It gives a comprehensive description of a person, covering both his/her basic properties such as gender, age, identity card number, phone number, ..., and so on, and the dynamic behavior of the person (e.g., online and browsing patterns on the web). At this moment, one important use of digital identity is to get a subset of the attributes as identifier for authentication and authority of the person in cybersecurity [8–13]. Note that however, the impact of digital identity has already gone beyond cybersecurity to explainable AI.

The increasing demand of sovereignty on digital identity results in the research on **"Self-Sovereign Identity" (SSI)**. SSI is a new decentralized identity model that has the potential to solve the problems of digital identification and authentication, and to give individuals full control of their digital identity. Research efforts of SSI in recent years result in the development of standards, protocols, reference architecture, and technologies that accelerate and facilitate the widespread adoption and deployment of digital identity.

This paper focuses on the sovereignty aspects of the extended concept of digital identity, performs a comprehensive survey on data and digital sovereignty, and conducts a systematic literature review on SSI to understand the current state of the art of this paradigm. The main contributions of this paper can be summarized as follows:

(a) This paper defines the concept behind digital sovereignty and provides an in-depth understanding about the nature of digital sovereignty and its importance to support digital transformation in the society/economy. It points out the ongoing emphasis and existing research of this topic, which covers from policymaking, regulations, and law enforcement to technical challenges of implementing digital sovereignty system platform.
(b) It introduces the data sovereignty concept which is often related to indigenous data governance on the right of indigenous peoples. This provides a more complete and holistic view of data sovereignty from the early days of the wishes and desires of the indigenous people to the challenges for the execution of the sovereignty rights.
(c) It conducts a systematic review on data and digital sovereignty as well as the implementation model of SSI based on a set of research questions as the review framework. This set of research questions is not restricted to the architecture implementation of traditional identity concept but also investigates the sovereignty requirements and guidelines of defining digital identity. Both the fundamental research questions and open challenges will be discussed.

The rest of the paper is organised as follows: Section 2 and 3 present a comprehensive systematic review of data sovereignty and digital sovereignty, respectively. Section 4 presents a detailed systematic review of SSI covering the research methodology adopted, the **research questions (RQs)** and their corresponding results. Section 5 gives a summary of the related work. Finally, Section 6 concludes the paper.

2 DATA SOVEREIGNTY

The notion of data sovereignty is actually not new. It often focuses on the struggle of indigenous peoples to reclaim sovereignty over their land, culture and heritage. Data sovereignty, according to Wikipedia [17], "is the idea that data are subject to the laws and governance structures within the nation it is collected". Earlier research on this topic was mobilized by indigenous scholars who

Fig. 1. Word cloud of "data sovereignty" papers.

voice the rights of indigenous peoples concerning the collection, ownership and application of data about their people, lifeways and territories [18]. Thereafter, a survey by Couture and Toupin [19] shows the trend of growing interests on the research topic "data sovereignty" since 2015 to 2018. In this period, the academic and non-academic literature contributions are 89 and 2,459, respectively. We also did a search in Proquest central on the keyword "data sovereignty" from 2018 till 2021 and after filtering out the non-relevant papers, we narrowed down to a collection of 263 academic peer-reviewed, full-text papers. A sample word cloud generated from these 263 papers is shown in Figure 1. Prominent words observed from the word cloud are 'indigenous', 'control', 'access', 'policy', 'personal' and 'protection'. This aligns with the long-term research that situates on the larger struggle and wishes of indigenous peoples, especially in the Anglo settler–colonial CANZUS countries (Canada, Australia, New Zealand and the United States) [20], to reclaim sovereignty over their land, body, and culture. The movement toward open data and open science does not fully engage well with the indigenous people's rights and interests. Their voices and rights are missing with respect to the concern of the "collection, ownership and application of data about their people, lifeways and territories" [21]. There needs to establish indigenous frameworks and executable principles on the collection and/or creation of indigenous data, to assert greater control over the application and the use of indigenous data and knowledge for collective benefit, and to facilitate increased data sharing among entities. This provides the initial humanity and sociological science research perspectives on data sovereignty.

With advances in technologies and access to the internet, coupled with the digital transformation for our society and economy, the discussion on data sovereignty has been expanding to cover a wider scope. According to the World Bank [22], by 2022, global internet traffic is expected to reach 150,000 GB of traffic per second, a 1,000-fold increase compared with the 156 GB in 2002. As organizations become increasingly global and interconnected, issues related to data security and privacy have become more complex. Questions like where the data actually resides, who owns it, and what regulations should it be subject to, …, and so on, all become critical considerations for organizations in compliance with the regulations of the country in which business is conducted. However, data should also be governed by regulations specific to the region from which the data originates. Failing to meet the compliance can result in hefty fines or criminal offenses. This was evident in the word cloud from words like 'governance', 'privacy', 'regulation', 'security', 'platforms', 'infrastructure' and 'cybersecurity'.

Usually, data sovereignty can be understood from two aspects. The first aspect is on the working principles. Data sovereignty will inevitably be related to the discussion on indigenous people's rights and wishes to have control over the access of their data. Two basic principles are proposed:

the CARE and FAIR principles. The CARE principles for Indigenous Data Governance are created to advance the legal principles underlying collective and individual data rights in the context of the **United Nations Declaration on the Rights of Indigenous Peoples (UNDRIP)** [23]. **CARE** is an acronym that stands for **Collective Benefit, Authority to Control, Responsibility, Ethics** [24]. It is the first attempt to outline collective rights as part of the movement to open up data. It provides external data stakeholders with guidance and advice on governance practices and stewardship responsibilities for indigenous data. While CARE can be considered as part of the open data movement, the principles are still at the conceptual level. There needs to develop criteria and tools to implement the principles, in particular those related to data sharing standards such as **FAIR (findable, accessible, interoperable, reusable)** [25] by considering power differentials and historical contexts. Developed in the Netherlands in 2015, the FAIR principles have since been taken up as a way of sharing data that will maximize the use and re-use of data. The rationale behind FAIR is to support data and knowledge integration and to promote sharing and re-use of data. Compared to CARE, the FAIR principles are more aligned to technical principles from an implementation viewpoint. They are related to how data can be searched, retrieved, shared, transferred, and reused across different IT systems, platforms, and environments. Technology will be a critical mean to operationalize these principles and to facilitate the execution of indigenous people's sovereignty and self-determination. In essence, the principles for indigenous data governance require enacting FAIR, but with CARE.

The second aspect is related to the regulatory requirements. Data sovereignty is closely linked to the laws and regulations of the countries where the data resides. It refers to the concept that the data an organization collects, stores, and processes are subject to the nation's laws and general best practices on where it is physically located. Another closely related term is data residency. Data residency is what a business or government specifies the geographical location where its data should be stored. Thus, at its core, data sovereignty is about protecting sensitive and private data, ensuring it remains under the control of its owner in the specified country. In simple terms, this means that a business has to store the personal information of its customers in a way that complies with all the data privacy regulations, best practices, and guidelines of the host country. With regulations like the European Union's **General Data Protection Regulation (GDPR)** [26] setting the bar for data privacy protection, it will be more important than ever for organizations to proactively safeguard their sensitive customer and employee data – where and how data is stored and shared. In the U.S., the **California Consumer Privacy Act (CCPA)** [27] has similar objective to give California residents greater control over how their data is used and stored.

The principle of the laws on data sovereignty is that data belongs to the jurisdiction of the nation-state where it is originally held in binary form. Earlier research efforts on data sovereignty mainly focused on the humanity and sociological aspects conducted by researchers who work with the data (or asset) owners. However, data sovereignty principles such as CARE and FAIR lack implementation and technical details to execute them. The same argument goes to the defined or legislated laws revolving around data sovereignty and intended protection, which also require technology support to enforce them and to execute the regulations and policies automatically. In summary, due to the lack of operational details, implementation mechanisms and enacting systems to execute and realize the data sovereignty principles and laws, it is a big challenge to govern the stewardship and application of data to fully assert sovereignty on the locally hosted data that the laws are designed to protect. With digital transformation that enables data to transcend geopolitics and economics, nation-states will need to reassert their authority to protect their citizens and businesses. They need to look at data sovereignty over a wider digitalized spectrum, that is, digital sovereignty. Thus, an understanding of what digital sovereignty is and its existing research is necessary.

Fig. 2. Word cloud for "digital sovereignty" papers.

3 DIGITAL SOVEREIGNTY

Digital sovereignty is a key issue of concern and challenge in digital economy and society. Digitalization has already become a major driver that transcends geopolitics and economics with waves of technological innovations, especially in AI. This acceleration in digitalization or digital transformation is aligned with the realization of the power of data in which a complex ecosystem of entities collects, analyzes, and trades the value that may be extracted from the collected data. In this transformation, despite the fact that lots of personal data are captured, individuals usually have no ownership or control over them, thus resulting in the privacy and sovereignty issue. It is therefore necessary to have control over the ownership, custody, and utilization of these data, and to protect one's digital assets and identity through proper data governance, cybersecurity control and privacy protection.

According to the definition document drafted by the EU Federal Chancellery [22]: "Digital sovereignty describes the ability to shape the digital transformation in a self-determined manner with regard to hardware, software, services, and skills. Being digitally sovereign does not mean resorting to protectionist measures or doing everything yourself. Being digitally sovereign means, within the framework of applicable law, making sovereign decisions about the areas in which independence is desired or necessary."

A survey by Couture and Toupin [19] shows the research focuses on the topic of "digital sovereignty" and the trend in the interests under this topic. In the period of 2015 to 2018, the academic and non-academic literature contributions are 22 and 239, respectively. We did a similar search in Proquest central using the keyword "digital sovereignty" from 2018 till 2021 and narrowed down to a collection of 39 relevant academic peer-reviewed, full-text papers. Compared to data sovereignty, digital sovereignty is still a relatively new topic area. A sample word cloud on these 39 papers is shown in Figure 2. A glance at the word cloud can reveal keywords such as governments, national, states related words like 'commission', 'Europe' ('European'), 'political' and 'regulation' which align with a nation's desires to have sovereignty over the control and autonomy of the data. The focuses are reflected in prominent words like 'technology', 'security', 'internet', 'infrastructure', 'services', 'cloud', 'platforms', 'privacy' and 'cybersecurity'. It is observed that recent papers have shifted their discussions and focuses to the execution of digital sovereignty and their corresponding implementation technologies, infrastructures, and cloud platforms. This is in sync with the time (April 2016) when Christopher Allen discussed in his blog on digital identity issues, internet identity and his vision which he called "Self-Sovereign Identity" [140].

There are also security and cybersecurity concerns relating to privacy and data integrity as a result of the AI movement [28]. In our study, we observe that EU is commonly cited and linked to their efforts of asserting digital sovereignty with implementations of regulations. A commonly cited regulation in supporting digital sovereignty requirement is the European General Data Protection Regulation (GDPR) which grants and protects privacy rights to individuals located in the EU when their personal data is processed by non-EU companies that offer goods or services to them or monitor their behaviors.

Most of the papers surveyed view digital sovereignty from the government and regulatory perspectives - the government as the central authority to protect its citizen's privacy and data rights. On closer examination of the word cloud, words like 'individual' are also observed. This suggests that the research focuses on individuals' human rights to have personal control on the ownership and ethical use of their data. Compared to the word cloud generated for data sovereignty in Figure 1, the words 'personal', 'control' and 'ownership' are also visible, suggesting the need of individuals to have sovereignty over their data, whether in analogue or digitalized form. Thus, the concepts of digital sovereignty are not limited to the control of the state over the use and design of critical digital systems and the data generated and stored therein, but also for its people, on the individual level [19], to regain custody and control over the security and privacy of their own data.

From the computer science research point of view, digital sovereignty requirements need to be aligned with data security and privacy issues [29–31] and require the adaptation of the information security triad principles: confidentiality, integrity, and availability. Digital sovereignty on an individual basis cannot be ignored. Therefore, in the next two sub-sections, digital sovereignty is going to be viewed from two perspective levels: the national level and the individual level [27, 32–34].

3.1 National Level

On the national or state level, the term digital sovereignty has been used by the governments to convey the idea that states should reassert their authority over the internet and protect their citizens and businesses by putting up regulatory requirements on the digital infrastructure within their territories and the data of their people inside these infrastructures. However, with the dominant position of big tech companies such as the **'GAFAM' (Google, Apple, Facebook, Amazon, Microsoft)** in the provisioning of cloud computing and social media, data of citizens and companies are now stored in the cloud of these big tech companies. Thus, nation-states of non-big tech companies would like to reassert their authority [35] and control over the data and its usage. In the case of EU, in recognizing the importance of retaining this sovereignty and to constrain the 'platform power', the EU has channeled its efforts into tighter and more comprehensive regulations over the tech sector. To some extent, this can be viewed as a full-blown defense of Europe's digital sovereignty, making Europe to become the world's leading "regulatory superpower" [33]. In addition, governments in the European Council, more specifically Germany, announced their intention "to establish digital sovereignty as a leitmotiv of European digital policy" [32].

This movement attempts to reassert the government's authority over cyberspace [25] and to protect their citizens and businesses from the challenges of not having control over their data. Two notable EU project initiatives are the GAIA-X cloud initiative by Germany with the support from France, and the European Cloud Federation initiative. The GAIA-X cloud initiative, launched in 2019 [36] by the German Federal Government, aims to create its own European offering of cloud infrastructure, services, with data being explicitly guided by the principles of sovereignty-by-design. Under this design, the customer has full control over the storage and processing of the data and access thereto. The infrastructure aims to meet the highest standards in terms of

digital sovereignty and foster innovation. GAIA-X is in line with the existing EU's GDPR that establishes key privacy requirements of handling data for European individuals or businesses. The European Cloud Federation initiative sets the standards for interoperability between providers and portability of data, where cloud providers will be expected to offer a choice as to where (personal) data are stored and processed, without otherwise requiring storage in Europe.

From the digital sovereignty perspective, the GAIA-X project is a more viable and promising initiative, but it requires a high degree of commitment and coordination from the EU member states. With the lack of binding European cloud policies, the cloud choices available to EU member states are through outsourcing, resulting in the requirements only subject to the country where the cloud services are provided. Outsourcing also brings the risk of vendor lock-in and this threatens digital sovereignty [37]. Usually, major market players only offer limited interoperability and portability of data and applications. They can use their own standards and build their own private internet infrastructure, making any interconnection with other parties difficult, both in terms of infrastructure and data exchange. In addition, dependence on foreign providers might result in the control and requirements from other countries, which might have different rules concerning espionage, privacy, and government access to data.

