







## Research

### A Bibliometric Study of Scientific Production on Self-Sovereign Identity

#### Estudio bibliométrico de la producción científica sobre identidad auto-soberana

Roberto Albeiro Pava Díaz<sup>1</sup>, Rafael Vicente Páez Méndez<sup>2</sup>, and Luis Fernando Niño Vásquez<sup>3</sup>

<sup>1</sup>Universidad Distrital Francisco José de Caldas.

<sup>2</sup>Pontificia Universidad Javeriana.

<sup>3</sup>Universidad Nacional de Colombia.

#### Abstract

**Context:** Self-sovereign identity (SSI) enables the creation of user-centric, privacy-by-design, secure, and decentralized identity management systems. The aim of this paper is to carry out a bibliometric analysis of the scientific production on SSI during the 2017-2022 period.

**Method:** A complete bibliometric analysis of all publications on SSI indexed in Scopus and Web of Science was carried out. A corpus of 143 articles was examined by processing their bibliographic metadata via a bibliometric tool. To this effect, the Bibliometrix package and the R programming language were used.

**Results:** A bibliometric characterization of the publications on SSI was obtained for the 2017-2022 period. The most important keywords used in these publications were identified, as well as their use tendencies throughout this period. Moreover, the most influential authors in the area and the most relevant publication sources were identified.

**Conclusions:** The results of the bibliometric analysis show that Lotka and Bradford's laws apply for academic publications on SSI, which means that the most relevant publications in this area are concentrated in a relatively small group of authors and journals. Paul Jenkins, Nitin Naik, Yang Liu, and Aijun An were the most impactful authors, and *Lecture Notes in Computer Science*, *Frontiers in Blockchain*, and *IEEE Access* were the most influential journals. Finally, the keyword analysis showed that *blockchain*, *authentication*, *identity management*, *electronic document identification systems*, and *digital identity* are currently the most relevant concepts for research on SSI.

**Keywords:** self-sovereign identity, SSI, DLT, blockchain, identity management, bibliometric analysis

#### Article history

**Received:**  
11<sup>th</sup>/Jul/2021

**Modified:**  
17<sup>th</sup>/Sep/2022


**Accepted:**  
20<sup>th</sup>/Jan/2023

*Ing.*, vol. 28,  
no. suppl, 2023.  
e19656

©The authors;  
reproduction  
right holder  
Universidad  
Distrital  
Francisco José de  
Caldas.

#### Open access



\* **Correspondence:** rapavad@udistrital.edu.co

## Resumen

**Contexto:** La identidad auto-soberana (SSI) permite la creación de sistemas de gestión de identidad centrados en el usuario, con privacidad desde el diseño, seguros y descentralizados. El objetivo de este artículo es realizar un análisis bibliométrico de la producción científica sobre SSI durante el periodo 2017-2022.

**Método:** Se realizó un análisis bibliométrico completo de las publicaciones sobre SSI indexadas en Scopus y Web of Science. Se examinó un corpus de 143 artículos mediante el procesamiento de sus metadatos bibliográficos mediante una herramienta de análisis bibliométrico. Para ello se utilizaron el paquete Bibliometrix y el lenguaje de programación R.

**Resultados:** Se obtuvo una caracterización bibliométrica de las publicaciones sobre SSI durante el periodo 2017-2022. Se identificaron las palabras clave más importantes, así como las tendencias de uso de las mismas durante este periodo. Además, se identificaron los autores de mayor influencia en el área y las fuentes de publicación más relevantes.

**Conclusiones:** Los resultados del análisis bibliométrico completo muestran que la ley de Lotka y la ley de Bradford se cumplen en las publicaciones sobre SSI. Esto quiere decir que las publicaciones de mayor alcance e impacto están concentradas en unos pocos autores y revistas. Paul Jenkins, Nitin Naik, Yang Liu y Aijun An resultaron ser los autores más representativos, y *Lecture Notes in Computer Science*, *IEEE Access* y *Frontiers in Blockchain* resultaron ser las revistas más influyentes. Finalmente, el análisis de las palabras clave mostró que *blockchain*, *authentication*, *identity management*, *electronic document identification systems* y *digital identity* son actualmente los conceptos más importantes para la investigación sobre SSI.

**Palabras clave:** identidad digital auto-soberana, SSI, DLT, blockchain, gestión de identidad, análisis bibliométrico

## Table of contents

		4.2.1. Geographical distribution . . . . .	12
	Page	4.3. Most relevant sources . . . . .	12
1. Introduction	2	4.4. Keywords . . . . .	14
2. Self-sovereign identity	3	4.5. Citations analysis . . . . .	19
3. Search protocol	4	4.5.1. Most cited articles . . . . .	19
4. Results and discussion	5	5. Conclusions	20
4.1. Productivity analysis . . . . .	5	6. Author contributions	20
4.2. Bibliometric analysis of authors	7	References	20

## 1. Introduction

Technological advancements can sometimes create as many problems as they solve. For instance, the Internet has improved the lives of billions of people around the world by

democratizing the access to information, but it has also facilitated breaches of privacy, identity theft, and the misuse of personal data. Protecting and preserving digital identity is thus one of the most pressing technological challenges currently weighing on researchers and developers. The purpose of this paper is to apply a bibliometric analysis to the scientific literature in order to determine the extent to which the protection of digital identity has been lately addressed. The authors of (39) claimed that individuals have a fundamental right to be able to prove, manage, and preserve their identities, both physical and digital. For this reason, governments around the world ought to establish mechanisms that allow for the identification of citizens, the regulation of personal data use, and the protection of their privacy. However, current research on centralized identity management systems shows that they are often vulnerable to attacks to user privacy and identity theft (39), (19), (18), (28). Moreover, a centralized or federated approach to personal data management has generated a fragmented identity, which entails a loss of control over the information on the part of its owners. Since people have no say in the way that their data are stored, they will remain vulnerable to having it used in unintended ways. For instance, when leaks from several centralized sources are contrasted with one another, the identity of users whose data was stored anonymously in any of those sources might be revealed (17), (12). Hence, self-sovereign identity (SSI) was created, in order to allow users to manage their own personal data. With this technology, for instance, users may have a multifaceted identity, composed of a subset of identities, which would allow them to provide any platform or company only with the bare minimum data necessary to obtain the service that users wish to acquire (24). Thus, SSI grants users control of all their information (60) in a secure and decentralized fashion. As an example, Sovrin (64), uPort (16), ShoCard (55), and Civic (15) are frameworks for the deployment of SSI. In order to evaluate the spread of SSIs, this study analyzes the metadata of all articles published on this topic between 2017 and 2022 in the Web of Science and Scopus scientific databases. Moreover, this article provides useful information on the publishing dynamics of articles about SSI. This study was carried out with the bibliometrix package (6) of R (53), given its mathematical functions, statistics, and graphical capacity. The partial results obtained of the bibliometric analysis are available in (52). This article is structured as follows: Section 2 presents a general contextualization of SSI; Section 3 explains the implemented search