From the cybersecurity perspective, undermining digital sovereignty results in risks such as the lack of control over the platforms and environment where the data reside and the level of data protection inside. Even though extra requirements can be put in the service level agreement as remedy, the nation-states will inevitably face a spectrum of threats to their data, ranging from systematic theft of intellectual property, digital extortion, targeted misinformation, systematic infiltration of social media [34], to the influence on elections and democratic processes. For digital sovereignty at the national level, to ensure the controls of the digital assets that the nation-states seek to protect, there is a need for them to re-evaluate their data infrastructure and digital policy that contribute to the development and execution of digital governance. At the same time, they also need to cope with its geo-economic and geopolitical implications to foster multilateral cooperation [35] at the international level.

3.2 Individual

Digital sovereignty is not only the wishes of nations but is also the wishes of their citizens as individuals. Emphasizing the importance of individual self-determination, digital sovereignty should focus on the autonomy of citizens, in their roles as employees, consumers, and individual users of digital technologies and services, to determine who gets access and uses their data [32]. An interesting aspect of this is the departure from a state-centered understanding of sovereignty to the individual perspective. Instead of viewing sovereignty as the prerequisite to exercise authority under the government regulation and policies in a specific territory, digital sovereignty should be viewed as the ability of an individual to take actions and decisions in a conscious, deliberate and independent manner over the access and handling of her/his data [31, 32], that is, the self-sovereignty over one's data [27].

Digital transformation has brought data to the core of innovation that requires sharing, exchange, and learning of the data that transcends through geopolitics and economics, for example, in AI [28] or IoT [38]. The data being shared or utilized by these innovative services are often personal and sensitive in nature. And individuals should have sovereignty to control and authorize the usage and access of their data, and the details of the disclosure [33]. In other words, digital transformation and utilization of these personal and sensitive data challenge the privacy and sovereignty of an individual.

Currently, big tech companies (e.g., GAFAM) are in a unique position to collect, harvest, and analyze data generated through the online activity of an individual user on their platforms [34].

There are significant concerns about the handling of these sensitive and personally identifiable data [27] because these data can provide crucial insights about individuals' behavior and online content consumption. There are also concerns that GAFAM's harvesting of data can open the door to manipulation of online public discourse. To make the situation worse, individuals' personal data and information are often captured and harvested by these corporations without the knowledge and approval from the data owners. All these explain why digital sovereignty is important to an individual. As a data owner, it is important for him/her to exercise her/his rights and control the usage and sharing of his/her data [32].

Another fundamental consideration of regaining control of an individual's data is related to the identity management for individuals [27]. Identity is the key to put access control policies in place, identify where data assets are stored, what can be accessed, and for establishing trust between parties. To an individual, the security and privacy protection of his/her online identity is of particular concern, and this has direct connection to the execution of one's data sovereignty right. To address the increasing need for online identification, there has been a proposal to create an **eID (electronic identity or digital identity)** of each citizen by the government. This is highly relevant in EU, where citizens are expected to have higher privacy preservation as users of online platforms. However, the solution being offered currently is mostly for access only to digital government services [37], while the authentication in the private domain is still left to the major foreign platforms, such as Facebook, Apple, Amazon, Google, Alibaba, or Tencent. Furthermore, in a centralized identity management system, individuals need to rely on a central authority to keep safe of their personal information that is used to verify their online identities. Cases of identity theft are prevalent in the past few years when identity data stored in central authority platforms are stolen and used for malicious purposes. On top of this, the central authority can also monitor the online transactions of an individual based on the identity verification requests that it receives for that individual. With these data likely to be kept by the platforms, this explains why individual privacy and digital sovereignty are becoming issues of concern.

In the context of sharing personal and sensitive data for use by AI and machine learning applications [28], privacy and data security, and the ability for individuals to control the level of access and disclosure of details are a deterrence to individual's decision to share. To ease the concerns, there are **privacy-preserving data mining (PPDM)** methodologies [14] to protect and preserve the privacy of data owners. The PPDM methods are designed to guarantee a certain level of privacy, while maximizing data utilization so that data mining can be performed on the transformed data efficiently and yet the privacy of the data owner is still preserved. The privacy-preserving methods work by withholding sensitive information about the data owner, and thereby controlling the selective use of information. Other than privacy preservation methods, there are also privacy-enhancing techniques that ensure the privacy of personal and sensitive data while supporting useful utilization of these data for machine learning and computation. Some of these privacy-enhancing techniques include secure multi-party computation [39], differential privacy [15], fully homomorphic encryption [16], and **zero-knowledge proof (ZKP)** [40].

Though these privacy-preserving and enhancing methods and techniques are available to protect data owners' privacy, the level of privacy is usually not decided by them but by the data custodians without seeking any consent from the data owners [41]. Most of the time, this does not align well with the privacy level that an individual, as a data owner, desires or is comfortable with. The data being shared by individuals can contain identifiable personal data that can relate and link to their personal lives and thus traceable to their online activities. Data owners should have the rights and authority to control the level and details of their own identifiable personal data to be shared. In addition, data owners should have the autonomy to decide whether they are willing to participate in the sharing, the destination of their data and how they should be used

or handled. The intention of sovereignty is not just to have individual's data records locked away in fragmented organizational silos and not easily accessible and resulting in reduced access and utility, but to leave it to individual, as the rightful owner of the data, to have the sovereignty to control and decide.

In addressing the governance of one's digitalized data, a new category of digital sovereignty claim has emerged in recent years. Emphasizing on the desire and the need of individuals to take control and claim back over the management of their digital identity (including personally identifiable information (PII) as well as traits and behavioral data that can identify the individual's online digital self), a category on self-sovereign identity (SSI) was proposed [27]. SSI is a paradigm to decentralize the storage and management of an individual's identity and credentials, and to allow him/her to maintain control over his/her identity across different services. With SSI, autonomy in managing services related to identity and credentials can be achieved. To have a deeper understanding of SSI, a systematic literature review is conducted to compile an in-depth study and analysis of the SSI paradigm.

4 SELF-SOVEREIGN IDENTITY (SSI)

SSI is a decentralized approach that empowers individuals to have complete control and ownership of their data and be able to decide what data to share and with whom. It enables an individual to have autonomy and control over their identity and personal data with self-determination on the level of data to be shared, used, and handled. SSI research has gained much research focus and interest in recent years. Hence, a detailed and comprehensive systematic literature review on SSI is conducted to provide a good understanding of this topic, covering both the current research focus and its trend.

4.1 Research Methodology

To conduct the systematic literature review of SSI, this paper uses the Kitchenham's guidelines [42] and follows the **software development lifecycle (SDLC)** methodology. The Kitchenham's guidelines cover three phases of a systematic literature review: planning the review, conducting the review and reporting the review. Software development lifecycle (SDLC) methodology refers to the continuous process which starts with the launch of a project and ends when the full exploitation is removed [146]. Usually, it covers five stages, namely, planning and requirements analysis, designing project architecture, development, testing, and deployment. With respect to this paper, the background information about SSI is first given before the SDLC phases are taken up. Then the stages that are adopted to the SDLC protocol are shown in Figure 3, with their research questions aligned to these three phases: requirements analysis, architectural design, and implementation.

4.2 Research Questions

In view of the various information, entities, stakeholders and technologies in SSI research and development, some background information of SSI is first provided before proceeding to provide a systematic literature review from a software engineering perspective. The study of SSI can be divided into three main SDLC phases, namely, requirement analysis, architectural design, and implementation. Furthermore, as is shown in Figure 3, this paper adopts the SDLC phases and groups the research questions according to the phase they belong to. The requirement analysis phase analyses the issues and problems that the SSI design needs to satisfy. The architectural design phase seeks to understand the design of the architectural structures and components for the SSI solution as well as the supporting standards and processes, Finally, the implementation phase gathers the development works and the challenges faced. Explanations on how the research questions are derived will be elaborated next.

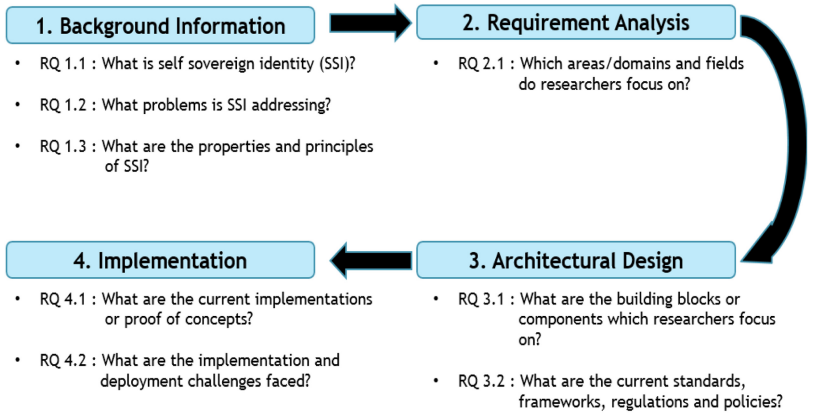

Fig. 3. Research questions mapping to software development lifecycle practice.

4.2.1 Background Information. To understand SSI, it is important to have a good understanding about the background of SSI and the problems it is addressing. Thus, there are three main questions listed below. Note that we use RQ to stand for "research question".

- RQ 1.1: What is self-sovereign identity (SSI)?
  - To understand the definitions of SSI provided by researchers.
- RQ 1.2: What problems is SSI addressing?
  - To know the benefits and objectives of SSI solution.

After learning the background and objectives of SSI and with an overview of the problems it is addressing,

- RQ 1.3: What are the properties and principles of SSI?
  - To list and explain the various principles and properties used in this field of SSI, as elaborated by the literatures surveyed in this paper.

4.2.2 Requirement Analysis. Knowing the problems that SSI wants to address,

- RQ 2.1: Which areas/domains and fields do researchers focus on?
  - To find out the areas and domains as well as the involved issues researchers are focusing and working on.

4.2.3 Architectural Design. With an overview on how SSI works in the requirements analysis, the next focus will be on the architectural design of the SSI building blocks and components and to understand what researchers focus on in their design.

- RQ 3.1: What are the building blocks and components of SSI that researchers are focusing on?
  - Architectural design may need to conform to industry standards and align with defined regulations and policies, and
  - Established frameworks may exist for adoption.
- RQ 3.2: What are the current standards, frameworks, regulations and policies for SSI?

These RQs aim to identify the possible approaches that address the requirements during development, and to extract the software components for the architecture design to fulfill them.

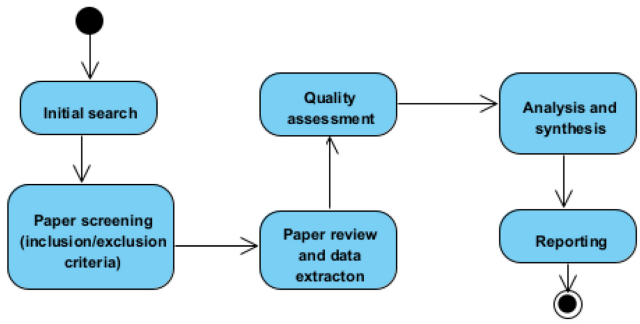

Fig. 4. Paper search and selection process map.

4.2.4 Implementation. In the implementation phase,

- RQ 4.1: What are the current implementations and proof of concepts of SSI available?
  - To know the existing implementations of SSI as well as their implementation challenges.
- RQ 4.2: What are the implementation and deployment challenges being faced?
  - To identify possible challenges that future implementation and deployment of SSI needs to address.

4.3 Sources, Selection, and Strategy

This paper searched through the following search engines and databases: (1) ACM Digital Library, (2) IEEE Xplorer, (3) ScienceDirect, (4) Springer Link, and (5) ArXiv. Figure 4 shows the adopted paper search and selection process.

4.3.1 Search String Definition. The following search terms were used to select the initial studies: ("self-sovereign identity" OR "self sovereign identity" OR "digital identity")

These three terms are used because the search wants to focus on the work related to the design, implementation, and operation/execution aspects of digital identity, and not on the conceptual understanding on the sovereignty need of digital identity by its owner. Hence, the single term "sovereign" and the abbreviation 'SSI' were omitted as they can be related to other fields, for example sovereign funds in the financial field and surgical-site infections (SSI) in the medical field. To increase the number of selected studies, the term "digital identity" which is closely related to the concepts of SSI is also included. Note that some of the digital identity papers might be filtered out later due to their relevance to the topic of interest. In addition, research showed that April 2016 was the time SSI was initially conceptualized when Christopher Allen discussed in his blog digital identity issues, internet identity and his vision which he called "Self-Sovereign Identity" [140]. Thus, we decided to base our studies on publications between 2016 and 2022. As a nascent topic, three publications mentioning SSI were found in that year. These papers attempted to provide preliminary formulation [155, 156] and critique [157] of SSI. However, since their fundamental science and system implementation perspectives were very weak, they are not included in the analysis in Section 4.4. Note that this argument is supported by other survey papers on SSI, e.g., [143, 145, 150], in which only SSI papers from 2017 onwards are included in their study. Table 1 summarizes the captured studies in the searches conducted in May 2022, starting from 2017.

4.3.2 Inclusion and Exclusion Criteria. The inclusion and exclusion criteria are formulated to effectively select relevant papers. The inclusion criteria restrict the scope of the selected studies to align with systematic review questions in this paper while the exclusion criteria remove unwanted

Table 1. Number of Selected Publications per Source on Self-Sovereign Identity

Sources	ACM	IEEE	Springer	ScienceDirect	ArXiv	Total
Paper count (initial)	26 (28)	94(100)	74	37 (43)	32 (43)	263 (288)
Paper count (filtered)	11 (13)	36 (40)	31	6(8)	13 (22)	97 (114)

studies in terms of irrelevant types, language, and related subjects. The final inclusion criteria are as follows:

- Long and short papers that elaborate on SSI and the related systems, with research works that give comprehensive explanations on the system components and functionalities.
- Survey and review papers that identify the open problems and future research trends.

The final exclusion criteria are as follows:

- Non-primary, short (less than five pages) and non-English studies.
- PhD dissertations, tutorials, editorials, and magazines.

4.3.3 Quality Assessment. A quality assessment scheme is developed to evaluate the quality of the papers. There are four quality criteria used to rate the papers:

- Citation rate. This is measured by checking the number of citations received by each paper according to Google scholar. The threshold rate for a paper to be included in the survey is five.
- Methodology contribution. The methodology contribution of the paper is measured by asking 2 questions: (1) Is this paper highly relevant to the research? (2) Is there a clear methodology that addresses its main research questions and goals?
- The sufficient presentation of the findings. Each paper is evaluated based on the availability of results and the quality of findings. The question being asked is: Are there any solid findings/results and clear-cut outcomes?

4.3.4 Data Extraction and Synthesis. All the selected papers are downloaded, and the essential information related to the papers are recorded in a data extraction sheet, including the title, source, year, paper type, venue, authors, affiliation, the number of citations of the paper, the answers for each RQ, and the research classification.

4.4 Results

Based on the sources and the selection strategy, a total of 97 papers are identified [43–139] for our study. Table 1 shows the sources of the paper; and Figure 5 shows the paper count by year and the paper type. From the figure about the paper count by year, the growing interest in SSI research can be observed. Furthermore, most of the papers focus either on the proof of concept or on the prototype implementation of the proposed SSI solutions. In the next sub-sections, the findings and the gathered statistics for the RQs are first presented before we consolidate them and provide the final analysis and discussions.

4.4.1 RQ 1.1: What is self-sovereign identity (SSI)?. Self-sovereign identity (SSI) is a term that refers to the ability of an individual to own, control, manage, and share their personal information (or credentials that make up the identity) in a secure and decentralized manner, without the need for a centralized authority or intermediary. With SSI, individuals can store their credentials in

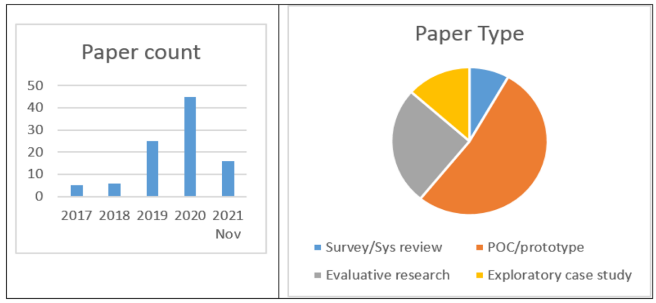

Fig. 5. Paper count by year and paper type.

Fig. 6. Word cloud for SSI papers.

a digital wallet that they control and manage. This digital wallet is secured using cryptography, and the individual can choose what information they want to share with others, and they can choose to revoke access to it at any time. The goal of SSI is to create a more secure and privacy-preserving system for managing personal data. By giving individuals control over their data, SSI can help to prevent data breaches and identity theft. SSI also has potential applications in areas such as healthcare, finance, and government services, where the ability to securely manage personal information is critical.

The first research question (RQ 1.1) is "What is self-sovereign identity?" To answer this question, the features and the breakdown of the words in self-sovereign identity reported by each study are recorded. This question helps the audience to understand: (1) what the definition of SSI is, and (2) the focuses of researchers on SSI. To get a visual understanding of these two aspects, a word cloud shown in Figure 6 is generated. It shows the frequency of the words that appear in the studied papers [51, 60, 84, 92, 97, 108–111, 113, 114, 120, 138] which provide an elaborative definition of "self-sovereign identity". The most frequently appeared and relatable words, other than the self-sovereign identity words, include: individual, control, data, user, consent, credential, disclosure, authority, autonomy, security, ownership, portability, transparency, minimal, correlation, and ecosystem.

With reference to the properties of SSI that are described above, which include decentralization, security, privacy, user control, interoperability, trust, and portability, the association of the frequently occurred words with these properties are given in Table 2.

Table 2. Association of SSI Properties with Frequently Occurred Words in SSI Papers

Property of SSI	Frequently Occurred Words in SSI Papers
Decentralization	consent, autonomy, correlation, minimal, consent
Security	data, security
Privacy	data, minimal, disclosure
User Control	individual, control, user, authority, ownership
Interoperability	data, transparency, ecosystem
Trust	data, credential
Portability	portability

4.4.2 RQ 1.2: What Problems is SSI Addressing?. SSI is seen as a solution to move away from a centralized model to a user-centric decentralized model. With the shift, it attempts to address the following problems:

(a) Identity theft and fraudulent transaction whereby a central data source captures the personal identifiable information of individual's citizenry, biometric data, and private data. Major research focuses include:
- Distributed ledger architecture and identity management
  - Distributed ledger architecture for identity management [48, 96, 138]
  - Requirements, models, and techniques for identity management in distributed ledger architecture [113,123,137]
- Passport-level legally valid identity
  - Approaches to define valid identity, including decentralized biometric [111], personal traces on smartphones [118], and proof of personhood [89]
  - Design of passport-level legally valid identity [53]
  - Case study of scam bank service calls and the proposed architecture solution [81]
- Trust environment for secure identification
  - Secure identification models, including structural and hierarchical model of identification that supports global interoperability and identification while preserving country sovereignty [57], **Self-Sovereign Identity Based Access Control (SSI-BAC)** access control model for cross-organization identity management [85], enriched digital identity model with multi-purpose and multi-origin attributes [101], and human-centric paradigm [131]
  - Case studies of blockchain-based self-sovereign identity system for open banking [87], for IoT devices [90], for refugee's handling from humanitarian perspective
- Verification and revocation of identity credentials
  - Management of dynamic identity credentials and their verification / revocation [62]
  - Revocation and offline verification [80]
- Securing data, its backup and recovery
  - Self-sovereign backup-and-restore protocol with audited by commits on a publicly accessible distributed ledger [64]
  - Decentralized key recovery digital wallet solution using Shamir's secret sharing scheme and Hyperledger Indy distributed ledger [75]
  - Decentralization of the authentication and authorization processes to overcome the single point of failure problems through the integration of a public permissioned SSI framework with a permissioned consortium BC based architecture [117]

- Decentralized accountability for trusted information sharing
  - Privacy for trusted SSI information sharing [115]
  - Study on decentralized control and accountability of identity related private data [97]
  - Case studies on blockchain-based solution for personal health data (including electronic health records) sharing [72, 134, 136], comparison of different SSI solutions [77, 98], and risk delegation and private key recovery of the mobile application [112]

(b) Alignment to local regulation and policy on managing data privacy. Major research focuses include:
- Conformance of architecture and platform design
  - Analysis of blockchain-based privacy preserving architecture with respect to regulations and policies [77]
  - Blockchain architecture of smart identity wallet for **Identity Internet of Things (IDoT)** [70], sovereign identity architecture [108], decentralized biometric-based authentication protocol for identity ecosystems [111], and user-controlled resilient identity management systems that can be used in a public transportation sector that spans different operators in multiple countries [137]
- Privacy preservation and mandate representation
  - Representation system that supports sovereignty and access right delegation [71]
  - Representation of SSI Based Access Control, and access control model for cross-organization identity management [85]
  - Case study on the enhancement of third-party trust delegated by banks [81]
  - Requirements of SSI with respect to privacy preservation [113]
- Personal data protection compliance
  - Data protection compliance in digital wallet and decentralized key recovery solution [74]
  - Data protection for SSI based human contact tracing [82]
  - Evaluation of current self-sovereign identity paradigm solutions against General Data Protection Regulation (EU) 2016/679 (GDPR) [91], SSI solutions with focus on identity proofing and authentication solutions [92], and dependency of SSI with blockchain [132]
- Social communication control and exchange of information and data, including personal and sensitive [78, 105]
  - Decentralized service architecture for social communication [78]
  - Discussion on federated identities and exchange of information in SSI [105]

- Proliferation of passwords stored in multiple systems and the risks involved. Note that this is a move towards password-less authentication. Major research focuses include:
  - Shared private data and protocol that prevent information leakage in cross-organization environment
    - Needs and requirements for shared private data protocol and system [68]
    - Design of shared private data protocol and system [43, 69, 72], decentralized service architecture for social communication [78], and decentralized biometric-based authentication protocol for identity ecosystems [111]
  - Passport-level identity within the context of mutual distrust with attestations of truth from third parties [53, 57, 70, 75, 81, 96, 109, 116]
    - Blockchain-based digital identity solution in the context of mutual distrust [53, 57, 75, 96]

- Assessment of blockchain-based digital identity solution in mutual distrust environment such as IoT [70], bank services [81]
- Evaluation of SSI as emerging identity management solution [109]
- Proposal of using SSI for decentralized identity management [116]

- IoT devices to fully own and manage their digital identities
  - Review on the challenges [55] and how [88] to use SSI for Industrial Internet of Things and IoT applications
  - Solutions of using SSI and distributed ledger technology for IoT device identity management [90, 106]
- Attribute-based identity as a means to authenticate an individual
  - Formal models and properties of attributes to be used in SSI [60]
  - Possible attributes inside an identity that can be used for authentication, including biometrics [56, 77] and human behavior signature [133]
  - Case studies for attribute-based authentication system for SSI, covering document management [58], backup and restore [64], mobile phone applications [71], digital wallet [74], multi-party ecosystems [101, 104], decentralized applications [108]
  - Comparison [98] and assessment [113] of existing blockchain-based SSI

(c) Providing digital identity as a form of verifiable document, for example as a national identity, for employment verification and even for refugees in crisis-prone countries [73, 89, 97, 131].
- Ethical issues involved using SSI for digital identity [97, 131]
- Case studies, including refugees' experience [73] and proof of personhood [89]

4.4.3 RQ 1.3: What are the Properties and Principles of SSI?. The paper that C. Allen outlined in 2016 [140] is often considered as the first draft of the properties and principles of SSI. It is highly cited by SSI community and publications, including most of the papers surveyed here. The 10 principles are existence, control, access, transparency, persistence, portability, interoperability, consent, minimalization and protection. Other principles such as provable are also proposed by [53, 69, 77, 84, 99]. The objectives of each of the 10 principles are outlined below:

- Existence: Users must have an independent existence.
- Control: Users must control their identities.
- Access: Users must have access to their own data.
- Transparency: Systems and algorithms must be transparent.
- Persistence: Identities must be long-lived.
- Portability: Information and services about identity must be transportable.
- Interoperability: Identities should be as widely usable as possible.
- Consent: Users must agree to the use of their identity.
- Minimalization: Disclosure of claims must be minimized.
- Protection: The rights of users must be protected.
- Provable: Claims must be shown to hold true.

From the SSI solution perspective, Naik and Jenkins [110] propose an extended set of 20 principles to cater to the evolving SSI requirements and standards used for the assessment of SSI solutions. These 20 principles are related to those provided by C. Allen [140] but with the addition of the following principles that relate to the SSI infrastructure and services, other than security, privacy, decentralized, availability and scalability:

- Recovery: Mechanisms must be in place to recover and re-assert identity due to complete loss of credential.

- Cost Free: Owning an identity should be free of cost or with negligible cost.
- Sustainable: An identity infrastructure and services should be environmentally, economically, technically and socially sustainable for the long term.

With respect to the categorization of the principles, Sovereign Foundation categorized Allen's principles into three groups: security, controllability, and portability [50]. Lastly, Ferdous et al. [60] analyzed the existing definitions, extracted properties and classified them into five categories, foundational, security, controllability, flexibility, and sustainability.

4.4.4 RQ 2.1: Which Areas/Domains and Fields do Researchers Focus on?. The research focus of the papers under our study is further broken down in terms of areas/domains so as to provide a better understanding of the current SSI research. The results on the areas/domains and their associated papers are listed as follows:

- Financial Banking, with focus on:
  - Electronic and cardless payment technologies [62, 81]
  - Digital asset and attribute authentication e.g., **Know-Your-Customer (KYC)** [63, 100, 101]
  - Open banking service and platform [87, 124]
- Education and certification [84, 85, 107]
- Healthcare, with focus on:
  - Decentralized accountability and solution [46, 127, 139]
  - Immunity and contact tracing platform [82, 135]
  - Patient identity and data security [95, 130, 134, 136]
- National, e-Gov, with focus on:
  - Framework and platform [57, 65]
  - Decentralized accountability [69, 72, 138]
  - eID derivation and compliance [79, 91, 96, 132]
- Transportation [137]
- IoT, with focus in the area of:
  - Home and Industrial [55, 66, 106]
  - Environmental [88, 117]
  - Smart identity and trust [70,103, 105, 126, 128]
- SSI framework and components design, with specific focus on:
  - Private data storage and security [43, 47, 75, 109, 122]
  - Key management [49, 76]
  - Authentication and authorization [54, 58, 80]
  - Decentralized platform and framework [59, 77, 83, 86, 97, 98, 102, 112, 114]
  - Verifiable credential and decentralized identifier [67, 104]
  - Trust model and governance [93, 110]
- Identity management (IdM), specifically:
  - Open challenges and evaluation [48, 50, 60, 61, 68, 74, 89, 92, 99, 113, 115, 131]
  - Using biometric and attribute-based solution [44, 45, 51, 56, 71, 94, 111]
  - Deployment platforms and authentication [52, 53, 73, 108, 116, 118, 119, 133]
  - Registration and revocation of the identity credentials [52, 129]
- Content management [100, 112, 120, 121, 125]

In terms of the application domains, SSI solutions mainly focus on financial banking, education and certification, healthcare, transport, national (e-government) and IoT. For financial banking, it is to improve loan application process and verify asset ownership. It provides a new digital

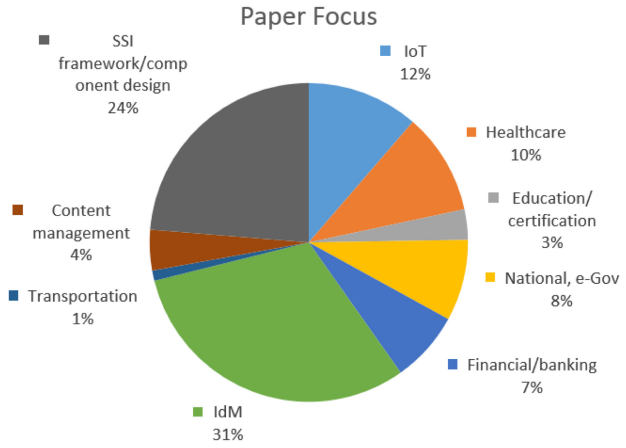

Fig. 7. Area of focus of the papers being surveyed.

approach to the KYC process and the prevention of identity scams, and the use of SSI solution in the rural areas for banking transactions. For education and certification, it is to certify the quality, authenticity of published datasets and sharing the academic credentials. In healthcare, SSI solution is to address healthcare data issues for data owners to control and delegate access of personal **EHR (electronic health record)** to relevant stakeholders, e.g., a medical practitioner. In transport, SSI approach to travel credentials for taking transportation and eventual improvement to the travel process is proposed. In national and e-Government, a verifiable national identity to support identification and authentication with cross-country interoperability is being researched. In IoT, a decentralized, transparent digital identity for remote devices for purposes of authentication, verification, and authorization is being investigated.

From the viewpoint of the SSI framework and the design of its components, the emphasis is on the performance improvement, extension, analysis, and evaluation of new SSI solutions over the existing ones for functions related to user model and authentication, verifiable credential, cryptographic schemes, key management, and digital wallet. For content management, it is to resolve issues related to digital rights management, data sharing, exchange and trading. For identity management, the focus is on resolving issues of digital identity management, authentication, access control and biometrics. Figure 7 shows the paper focus by percentage.

4.4.5 RQ 3.1: What are the Building Blocks and Components which Researchers Focus on?. There are several main building blocks and components in the SSI solution architecture. They are: Verifiable Credential/Claim (VC), Decentralized Identifier (DID), Decentralized Verifiable Data Registries (DVDR), Privacy-Promoting Credential and Claim Checks (PP), Personal Data Stores (PDS), DID Communication (DIDComm), Governance Frameworks (GF), and decentralized Distributed Ledger Technologies (DLT) or Blockchain. Blockchain is often the underlying platform used by most of the proposed solutions in the papers being surveyed here. Other papers focus on the evaluation of various blockchain platforms or DLT. There are also proposed solutions to improve the SSI building blocks and components. Figure 8 shows the coverage of the papers.

For verifiable credential and claim (VC), research looks into the backup and restore/recovery of user's verifiable credential, identity data and cryptographic keys, and supports flexible proof mechanisms to ensure that the credentials are cryptographically reliable as established by the issuer [64, 75, 79, 119, 112, 125, 132]. For decentralized identifier (DID), attribute certificate, cryptographic key creation and management, and DID search are the main focuses to provide means

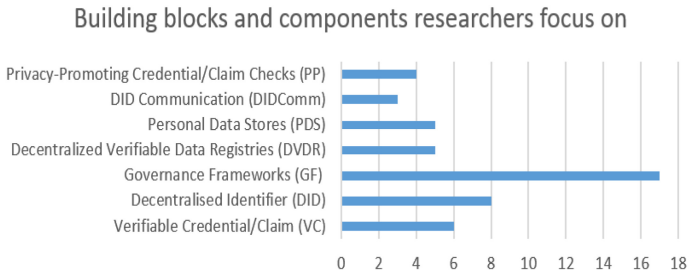

Fig. 8. Building blocks and components researchers focus on.

for both issuer and verifier to establish the identity of the holder without reliance on a centralized party [49, 51, 58, 67, 76, 88, 105, 116]. For governance framework (GF), research efforts are on new ways to improve GF and its protocols [44, 45, 54, 56, 62, 69, 81, 84, 90, 93, 101, 104, 106, 107, 111, 117]. Authentication, authorization protocol, identity/credential proofing and verification mechanism with creation of structures, roles, and policies are some of the main investigation focuses for SSI approaches to be adaptable in different domains. For DVDR, attribute access control and the suspension or revocation of verifiable credentials are being studied [46, 71, 72, 80, 85, 122, 129]. For PDS, the main focus is on the VC and DID store in the digital wallet and passports to support individual control over sharing and access to these data [52, 94, 118, 121, 124, 137]. DIDComm study is on the communication channel to facilitate DID sharing and verification regarding the acquisition, processing, and distribution of personal information [78, 126, 128, 131]. Finally, PP investigates privacy preservation techniques, specifically zero-knowledge proof schemes, and inspects the status of a credential without revealing any additional co-relatable data about an individual [63, 79, 82, 99, 122, 123].

4.4.6 RQ 3.2: What are the Current Standards, Frameworks, Regulations and Policies?. There are various standards, frameworks, regulations, and policies that support SSI architecture and solutions as well as conformance to the government legal framework. A consolidation of the SSI-related standards, frameworks, regulations and policies, as were mentioned and discussed in the surveyed papers here, are tabulated in Tables 3, 4, and 5, respectively.

4.4.7 RQ 4.1: What are the Current Implementations or Proof of Concepts?. In the categorization of the papers in RQ 3.1, some papers provided either their implementations or proof-of-concepts. According to their focuses on the different SSI building blocks and components, a tabulation of these papers with their corresponding focus on the SSI building blocks and components is described in Figure 9, and a summary of their objectives are given below.

In the management of DID and user attributes, Soltani et al. [52] implemented a decentralized service for SSI management. It defines user-managed attributes using a decentralized query protocol without a centralized party, and with privacy-preserving features such as ABE-based access control on sensitive attribute data. By allowing patients to have control over their personal health data, it also addresses the issues of data ownership and isolation, lack of accountability, and high privacy risks existing in current electronic health record (EHR) systems. Brunner et al. [58] implemented a decentralized platform that supports the issuing, management, and verification of digital documents using the Ethereum blockchain. It supports attribute-based authentication and facilitates the verification of the origin and integrity of paper documents as well as the recovery of lost documents. To address the need for trusted third party **KGC (Key Generator Center)** to generate and manage the keys for the user, Tu et al. [76] proposed a decentralized identity authentication and key management scheme using blockchain technology. To perform

Table 3. Standards Referenced by Surveyed Papers

Standards	Description	Origin
Biometrics Open Protocol Standard (BOPS)	A standard for a software-based system on identity assertion.	IEEE
Fast Identity Online (FIDO)	An open authentication standard that enables a service provider to leverage existing. technologies for passwordless authentication.	Alliance
Federated Identity Management (FIdM)	A federated SSO to establish a trusted relationship between separate organizations and third parties to share identities and authenticate users across domains.	
Identity-Mixer	A cryptographic protocol suite for privacy-preserving authentication and transfer of certified attributes.	IBM
Open Authentication (OAuth)	An open standard for access delegation used by Internet users to grant websites or applications access to their information on other websites without giving them the passwords.	IETF
OpenID Connect (OIDC)	An open authentication protocol that profiles and extends OAuth 2.0 to add an identity layer.	Foundation
Security Assertion Markup Language (SAML)	An open standard that allows identity providers (IdP) to pass authorization credentials to service providers (SP).	OASIS
U-Prove	A SDK for user-centric identity management which enables application developers to reconcile security and privacy objectives (including anonymity), and allows for digital identity claims to be efficiently tied to the use of tamper-resistant devices.	Microsoft
W3C Verifiable Credentials	Specifications for an open standard for digital credentials to support expressing and exchanging credentials that have been verified by a third party easier and more secure on the Web.	W3C

password-less authentication, Szalachowski [116] implemented and evaluated a **password-authenticated decentralized identity (PDID)** framework with global and human-meaningful names. It uses the OPAQUE protocol, which is a secure asymmetric password authenticated key exchange, and provides the performance metric of PDID operations.

For verifiable credential and claim, Abraham et al. [79] implemented a decentralized eID derivation system that enables users to selectively disclose only relevant parts of the imported identity assertion according to the services' requirements. This addresses the lack of qualified identity data that satisfies the services' requirements; it also protects users' privacy during the derivation of eID data. Jakubeit et al. [64] combined SSI sustaining aspects and extended them to create a backup-and-restore protocol with authenticated backup and auditing by remote entities.

In the area of decentralized verifiable data registries, Liang et al. [46] implemented a decentralized blockchain that made use of the trusted execution platform enabled by Intel SGX. It addresses the major privacy concern over health data collected from wearable devices, which can reflect patients' health conditions and habits, and the increased data disclosure risks among the healthcare providers and application vendors. It also provided accountability for data access. Shetty et al. [72] implemented a mobile healthcare system for personal health data collection and for data sharing and collaboration among individuals, healthcare providers, and insurance companies. Its implementation adopted blockchain and Intel SGX technology. Belchior et al. [85] developed an SSI based access control model for cross-organization identity management. They showed how SSI can fit within the context of established enterprise identity and access management technologies to alleviate data breach and user privacy problems.

Table 4. SSI Related Frameworks

Frameworks (SSI)	Description	Origin
European Self Sovereign identity framework (ESSIF)	Part of the European blockchain service infrastructure (EBSI) to provide an interoperable and decentralized framework for people to control their own identities across EU member states, without having to rely on individual government.	EU
Blockcert	An open standard for creating, issuing, viewing, and verifying blockchain-based digital records registered on the blockchain.	
Civic	A blockchain-based identity management solution that gives individuals and businesses the tools they need to control and protect personal identity information.	Commercial
Hyperledger Indy, Aries, Ursa	Hyperledger sovereign identity blockchain solutions. Indy is a distributed ledger, purposely built for decentralized identity. Aries provides an infrastructure for peer-to-peer network and blockchain-rooted interactions within platforms to offer key management and secret management systems. Ursa provides cryptographic support.	Open source
Jolocom	Open-source protocol for decentralized digital identity and access right management.	Open source
Shocard (PingID)	Store one's identity onto bitcoin's blockchain so that one can prove her/his identity whenever needed.	Commercial
Sovrin	An identity metasystem for SSI with a public service utility enabling SSI on the internet.	Non-profit organization
uPort	An SSI and user-centric data platform on Ethereum.	Donated to the Decentralized Identity Foundation
Veres One	A blockchain for acquiring and managing decentralized identifiers.	Global network for identity

For secure decentralized identifier communications and exchange of attributes, Fedrecheski et al. [126] presented a solution for self-sovereign identification and communication of IoT agents under constrained networks using the SWARM framework. It implemented an optimization layer for the protection of messages exchanged between self-sovereign agents; it also used binary encoding to achieve size reduction for signed and encrypted messages.

In the area of personal data store, the challenges are to address the closed, proprietary data silos of existing web-based social communication platforms which not only lock in users into their service platforms, but also control and exchange their personal and sensitive data such as photos, messages, or contact information, Westerkamp et al. [78] implemented a decentralized service architecture, using Ethereum blockchain, for users to take full control of their own personal data. To address the backup and restore issue of user's identity data. Grabatin and Hommel [128] developed a hardware solution based on ESP32 using SX1276/SX1278 LoRa chips, with adaptation made to the lmic- and MbedTLS-based software stack. This is to address the challenge of verifying the identity of nodes within a wireless ad-hoc mesh network and the authenticity of their messages in sufficiently secure, yet power-efficient ways. Another set of challenges is related to the limited ATM facilities in rural areas, the high initial cost of ATM deployment, the potential security issues in ATM systems, and high inter-bank transaction fees. Bandara et al. [124] implemented a

Table 5. Regulations and Policies

Regulation and policy	Description	Origin
California Consumer Privacy Act (CCPA)	To protect the data privacy rights of citizens living in California.	US
Convention on the Rights of the Child 41(CRC, non-US)	The most universally accepted human rights instrument, ratified by almost every country in the world (except two).	UN
electronic identification and trust services (eIDAS, EU)	An EU regulation on electronic identification and trust services for electronic transactions in the European Single Market.	EU
European Convention on Human Rights (ECHR), Article 8	To protect the human right of people in countries that belong to the Council of Europe.	EU
General Data Protection Regulation (GDPR, EU)	A regulation in the EU law on data protection and privacy in the European Union (EU) and the European Economic Area (EEA).	EU
Health Insurance Portability and Accountability Act of 1996 (HIPAA)	A federal law that requires the creation of national standards to protect sensitive patient health information from being disclosed without the patient's consent or knowledge.	US
International Covenant on Civil and Political Rights (ICCPR, US)	Commit its parties to respect the civil and political rights of individuals, including the right to life, freedom of religion, freedom of speech, freedom of assembly, electoral rights, and rights to due process and a fair trial.	UN
Payment Services Directive Two (PSD2)	To make online payment safer for customers, improve protection of consumer information, address payment fraud, and provide a common platform for competitors.	EU

blockchain-based, low cost, peer-to-peer money transfer system as an alternative for traditional ATM system and debit/credit card system. It also provided an SSI empowered mobile wallet for its end users. Wu et al. [121] implemented a self-sovereign blockchain-based privacy-preserving matchmaking platform that enables its users to treat their personal data as a digital asset and to trade it according to the matching score with other users.

In governance frameworks, Hammudoglu et al. [45] explored the capability of current smartphones to acquire, process, and match fingerprints using only its built-in hardware. They devised a mobile biometric-based authentication system that only relies on local processing without requiring any cloud service, server, or permissioned access to fingerprint reader hardware. In addition, they proposed this as a key building block for an SSI solution that integrates permissionless blockchain for identity and key attestation. To address the privacy concerns of sharing digital credentials for secure verification of participants' identities and credentials to increase trust, and to allow data minimization mechanism to reduce the risk of oversharing the credential data, Mukta et al. [107] implemented a blockchain-based SSI platform architecture that allows secure creation, sharing and verification of credentials.

With respect to the privacy and security requirements for users' digital identity information without a centralized party managing the identity exchange transactions, Gunasinghe et al. [63] presented a decentralized protocol for privacy-preserving exchange of users' identity information and digital assets using a permissioned blockchain network. In the light of the COVID-19 pandemic and the need for automated and efficient human contact tracing that needs to be non-intrusive and effective as well as with privacy-preserving, Bandara et al. [82] developed "Connect", a Blockchain and SSI based digital contact tracing platform. Xiao et al. [122] implemented a blockchain based SSI management solution which used threshold CP-ABE to achieve access control. To resolve issues of user attribute revocation existing in multi- **attribute authority (AA)** and threshold multi-AA schemes, the solution is deployed in large-scale cloud or cross-cloud access. Yang and Li [123]

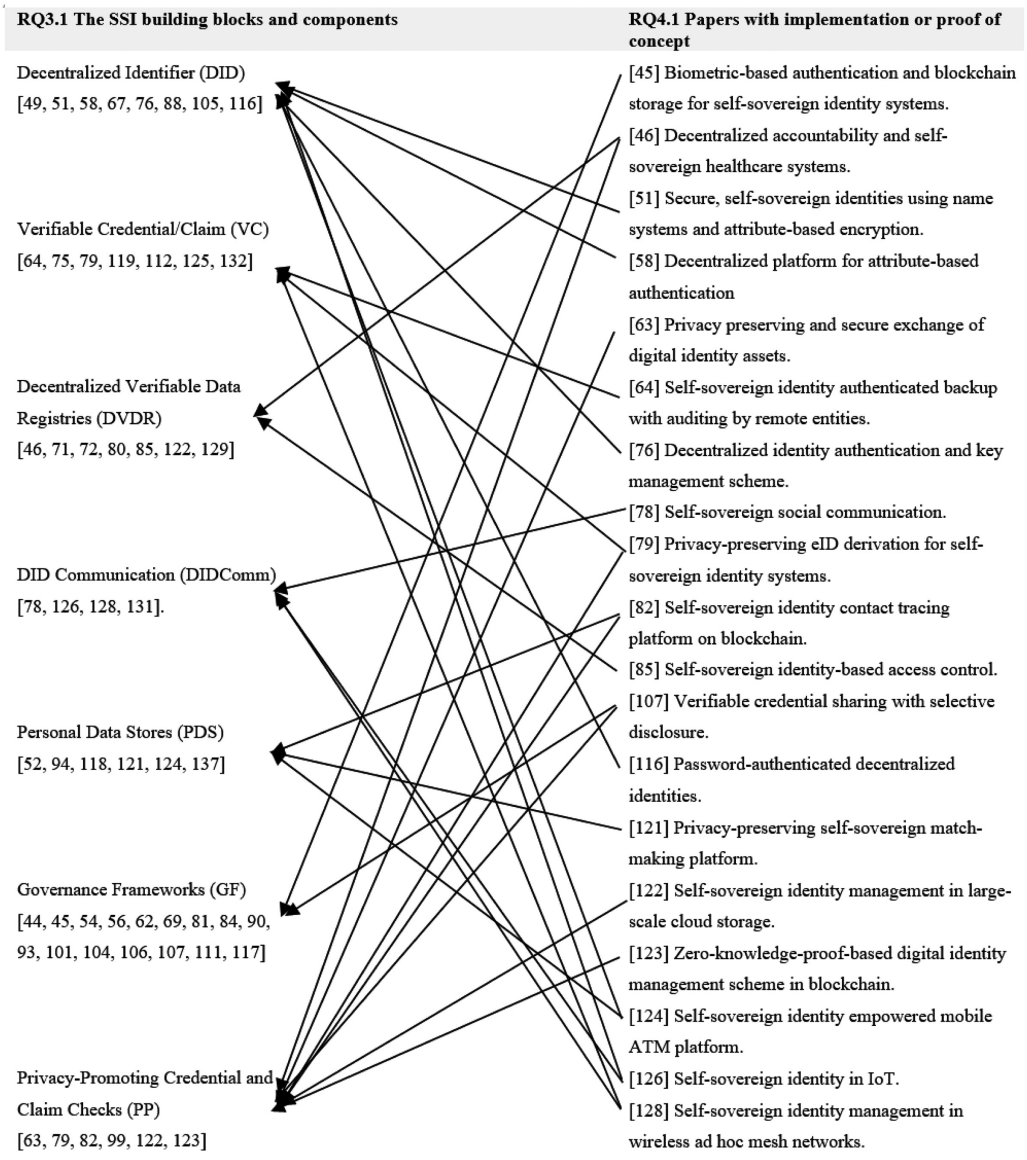

Fig. 9. The SSI building blocks and components papers with implementation or proof of concept.

leveraged the smart contracts and ZKP algorithms and implemented a system prototype named BZDIMS that includes a challenge-response protocol and allows users to selectively disclose the ownership of their attributes to service providers, thus protecting users' behavior privacy.

4.4.8 RQ 4.2: What are the Implementation and Deployment Challenges Faced?. There are open questions and challenges identified for the implementation and deployment of SSI models and solutions. They can be grouped into the following areas:

(a) Key management. Challenges in decentralized Identifiers (DIDs) are often related to the distribution of public keys, keeping track of the changes made to the keys, storage

capacity and scalability [105]. DIDs and the accompanying **DID Documents (DDoc)** enable individuals to share abstract identifiers (DIDs) with an associated key pair and a resolution end point. Depending on the individual's need to have DID to verify one's identity credential or claim, each individual can have multiple DIDs with corresponding DDocs and associated key pairs. The identity could use one key pair for authentication, another one for encrypted messages, and yet another one for its verifiable credentials (VC). The keys can also be defined to permit which keys are allowed to authenticate changes to the DID Document itself. VC issuers and holders may also revoke the VCs if they are no longer valid, and this will involve the revocation of the associated keys. Thus, maintaining and managing multiple key pairs in SSI is a challenge with regards to the distribution, update, storage, recovery and revocation of the keys that are tied closely to the DID, DDoc and VC [122]. From the IoT deployment perspective, the availability of efficient and less resource intensive open software libraries to facilitate the key management and operation that manipulates DIDs and VCs on the IoT devices [76, 88] is also a challenge.

(b) Security of personal datastore/wallet. DIDs, DDoc and VCs are stored in the individual's personal data store or wallet. Thus, the implementation needs to ensure the security of the personal data store and the ability for users to securely manage their own identities. The solution will also require efficient and secure backup, restore and recovery of the user's identity data and cryptographic public keys [49, 64, 75] as well.

(c) Scalability and reliability of solutions. DLT is used as an immutable, transparent, and shared storage whereby transaction records and data are accessible by parties within the ecosystem. Scalability and reliability are intrinsic challenges of DLTs themselves [48]. Blockchain is the most common DLT for the implementation and deployment of SSI solutions. Note that other than cryptocurrencies, blockchain as DLT is still attractive as a disruptive technology for real-world applications.

(d) Trust and assurance. For wide deployment of SSI solutions for identity management with eID, the building-up of a chain-of-trust connecting various eID systems via the SSI ecosystem with different stakeholders [79] is necessary. This requires buy-in and adoption from stakeholders to participate and build up the trust to evolve the ecosystem [93]. From the users' perspective, data link to their identity, PII may be stored on DLT, and private information may be shared. Thus, in order for users to participate in the SSI ecosystem, there must be means for users as individuals to verify that the privacy of their identities is protected in terms of ensuring confidentiality, anonymity, and unlinkability [63].

(e) Standardization. The fragmented nature of the SSI market, the immaturity and incompatibility of standards, the legal and regulatory uncertainty, . . . , and so on, all have impacts to the SSI ecosystems and their possible adoption [131]. In terms of IdM, its design needs to be based on open standards and established protocols to ensure maximum transparency and adoption [92]. The assurance of interoperability of the solution in real world applications and backward compatibility are also much needed [113] to ensure the sustainability of the solution.

4.5 Discussions

The phase SSI, contains three words, "self", "sovereign", and "identity", each of which carries certain special meaning:

- Self: an individual, own, personal, user centric.
- Sovereign: independent, self-governing, control, authority, autonomy, empowerment, without asking for permission.

- Identity: a distinguishing character, personality, attributes and behaviors by which a person or thing is recognized.

Using these three words as the guidelines for the understanding, SSI can be viewed from both technical SSI and conceptual digital identity perspectives.

(a) From the Technical SSI Perspective

From a technical point of view, an SSI solution is to build a decentralized user-centric ecosystem that grants an individual user or entity with the rights to control, authorize and consent on the disclosure and usage of its own digital identity or credentials in order to fulfill a transaction. The digital identity which is a user's personally identifiable information can be in the static data form like credential (example, user identity, password and social security number) and personal attributes (trait like height, weight, hair color, biometric characteristic) or dynamic data like individual behavior (in the form of gait like walking posture, strength of keystrokes, speaking tones).

To an individual, data security and privacy are of utmost importance. Most of the papers that we surveyed focus on addressing problems related to the protection of an individual identity against theft and exploitation. The other shared problem of interest is the proliferation of one's password against multiple online platforms that are a common target for malicious entities. The SSI concept of decentralizing the storage for these credentials in the form of verifiable credentials, issued by authorized and trust issuers, and stored within one's digital wallet enables the individual entities to control the usage of their own data in a decentralized manner. There are varying properties and principles defined for SSI that attempt to serve as a bounding guide for the proposed SSI solutions. Controllability and security, as well as sustainability, are prominent as the principles proposed by researchers. In terms of wide adoption, portability and interoperability are also desired properties which are also current challenges faced in order to have a wider deployment.

World Wide Web Consortium (W3C) [141] and **Decentralized Identity Foundation (DIF)** [142] have been striving to establish standardization for new decentralized identity ecosystems. However, SSI is still in its infancy stage. Current proposed standards are with ambiguity, misunderstanding, and disagreement on key concepts and definitions with respect to the proposed workable and adaptable frameworks. The standards are created to establish interoperability and portability of the identity and credentials and to facilitate the exchange and management of identities and credentials. The deployment of SSI solutions and architectures requires the use of DLT as an essential building block, and blockchain is the most preferred DLT platform for the development and deployment of SSI solutions. To facilitate the creation of SSI ecosystems, frameworks are built to support the development of SSI solutions. Among the currently available frameworks, the most commonly used one is HyperLedger Indy (with the close support of its two sibling frameworks, Aries and Ursa) from the Blockchain Foundation established by the Linux Foundation and its members as well as the HyperLedger Fabric community. Sovrin and Uport are the other two SSI blockchain platforms commonly used for the deployment of decentralized identities and verifiable credentials.

SSI is still an emerging trend. Attention of researchers toward proposed SSI solutions also focuses on the design of governance framework to establish a secure, reliable protocol and mechanism. This is done through creating structures, roles, and policies for organizations or governments to facilitate the adoption and adaptation of SSI approaches in different domains. DID is also a focus as it is associated to document that holds critical contents to establish the authenticity and verifiability of claims and identity, while at the same time protecting the privacy of the holder through selective disclosure on a need-to-know basis. This is of utmost concern to an individual in protecting her/his identifiable information but at the same time able to provide the needed authentication

and verification. The suspension, revocation and recovery of VC and DID, both in individual's personal storage and in issuer's data registries are also of major focuses for researchers.

There are proposals and solutions to resolve the various challenges and gaps in the current SSI developments and deployments. However, as was discussed in the elaboration of RQ 4.2 above, there are still open problems. From the computer science perspective, open problems related to the portability and interoperability of the solutions, minimalization, protection and provability of verifiable claims, and better distributed key management system still remain to be solved. And from a social and national perspective, legal supports (including regulation and policy) as well as institutional, organizational and user adoptions are key challenges to the SSI deployment.

(b) From the Conceptual Digital Identity Perspective

There are substantial research and engineering efforts on the advancement of SSI, in terms of both the architectural design and the functionalities of the components. While they are important in the realization of SSI on system platforms, there is one fundamental question that often seems to be overlooked. It is the question of "what is identity", which in turn dictates (or at least influences) the design of SSI. In the current SSI implementation, the attributes to be managed are often the static identity properties of a person such as citizen identity card number. However, attributes in a digital identity, in theory, can be anything, from static attributes to dynamic attributes (e.g., behavior tracing). Even more complicated, the data owner and the data holder of the attributes might not be different persons. This will definitely complicate the architecture design of SSI. Another important consideration is the requirements that are spelled out by the FAIR and CARE principles. It will be interesting to see how the current SSI architecture design can be extended to support requirements from the FAIR and CARE guidelines and principles.

5 RELATED WORK

With respect to the related work, there are at least three perspectives. They are data sovereignty, digital sovereignty, and self-sovereign identity. While data sovereignty and digital sovereignty define the requirements for digital identity and the attributes that should be included inside, self-sovereign identity focuses more on the standards, requirements (from the viewpoint such as communication and interoperability), architecture and its implementation. In terms of systematic reviews on these three perspectives, many more papers are found for self-sovereign identity and very few for data and digital sovereignty.

For data sovereignty, two papers (from computer science related sources) are found. The first one is by Asswad and Gómez [153]. The paper gives an overview of the different aspects of data ownership, ranging from the personal ethical dimension to organizational and beyond organizational contexts. It describes new challenges of data ownership from IoT, where the amount of data is huge, and the number of stakeholders interested is tremendous. Furthermore, laws and regulations related to data ownership are also discussed. Finally, it covers the state of the art on the conceptualization and the implementation of data ownership concepts in five major domains: health, transportation, industry, energy, and smart cities. The second paper is by Hummei et al. [154]. This paper tries to sort out the difference among data sovereignty, digital sovereignty, and cyber sovereignty. It reports that there is a considerable degree of divergence and an occasional lack of clarity about intended meanings of data sovereignty and digital sovereignty. Then it proposes a conceptual grid to systematize different dimensions and connotations of data sovereignty with respect to meaningful control, ownership, and other claims to data articulated by a variety of agents ranging from individuals to countries. Both papers give a more in-depth understanding of what digital identity should contain, and this should be translated to the requirements for the implementation of self-sovereign identity to support.

For digital sovereignty survey, apparently, no literature is found. Rather, literatures related to self-sovereign identity often come up. Perhaps this is due to the fact that from the computer science perspective, the emphasis is much more on the architecture and system implementation rather than diving into the fundamental question of what is "identity".

For self-sovereign identity, there are quite a number of systematic reviews or surveys found. Siqueira et al. [136] provided a systematic literature review focusing on healthcare to investigate state-of-the-art measures based on SSI and blockchain technologies for dealing with electronic health records (EHRs), identifying gaps, and determining the key questions for future research. It looked at questions in the area of the access control mechanisms for EHR access and sharing, the privacy and security risks regarding unauthorized medical information disclosure and its storage. It concluded that it is still a novel subject that medical institutions have been exploring for some time, but still face the challenge of breaking the silos of health information and privacy concerns. Blockchain technologies have provided a viable means of addressing the fundamental challenges of EHR solutions such as access control, data integrity, interoperability, and auditing. The principles of SSI could be adopted to provide patient-centric healthcare solutions to ensure that patients, as the data owner, have complete control over their data. Houtan et al. also gave similar blockchain-based self-sovereign patient identity in healthcare [152]. However, the paper's coverage is mainly on healthcare and SSI adoption in this domain and does not cover other domains like this paper.

Schardong and Custódio et al. [143] presented a systematic mapping and systematic literature review covering theoretical and practical advances in SSI. As part of their research questions on SSI, it identified the authors of the papers, the co-references and co-authorship in this SSI ecosystem, the conceptual ideas introduced and refuted, a formal definition of SSI and the practical problems that have been introduced and solved. It also attempted to introduce mathematical formulations to precisely define one or more SSI-related problems and presented a solution to the pragmatical problem related to the SSI ecosystem. Mühle et al. [150] surveyed on the essential components of a self-sovereign identity. They focused on four basic components of SSI, namely identification, authentication, verifiable claims and storage. They also discussed two major operation approaches for SSI: the identifier registry model and its extension the claim registry model. Soltani et al. [148] gave related survey of self-sovereign identity ecosystem. On top of the technical implementation and regulations issues, they also listed down eight challenges of SSI, namely, standards for data management and wallets, key management, consent, access, accountability and governance, trust in data, new technology adoption, and investment and commercialization.

Čučko and Turkanović [144] presented a systematic mapping methodology to provide a coarse-grained overview of decentralized and SSI and structure the research area by identifying, analyzing, and classifying the research papers according to their contribution, application domain, IT field, research type, research method, and place of publication. In contrast to these two review papers, this paper provides a wider study to include data and digital sovereignty to lay a background for SSI and also to establish an overview understanding of sovereignty as a whole on digitalized data. Zaeem et al. [149] compiled a comprehensive list of functional and non-functional requirements of SSI and compare an extensive number of existing SSI/blockchain-based identity management solutions with respect to these requirements. Kaneriya and Patel [151] presented an exhaustive study on different blockchain based self-sovereign identity implementations (such as Sovrin, Uport, EverID, LifeID, Sora, SelfKey) along with its architectural components and discussed about use case of self-sovereign identity in details.

All these survey works about self-sovereign identity are good. However, they are often limited to the architecture and implementation aspects, without linking back to the fundamental question of the intrinsic properties and requirements of digital identity as well as their implications to SSI implementation.

6 CONCLUSION

This paper conducted a systematic study of data and digital sovereignty as well as SSI and presented the findings, observations with analysis and discussions.

Data sovereignty is not a new research topic as research had been performed by humanity and sociological researchers to look into the concerns and wishes of the indigenous peoples to protect their land, cultural heritage and assets. The CARE and FAIR principles are defined to advance the legal principles underlying the collective and individual data rights in the context of the UNDRIP [23] as well as to support data and knowledge integration and promote sharing and re-use of data. With the movement of digitization to turn analogue data to digital, nation-states start to be aware of the need to have sovereignty over their citizens' data and to protect their well-being and privacy. Regulations, like the European Union's GDPR, policies and laws are created to protect the data, specifically how data are used, where they are stored and their physical location. At its core, data sovereignty is about protecting sensitive, private data and ensuring it remains under the control of its owner within the specified country. However, with the lack of operational details on implementation mechanisms and enacting systems to execute and realize the data sovereignty principles and law, it is a challenge to govern the stewardship and application of data to fully assert the sovereignty of the locally hosted data of the people they are designed to protect. In our study, we found that while these FAIR and CARE principles are quite mature in terms of the requirements specification, they are seldom connected to the architecture design and implementation work of self-sovereign identity. We argue that the gap between FAIR/CARE and SSI should be bridged so that the original goals of data sovereignty can be achieved.

Digital sovereignty is a more recent topic emerging from the acceleration of digital transformation. With the dominant position of big tech companies in the field of cloud computing and social media, data of citizens and companies are virtually stored and utilised in the cloud of these big tech companies. To wrestle back the control and reassert their authority, nations like the EU initiated the GAIA-X cloud project and the European Cloud Federation initiative to create its own European offering of cloud infrastructure where the customer has full control over the storage, access, and processing of the data. However, the project requires a high degree of commitment and coordination from the EU member states which can be a challenge. In addition, with the lack of binding European cloud policies, currently the cloud choices made by the EU member states are through outsourcing. This brings out the risk of vendor lock-in that threatens digital sovereignty. On a personal basis, digital sovereignty is also the wish of an individual, as individual users of digital technologies and services, to have the autonomy to determine who gets to access and use the data about them as individuals. To an individual, the protection of their online identity is of particular concern and an individual's wish to be able to exercise sovereignty over one's identity. Cases of identity theft and the ability of central authority platforms to monitor the individual's online transactions have a direct impact on an individual's privacy and digital sovereignty. The evolution of the web, from Web 1.0 (read only), Web 2.0 (read/write) to Web 3.0 (read/write/own) clearly shows the trend and movement of digital sovereignty.

With respect to the implementation of digital sovereignty, the research topic on SSI has gained popularity and interest among computer science academia. The concept of SSI is to empower an individual to take ownership and control of her/his own personal and distinguishable data and be able to independently authorize the disclosure of her/his personal data. From the technical viewpoint, an SSI solution is a user-centric, decentralized ecosystem that empowers an individual user with the ability to control its own digital identity to authorize and consent on its disclosure and usage to fulfill a transaction. Data security and privacy are of utmost importance and focus. There are varying properties and principles defined for SSI; they serve as a bounding guide for proposed SSI solutions. For wide adoption, **World Wide Web Consortium (W3C)** [141] and

Decentralized Identity Foundation (DIF) [142] have defined standards for a new decentralized identity ecosystem. SSI is still in its infancy of establishing standards and frameworks. While substantial research efforts focus on improving the SSI frameworks and components design and the SSI core – identity management and governance, there are open problems in regard to the portability and interoperability of the solutions, minimalization, protection and provability of verifiable claims and a better distributed key management system. From the social and national perspectives, legal supports (including regulation and policy) as well as institutional, organizational and user adoptions are key challenges to SSI deployment.

REFERENCES

[1] A. F. Westin. 1968. Privacy and freedom. *Washington and Lee Law Review* 25, 1 (1968), 166.

[2] Google Inc. 2021. Helping build the digital future. On Europe's terms. Nov. 2021. [Online]. Available: https://cloud.google.com/blog/products/identity-security/helping-build-the-digital-future-on-europes-terms/

[3] K. Y. Lam and C. H. Chi. 2016. Identity in the Internet-of-Things (IoT): New challenges and opportunities. In *Proceedings of International Conference on Information and Communications Security*, Springer. 18–26.

[4] H. W. Sun, K. Y. Lam, M. Gu, and J. G. Sun. 2006. An efficient algorithm for fingercode-based biometric identification. In *Proceedings of the 1st International Workshop on Information Security (IS'06)*, Springer-Verlag LNCS 4277, 469—478.

[5] F. Gondesen, S. Mitra, and K. Y. Lam. 2020. Feasibility of PUF based authentication on ATtiny devices with off-the-shelf SRAM. In *Proceedings of the 6th ACM Cyber-Physical System Security Workshop (CPSS 2020)*, Taipei, Taiwan, 2020.

[6] B. Srinivasu, P. Vikramkumar, A. Chattopadhyay, and K. Y. Lam. 2018. CoLPUF: A novel configurable LFSR-based PUF. In *Proceedings of the 14th Asia Pacific Conf. on Circuits and Systems (APCCAS 2018)*, Chengdu, China.

[7] J. Y. Shi and K. Y. Lam. 2009. MinuCode: A fixed-value representation of fingerprint minutiae for biometric cryptosystem. In *Proceedings of the 3rd International Conference on Information Security and Assurance (ISA'2009)*, LNCS, Springer.

[8] X. B. Zhao, K. Y. Lam, S. L. Chung, M. Gu, and J. G. Sun. 2004. Authorization mechanisms for virtual organizations in distributed computing systems. In *Proceedings of the 9th Australasian Conference on Information Security and Privacy (ACISP'04)*, Springer-Verlag LNCS 3108, Sydney, Australia. 414—426.

[9] J. P. Yong, K. Y. Lam, S. L. Chung, M. Gu, and J. G. Sun. 2004. Enhancing the scalability of the community authorization service for virtual organizations. In *Proceedings of the 1st Advanced Workshop on Content Computing (AWCC 2004)*, Spring-Verlag LNCS 3309, 182—193.

[10] M. Ge and K. Y. Lam. 2009. Self-initialized distributed certificate authority for mobile ad hoc network. In *Proceedings of the 3rd International Conference on Information Security and Assurance (ISA'09)*, Seoul, Korea, LNCS 5576, Springer.

[11] J. L. Guo, W. Z. Yang, K. Y. Lam, and X. Yi. 2018. Using blockchain to control access to cloud data. In *Proceedings of the 14th International Conference on Information Security and Cryptology (Inscrypt 2018)*, Fuzhou, China, Springer LNCS. 11449.

[12] H. Y. Ma, E. Huang, and K. Y. Lam. 2020. Blockchain-based mechanism for fine-grained authorization in data crowdsourcing. *Future Generation Computer Systems, Elsevier*, 106 (2020), 121–134.

[13] K. Y. Lam, S. Mitra, F. Gondesen, and X. Yi. 2021. ANT-Centric IoT security reference architecture - Security-by-Design for satellite-enabled smart cities. *IEEE Internet of Things Journal* 9, 8 (2021), 5895–5908.

[14] R. Agrawal, and R. Srikant. 2000. Privacy-preserving data mining. In *Proceedings of the ACM SIGMOD International Conference on Management of Data*. 439–450.

[15] C. Dwork. 2008. Differential privacy: A survey of results. In *Proceedings of International Conference on Theory and Applications of Models of Computation*. Springer, Berlin. 1–19.

[16] Z. Brakerski and V. Vaikuntanathan. 2014. Efficient fully homomorphic encryption from (standard) LWE. *SIAM Journal on Computing* 43, 2 (2014). 831–871.

[17] Wikipedia. 2021. Data sovereignty. Nov. 2021. [Online]. Available: https://en.wikipedia.org/wiki/Data_sovereignty

[18] B. Q. S. M. Smith. 2017. Indigenous data sovereignty: Toward an agenda: Official Newsletter of the New Zealand demographic society. *New Zealand Population Review* 43 (2017), 159–161.

[19] S. Couture and S. Toupin. 2019. What does the notion of "sovereignty" mean when referring to the digital?. *New Media & Society* 21, 10 (2019), 2305–2322.

[20] M. Walter, R. Lovett, B. Maher, B. Williamson, J. Prehn, G. Bodkin-Andrews, and V. Lee. 2021. Indigenous data sovereignty in the era of big data and open data. *Australian Journal of Social Issues* 56, 2 (2021), 143–156.

[21] T. Kukutai and J. Taylor. 2016. *Indigenous Data Sovereignty: Toward an Agenda*. ANU press, 2016.

[22] The World Bank. 2021. World Development Report 2021. 2021. [Online]. Available: https://wdr2021.worldbank.org/stories/crossing-borders/
[23] The Indigenous Peoples and Development Branch/Secretariat of the Permanent Forum on Indigenous Issues. Issues, "United Nations Declaration on the Rights of Indigenous Peoples."2021. [Online]. Available: https://www.un.org/development/desa/indigenouspeoples/declaration-on-the-rights-of-indigenous-peoples.html
[24] T. G. I. D. ALLIANCE. 2021. CARE principles for indigenous data governance. 16 Oct 2021, 2021; [Online]. Available: https://www.gida-global.org/care
[25] S. R. Carroll, E. Herczog, M. Hudson, K. Russell, and S. Stall. 2021. Operationalizing the CARE and FAIR principles for indigenous data futures. *Scientific Data* 8, 1 (2021), 1–6.
[26] S. C. Rainie, T. Kukutai, M. Walter, O. L. Figueroa-Rodríguez, J. Walker, and P. Axelsson. 2019. Indigenous data sovereignty. In *The State of Open Data. Histories and horizons: African Minds and the International Development Research Centre (IDRC*. 300–319.
[27] E. Digital. 2020. EIT Digital Report on European Digital Infrastructure and Data Sovereignty. 2020. [Online]. Available: https://www.earto.eu/eit-digital-report-on-european-digital-infrastructure-and-data-sovereignty/
[28] C. Stix. 2021. The ghost of AI governance past, present and future: AI governance in the European Union. *Cornell University Library, arXiv.org*, (2021).
[29] J. Müller-Quade, M. Backes, P. Buxmann, C. Eckert, and T. Holz. 2019. Cybersecurity research: Challenges and course of action. *Tech. Rep.*, Karlsruher Institut fur Technologie (KIT), 2019.
[30] N. Möllers. 2021. Making digital territory: Cybersecurity, techno-nationalism, and the moral boundaries of the state. *Science, Technology, & Human Values* 46, 1 (2021), 112–138.
[31] S. Geisler, M.-E. Vidal, C. Cappiello, B. F. Lóscio, A. Gal, M. Jarke, M. Lenzerini, P. Missier, B. Otto, and E. Paja. 2021. Knowledge-Driven data ecosystems toward data transparency. *ACM Journal of Data and Information Quality (JDIQ)* 14, 1 (2021), 1–12.
[32] J. Pohle and T. Thiel. 2021. Digital sovereignty. *Practicing Sovereignty: Digital Involvement in Times of Crises*. Bielefeld: transcript Verlag. 47–67.
[33] L. Floridi. 2020. The fight for digital sovereignty: What it is, and why it matters, especially for the EU. *Philosophy & Technology* 33, 3 (2020), 369–378.
[34] P. Chapdelaine and R. Jaqueline McLeod. 2021. Contested sovereignties: States, media platforms, peoples, and the regulation of media content and big data in the networked society. *Laws* 10, 3 (2021), 66.
[35] A. Cattaruzza, D. Danet, S. Taillat, and A. Laudrain. 2016. Sovereignty in cyberspace: Balkanization or democratization" In *Proceedings of International Conference on Cyber Conflict (CyCon US)*. IEEE. 1–9.
[36] A. Braud, G. Fromentoux, B. Radier, and O. Le Grand. 2021. The road to european digital sovereignty with Gaia-X and IDSA. *IEEE Network* 35, 2 (2021), 4–5.
[37] L. Moerel and P. Timmers. 2021. Reflections on digital sovereignty. *EU Cyber Direct, Research in Focus Series* (2021).
[38] Fitch Solutions Group Limited. 2021. Internet of things: The industry connection - 27 September 2021. London, 2021.
[39] D. Evans, V. Kolesnikov, and M. Rosulek. 2017. A pragmatic introduction to secure multi-party computation. *Foundations and Trends in Privacy and Security* 2 (2017), 2–3.
[40] M. Blum, P. Feldman, and S. Micali. 2019. Non-interactive zero-knowledge and its applications. *Providing Sound Foundations for Cryptography: On the Work of Shafi Goldwasser and Silvio Micali* (2019), 329–349.
[41] E. Celeste and F. Fabbrini. 2020. Competing jurisdictions: Data privacy across the borders. *Data Privacy and Trust in Cloud Computing* (2020), 43–58.
[42] S. Keele. 2007. Guidelines for performing systematic literature reviews in software engineering. *Citeseer* (2007).
[43] S. Alboaie and D. Cosovan. 2017. Private data system enabling self-sovereign storage managed by executable choreographies. In *Proceedings of IFIP International Conference on Distributed Applications and Interoperable Systems*. Springer, Cham, 83–98.
[44] D. Augot, H. Chabanne, O. Clémot, and W. George. 2017. Transforming face-to-face identity proofing into anonymous digital identity using the Bitcoin blockchain. In *Proceedings of 15th Annual Conference on Privacy, Security and Trust (PST)*. IEEE. 25–2509.
[45] J. Hammudoglu, J. Sparreboom, J. Rauhamaa, J. Faber, L. Guerchi, I. P. Samiotis, S. Rao, and J. A. Pouwelse. 2017. Portable trust: Biometric-based authentication and blockchain storage for self-sovereign identity systems. *arXiv preprint* arXiv:1706.03744.
[46] X. Liang, S. Shetty, J. Zhao, D. Bowden, D. Li, and J. Liu. 2017. Towards decentralized accountability and self-sovereignty in healthcare systems. In *Proceedings of International Conference on Information and Communications Security*. Springer, Cham, 387–398.
[47] N. Naik and P. Jenkins. 2017. Securing digital identities in the cloud by selecting an apposite Federated Identity Management from SAML, OAuth and OpenID Connect. In *Proceedings of 11th International Conference on Research Challenges in Information Science (RCIS)*. IEEE. 163–174.

[48] P. Dunphy, L. Garratt, and F. Petitcolas. 2018. Decentralizing digital identity: Open challenges for distributed ledgers. In *Proceedings of IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*. IEEE, 75–78.
[49] G. Linklater, C. Smith, A. Herbert, and B. Irwin. 2018. Toward distributed key management for offline authentication. In *Proceedings of the Annual Conference of the South African Institute of Computer Scientists and Information Technologists*. 10–19.
[50] A. Mühle, A. Grüner, T. Gayvoronskaya, and C. Meinel. 2018. A survey on essential components of a self-sovereign identity. *Computer Science Review* 30 (2018), 80–86.
[51] M. Schanzenbach, G. Bramm, and J. Schütte. 2018. reclaimID: Secure, self-sovereign identities using name systems and attribute-based encryption. In *Proceedings of 17th IEEE International Conference on Trust, Security and Privacy in Computing and Communications/12th IEEE International Conference on Big Data Science and Engineering (TrustCom/BigDataSE)*. IEEE. 946–957.
[52] R. Soltani, U. T. Nguyen, and A. An. 2018. A new approach to client onboarding using self-sovereign identity and distributed ledger. In *Proceedings of 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)*. IEEE. 1129–1136.
[53] Q. Stokkink and J. Pouwelse. 2018. Deployment of a blockchain-based self-sovereign identity. In *Proceedings of 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)*. IEEE, 1336–1342.
[54] H. Anada. 2019. Decentralized multi-authority anonymous authentication for global identities with non-interactive proofs. In *Proceedings of IEEE International Conference on Smart Computing (SMARTCOMP)*. IEEE, 25–32.
[55] P. C. Bartolomeu, E. Vieira, S. M. Hosseini, and J. Ferreira. 2010. Self-sovereign identity: Use-cases, technologies, and challenges for industrial IoT. In *Proceedings of 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*. IEEE, 1173–1180.
[56] L. Bathen, G. H. Flores, G. Madl, D. Jadav, A. Arvanitis, K. Santhanam, C. Zeng, and A. Gordon. 2019. Selfis: Self-sovereign biometric IDs. In *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. 2847–2856.
[57] A. Berbar and A. Belkhir. 2019. Identification in the service of national solidarity. In *Proceedings of the 4th International Conference on Smart City Applications*. 1–6.
[58] C. Brunner, F. Knirsch, and D. Engel. 2019. SPROOF: A decentralized platform for attribute-based authentication. In *Proceedings of International Conference on Information Systems Security and Privacy*. Springer, Cham, 1–23.
[59] A. B. Chavan and K. Rajeswari. 2019. Design and development of self-sovereign identity using Ethereum blockchain. In *Proceedings of International Conference on Sustainable Communication Networks and Application*. Springer, Cham, 523–531.
[60] M. S. Ferdous, F. Chowdhury, and M. O. Alassafi. 2019. In search of self-sovereign identity leveraging blockchain technology. *IEEE Access* 7 (2019), 103059–103079.
[61] A. Grüner, A. Mühle, T. Gayvoronskaya, and C. Meinel. 2019. A comparative analysis of trust requirements in decentralized identity management. In *Proceedings of International Conference on Advanced Information Networking and Applications*. Springer, Cham, 200–213.
[62] H. Gulati and C.-T. Huang. 2019. Self-sovereign dynamic digital identities based on blockchain technology." In *Proceedings in SoutheastCon*. IEEE, 1–6.
[63] H. Gunasinghe, A. Kundu, E. Bertino, H. Krawczyk, S. Chari, K. Singh, and D. Su. 2019. PrivIdEx: Privacy preserving and secure exchange of digital identity assets. In *Proceedings in The World Wide Web Conference*, 594–604.
[64] P. Jakubeit, A. Dercksen, and A. Peter. 2019. SSI-AWARE: Self-sovereign identity authenticated backup with auditing by remote entities. *Information Security Theory and Practice* (2019). 202.
[65] A. Jha, R. K. Bhattacharjee, M. Nandi, and F. A. Barbhuiya. 2019. A framework for maintaining citizenship record on blockchain. In *Proceedings of ACM International Symposium on Blockchain and Secure Critical Infrastructure*. 29–38.
[66] N. Kulabukhova, A. Ivashchenko, I. Tipikin, and I. Minin. 2019. Self-sovereign identity for IoT devices. In *Proceedings of International Conference on Computational Science and Its Applications*. Springer, Cham, 472–484.
[67] Z. A. Lux, F. Beierle, S. Zickau, and S. Göndör. 2019. Full-text search for verifiable credential metadata on distributed ledgers. In *Proceedings of the 6th International Conference on Internet of Things: Systems, Management and Security (IOTSMS)*. IEEE, 519–528.
[68] M. M. Nielsen. 2019. Tackling identity management, service delivery, and social security challenges: Technology trends and partnership models. In *Proceedings of the 12th International Conference on Theory and Practice of Electronic Governance*. 1–5.
[69] K. Pinter, D. Schmelz, R. Lamber, S. Strobl, and T. Grechenig. 2019. Towards a Multi-party, Blockchain-Based identity verification solution to implement clear name laws for online media platforms. In *Proceedings of International Conference on Business Process Management*. Springer, Cham, 151–165

[70] S. Sahmim, H. Gharsellaoui, and S. Bouamama. 2019. Edge computing: Smart identity wallet based architecture and user centric. *Procedia Computer Science* 159 (2019), 1246–1257.

[71] A.-S. Shehu, A. Pinto, and M. E. Correia. 2019. Privacy preservation and mandate representation in identity management systems. In *Proceedings of 14th Iberian Conference on Information Systems and Technologies (CISTI)*. IEEE, 1–6.

[72] S. Shetty, X. Liang, D. Bowden, J. Zhao, and L. Zhang. 2019. Blockchain-based decentralized accountability and self-sovereignty in healthcare systems. *Business Transformation through Blockchain*. Springer, 119–149.

[73] E. Shoemaker, G. S. Kristinsdottir, T. Ahuja, D. Baslan, B. Pon, P. Currion, P. Gumisizira, and N. Dell. 2019. Identity at the margins: Examining refugee experiences with digital identity systems in Lebanon, Jordan, and Uganda. In *Proceedings of the 2nd ACM SIGCAS Conference on Computing and Sustainable Societies*, 206–217.

[74] W. L. Sim, H. N. Chua, and M. Tahir. 2019. Blockchain for identity management: The implications to personal data protection. In *Proceedings of IEEE Conference on Application, Information and Network Security (AINS)*. IEEE, 30—35.

[75] R. Soltani, U. T. Nguyen, and A. An. 2019. Practical key recovery model for self-sovereign identity based digital wallets. In *Proceedings of IEEE International Conference on Dependable, Autonomic and Secure Computing, International Conference on Pervasive Intelligence and Computing, International Conf. on Cloud and Big Data Computing, Intl. Conference on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech)*. IEEE, 320–325.

[76] Y. Tu, J. Gan, Y. Hu, R. Jin, Z. Yang, and M. Liu. 2019. Decentralized identity authentication and key management scheme. In *Proceedings of IEEE 3rd Conference on Energy Internet and Energy System Integration (EI2)*. IEEE, 2697–2702.

[77] D. van Bokkem, R. Hageman, G. Koning, L. Nguyen, and N. Zarin. 2019. Self-sovereign identity solutions: The necessity of blockchain technology. *arXiv preprint* arXiv:1904.12816, 2019.

[78] M. Westerkamp, S. Göndör, and A. Küpper. 2019. Tawki: Towards self-sovereign social communication. In *Proceedings of IEEE International Conference on Decentralized Applications and Infrastructures (DAPPCON)*. IEEE, 29–38.

[79] A. Abraham, F. Hörandner, O. Omolola, and S. Ramacher. 2019. Privacy-preserving eID derivation for self-sovereign identity systems. In *Proceedings of International Conference on Information and Communications Security*. Springer, Cham, 307–323.

[80] A. Abraham, S. More, C. Rabensteiner, and F. Hörandner. 2020. Revocable and offline-verifiable self-sovereign identities. In *Proceedings of IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*. IEEE, 1020–1027.

[81] K. A. Ahmed, S. F. Saraya, J. F. Wanis, and A. M. Ali-Eldin. 2020. A self-sovereign identity architecture based on blockchain and the utilization of customer's banking cards: The case of bank scam calls prevention. In *Proceedings of 15th International Conference on Computer Engineering and Systems (ICCES)*. IEEE, 1–8.

[82] E. Bandara, X. Liang, P. Foytik, S. Shetty, C. Hall, D. Bowden, N. Ranasinghe, K. De Zoysa, and W. K. Ng. 2021. Connect-blockchain and self-sovereign identity empowered contact tracing platform. In *Proceedings of Wireless Mobile Communication and Healthcare: 9th EAI International Conference, MobiHealth 2020, Virtual Event*, 2020, Proceedings. Springer Nature. 362 (2021), 208.

[83] I. Barclay, M. Freytsis, S. Bucher, S. Radha, A. Preece, and I. Taylor. 2020. Towards a modelling framework for self-sovereign identity systems. *arXiv preprint* arXiv:2009.04327.

[84] I. Barclay, S. Radha, A. Preece, I. Taylor, and J. Nabrzyski. 2020. Certifying provenance of scientific datasets with self-sovereign identity and verifiable credentials. *arXiv preprint* arXiv:2004.02796.

[85] R. Belchior, B. Putz, G. Pernul, M. Correia, A. Vasconcelos, and S. Guerreiro. 2020. SSIBAC: Self-Sovereign identity based access control. In *Proceedings of IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*. IEEE, 1935–1943.

[86] M. P. Bhattacharya, P. Zavarsky, and S. Butakov. 2020. Enhancing the security and privacy of self-sovereign identities on Hyperledger Indy blockchain. In *Proceedings of International Symposium on Networks, Computers and Communications (ISNCC)*. IEEE, 1–7.

[87] C. Dong, Z. Wang, S. Chen, and Y. Xiang. 2020. BBM: A blockchain-based model for open banking via self-sovereign identity. In *Proceedings of International Conference on Blockchain*. Springer, Cham, 61–75.

[88] G. Fedrecheski, J. M. Rabaey, L. C. Costa, P. C. C. Ccori, W. T. Pereira, and M. K. Zuffo. 2020. Self-sovereign identity for IoT environments: A perspective. In *Proceedings of 2020 Global Internet of Things Summit (GIoTS)*. IEEE, 1–6.

[89] B. Ford. 2020. Identity and personhood in digital democracy: Evaluating inclusion, equality, security, and privacy in pseudonym parties and other proofs of personhood. *arXiv preprint* arXiv:2011.02412.

[90] S. K. Gebresilassie, J. Rafferty, P. Morrow, L. L. Chen, M. Abu-Tair, and Z. Cui. 2020. Distributed, secure, self-sovereign identity for IoT devices. In *Proceedings of 2020 IEEE 6th World Forum on Internet of Things (WF-IoT)*. IEEE. 1–6.

[91] A. Giannopoulou. 2020. Data protection compliance challenges for self-sovereign identity. In *Proceedings of International Congress on Blockchain and Applications*. Springer, Cham, 91–100.

[92] K. Gilani, E. Bertin, J. Hatin, and N. Crespi. 2020. A survey on blockchain-based identity management and decentralized privacy for personal data. In *Proceedings of 2020 2nd Conference on Blockchain Research & Applications for Innovative Networks and Services (BRAINS)*. IEEE, 97–101.

[93] A. Grüner, A. Mühle, M. Meinig, and C. Meinel. 2020. A taxonomy of trust models for attribute assurance in identity management. In *Proceedings in Workshops of the International Conference on Advanced Information Networking and Applications*. Springer, Cham. 65–76.

[94] H. Halpin. 2020. A critique of immunity passports and W3C decentralized identifiers. *arXiv preprint* arXiv: 2012.00136.

[95] B. Houtan, A. S. Hafid, and D. Makrakis. 2020. A survey on blockchain-based self-sovereign patient identity in healthcare. *IEEE Access* 8 (2020), 90478–90494.

[96] J. Huang, M. Wu, and Y. Huang. 2020. Research and application of eID digital identity. In *Proceedings of the 2nd International Conference on Artificial Intelligence and Advanced Manufacture*. 266–270.

[97] G. Ishmaev. 2020. Sovereignty, privacy, and ethics in blockchain-based identity management systems. *Ethics and Information Technology* (2020), 1–14.

[98] J. Kaneriya and H. Patel. 2020. A comparative survey on blockchain based self-sovereign identity system. In *Proceedings of 3rd International Conference on Intelligent Sustainable Systems (ICISS)*. IEEE, 1150–1155.

[99] N. Kulabukhova. 2020. Self-sovereign identity as trusted root in knowledge based systems. In *Proceedings of International Conference on Computational Science and Its Applications*. Springer, Cham, 14–24.

[100] Y. Kurihara. 2020. Self-sovereign identity and blockchain-based content management. In *Proceedings of IFIP International Conference on Human Choice and Computers*. Springer, Cham, 130–140.

[101] R. Laborde, A. Oglaza, S. Wazan, F. Barrère, A. Benzekri, D. W. Chadwick, and R. Venant. 2020. A user-centric identity management framework based on the W3C verifiable credentials and the FIDO universal authentication framework. In *Proceedings of IEEE 17th Annual Consumer Communications & Networking Conference (CCNC)*. IEEE, 1–8.

[102] Y. Liu, Q. Lu, H.-Y. Paik, X. Xu, S. Chen, and L. Zhu. 2020. Design pattern as a service for blockchain-based self-sovereign identity. *IEEE Software* 37, 5 (2020), 30–36.

[103] M. Luecking, C. Fries, R. Lamberti, and W. Stork. 2020. Decentralized identity and trust management framework for Internet of Things. In *Proceedings of IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*. IEEE, 1–9.

[104] Z. A. Lux, D. Thatmann, S. Zickau, and F. Beierle. 2020. Distributed-ledger-based authentication with decentralized identifiers and verifiable credentials. In *Proceedings of 2nd Conference on Blockchain Research & Applications for Innovative Networks and Services (BRAINS)*. IEEE. 71–78.

[105] P. N. Mahalle, G. Shinde, and P. M. Shafi. 2020. Rethinking decentralised identifiers and verifiable credentials for the internet of things. In *Internet of Things, Smart Computing and Technology: A Roadmap Ahead*. Springer, 361–374.

[106] P. N. Mahalle and G. R. Shinde. 2020. OAuth-based authorization and delegation in smart home for the elderly using decentralized identifiers and verifiable credentials. In *Security Issues and Privacy Threats in Smart Ubiquitous Computing*. 95–109.

[107] R. Mukta, J. Martens, H.-y. Paik, Q. Lu, and S. S. Kanhere. 2020. Blockchain-based verifiable credential sharing with selective disclosure. In *Proceedings of IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*. IEEE, 959–966.

[108] N. Naik and P. Jenkins. 2020. uPort open-source identity management system: An assessment of self-sovereign identity and user-centric data platform built on blockchain. In *Proceedings of IEEE International Symposium on Systems Engineering (ISSE)*. IEEE, 1–7.

[109] N. Naik and P. Jenkins. 2020. Self-sovereign identity specifications: Govern your identity through your digital wallet using blockchain technology. In *Proceedings of 8th IEEE International Conference on Mobile Cloud Computing, Services, and Engineering (MobileCloud)*. IEEE, 90–95.

[110] N. Naik and P. Jenkins. 2020. Governing principles of self-sovereign identity applied to blockchain enabled privacy preserving identity management systems. In *Proceedings of IEEE International Symposium on Systems Engineering (ISSE)*. IEEE, 1–6.

[111] A. Othman and J. Callahan. 2020. A protocol for decentralized biometric-based self-sovereign identity ecosystem. *Securing Social Identity in Mobile Platforms*, Springer. 217–234.

[112] A.-E. Panait, R. F. Olimid, and A. Stefanescu. 2020. Analysis of uPort Open, an identity management blockchain-based solution. In *Proceedings of International Conference on Trust and Privacy in Digital Business*. Springer, Cham, 3–13.

[113] D. Pöhn and W. Hommel. 2020. An overview of limitations and approaches in identity management. In *Proceedings of the 15th International Conference on Availability, Reliability and Security*. 1–10.

[114] A. Satybaldy, M. Nowostawski, and J. Ellingsen. 2020. Self-sovereign identity systems evaluation framework. *Privacy and Identity Management. Data for Better Living: AI and Privacy: 14th IFIP WG 9.2, 9.6/11.7, 11.6/SIG 9.2. 2 International Summer School*, Windisch, Switzerland, 447–461.

[115] Q. Stokkink, D. Epema, and J. Pouwelse. 2020. A truly self-sovereign identity system. *arXiv preprint* arXiv:2007.00415.
[116] P. Szalachowski. 2020. Password-authenticated decentralized identities. *arXiv preprint* arXiv:2007.15881.
[117] S. Terzi, C. Savvaidis, K. Votis, D. Tzovaras, and I. Stamelos. 2020. Securing emission data of smart vehicles with blockchain and self-sovereign identities. In *Proceedings of IEEE International Conference on Blockchain*. IEEE. 462–469.
[118] P. Vachon. 2020. The identity in everyone's pocket. *Communications of the ACM* 64 1 (2020), 46–55.
[119] R. van Kranenburg, L. Anania, G. Le Gars, M. Arniani, D. Fantini van Ditmar, M. Kaili, and P. Kavassalis. 2020. Future urban smartness: Connectivity zones with disposable identities. *Handbook of Smart Cities* (2020). 1–29.
[120] K. Wittek, L. Lazzati, D. Bothe, A.-J. Sinnaeve, and N. Pohlmann. 2020. An SSI based system for incentivized and self-determined customer-to-business data sharing in a local economy context. In *Proceedings of IEEE European Technology and Engineering Management Summit (E-TEMS)*. IEEE, 1–5.
[121] C. Wu, F. Yuan, Y. Lu, and S. Qi 2020. PSM2: A Privacy-preserving self-sovereign match-making platform. In *Proceedings of International Conference on Blockchain and Trustworthy Systems*. Springer, Singapore, 126–141.
[122] M. Xiao, Z. Ma, and T. Li. 2020. Privacy-preserving and scalable data access control based on self-sovereign identity management in large-scale cloud storage. In *Proceedings of International Conference on Security, Privacy and Anonymity in Computation, Communication and Storage*. Springer, Cham, 1–18.
[123] X. Yang and W. Li. 2020. A zero-knowledge-proof-based digital identity management scheme in blockchain. *Computers & Security* 99 (2020), 102050.
[124] E. Bandara, X. Liang, P. Foytik, S. Shetty, N. Ranasinghe, K. De Zoysa, and W. K. Ng. 2021. Promize-blockchain and self-sovereign identity empowered mobile ATM platform. *Intelligent Computing*. Springer. 891–911.
[125] I. Barclay, A. Preece, I. Taylor, S. K. Radha, and J. Nabrzyski. 2021. Providing assurance and scrutability on shared data and machine learning models with verifiable credentials. *arXiv preprint* arXiv:2105.06370.
[126] G. Fedrecheski, L. C. Costa, S. Afzal, J. M. Rabaey, R. D. Lopes, and M. K. Zuffo. 2021. A low-overhead approach for self-sovereign identity in IoT. *arXiv preprint* arXiv:2107.10232.
[127] J. Geng, N. Kanwal, M. G. Jaatun, and C. Rong. 2021. DID-eFed: Facilitating federated learning as a service with decentralized identities. *Evaluation and Assessment in Software Engineering* (2021), 329–335.
[128] M. Grabatin and W. Hommel. 2021. Self-sovereign identity management in wireless ad hoc mesh networks. In *Proceedings of IFIP/IEEE International Symposium on Integrated Network Management (IM)*. IEEE, 480–486.
[129] L. Helminger, D. Kales, S. Ramacher, and R. Walch. 2021. Multi-party revocation in Sovrin: Performance through distributed trust. In *Proceedings of Cryptographers' Track at the RSA Conference*. Springer, Cham, 527–551.
[130] D. N. Kirupanithi and A. Antonidoss. 2021. Self-sovereign identity creation on blockchain using identity based encryption. In *Proceedings of 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS)*. IEEE, 299–304.
[131] G. Laatikainen, T. Kolehmainen, M. Li, M. Hautala, A. Kettunen, and P. Abrahamsson. 2021. Towards a trustful digital world: Exploring self-sovereign identity ecosystems. *arXiv preprint* arXiv:2105.15131.
[132] S. Mahula, E. Tan, and J. Crompvoets. 2021. With blockchain or not? Opportunities and challenges of self-sovereign identity implementation in public administration: Lessons from the Belgian case. In *Proceedings of DG. O2021: The 22nd Annual International Conference on Digital Government Research*. 495–504.
[133] A. Pfeiffer and M. Bugeja. 2021. Introducing the concept of "digital-agent signatures": How SSI can be expanded for the needs of industry 4.0. *Artificial Intelligence in Industry 4.0: A Collection of Innovative Research Case-studies that are Reworking the Way We Look at Industry 4.0 Thanks to Artificial Intelligence* (2021), 213–233.
[134] M. Shuaib, S. Alam, M. S. Alam, and M. S. Nasir. 2021. Self-sovereign identity for healthcare using blockchain. *Materials Today: Proceedings*.
[135] M. Shuaib, S. Alam, M. S. Nasir, and M. S. Alam. 2021. Immunity credentials using self-sovereign identity for combating COVID-19 pandemic. *Materials Today: Proceedings*.
[136] A. Siqueira, A. F. Da Conceição, and V. Rocha. 2021. Blockchains and self-sovereign identities applied to healthcare solutions: A systematic review. 2021. *arXiv preprint* arXiv:2104.12298.
[137] L. Stockburger, G. Kokosioulis, A. Mukkamala, R. R. Mukkamala, and M. Avital. 2021. Blockchain-enabled decentralized identify management: The case of self-sovereign identity in public transportation. *Blockchain: Research and Applications* (2021), 100014.
[138] C. Sullivan. 2021. *Blockchain-Based Identity: The Advantages and Disadvantages*. Blockchain and the Public Sector. Springer, 197–218.
[139] P. Zhang and T.-T. Kuo. 2021. The feasibility and significance of employing blockchain-based identity solutions in health care. In *Proceedings Blockchain Technology and Innovations in Business Processes*. Springer, 189–208.
[140] C. Allen. 2021. The path to self-sovereign identity. 2021; [Online]. Available: http://www.lifewithalacrity.com/2016/04/the-path-to-self-soverereign-identity.html

[141] W3C. 2021. Decentralized Identifiers (DIDs) core architecture, data model, and representations. 2021. [Online]. Available: https://www.w3.org/TR/did-core/
[142] Identity.Foundation. 2021. Decentralized Identity Foundation. 11 Dec. 2021; [Online]. Available: https://identity.foundation/
[143] F. Schardong and R. Custódio. 2022. Self-Sovereign Identity: A systematic review, mapping and taxonomy. *Sensors* 22 (2022), 5641.
[144] Š. Čučko and M. Turkanović. 2021. Decentralized and self-sovereign identity: Systematic mapping study. *IEEE Access* 9 (2021), 139009–139027.
[145] C. Nyst, P. Makin, S. Pannifer, and E. A. Whitley. 2016. Digital identity: Issue analysis: Executive summary. 2016. [Online]. Available: https://chyp.com/wp-content/uploads/2020/06/PRJ.1578-Omidyar-Network-Digital-Identity-Issue-Analysis-Executive-Summary-v1_2-1.pdf
[146] B. Tarika. 2019. A review on models of software development life cycle. *Journal of Data Research* 3 (2019). DOI: https://doi.org/10.31058/j.data.2019.34002
[147] Web 3.0 Foundation. 11 Dec. 2021; [Online]. Available: https://web3.foundation/
[148] R. Soltani, U. T. Nguyen, and A. An. 2021. A survey of self-sovereign identity ecosystem. Volume 2021, Article ID 8873429, Security and Communication Networks.
[149] R. Nokhbeh Zaeem, K. C. Chang, T. C. Huang, D. Liau, W. Song, A. Tyagi, M. Khalil, M. Lamison, S. Pandey and K. S. Barber. 2021. Blockchain-Based Self-Sovereign identity: Survey, requirements, use-cases, and comparative study. *Proceedings of IEEE/WIC/ACM International Conference on Web Intelligence*, 2021.
[150] A. Mühle, A. Grüner, T. Gayvoronskaya, and C. Meinel. 2018. A survey on essential components of a self-sovereign identity. *Computer Science Review* 30, (November 2018). 80–86.
[151] J. Kaneriya and H. Patel. 2022. A comparative survey on blockchain based self sovereign identity system. In *Proceedings of the 3rd International Conference on Intelligent Sustainable Systems (ICISS)*.
[152] B. Houtan, A. Hafid, and D. Makrakis. 2020. A survey on blockchain-based self-sovereign patient identity in healthcare. *IEEE Access*, Volume 8, 2020. Corpus ID: 218895317.
[153] J. Asswad and J. M. Gómez. 2021. Data ownership: A survey. *Information* 12, 11 (2021), 465.
[154] P. Hummei, M. Braun, M. Tretter, and P. Dabrock. 2021. Data sovereignty: A review. *Journal of Big Data and Society* 8, 1 (January, 2021), 205395172098201.
[155] J. Andrieu. 2016. A technology-free definition of self-sovereign identity. In *Proceedings of the Rebooting the Web of Trust III; Web of Trust*: San Francisco, CA, USA, 2–5. https://github.com/WebOfTrustInfo/self-sovereign-identity
[156] A. Tobin and D. Reed. 2016. The inevitable rise of self-sovereign identity. *Sovrin Foundation*. Apr. 15 (2016); [Online]. Available: https://sovrin.org/wp-content/uploads/2018/03/The-Inevitable-Rise-of-Self-Sovereign-Identity.pdf.
[157] M. Schutte. 2016. Schutte's critique of the self-sovereign identity principles. 2016. [Online]. Available: http://matthewschutte.com/2016 /10/25/schuttes-critique-of-the-self-sovereign-identity-principles

Received 2 March 2022; revised 2 May 2023; accepted 3 August 2023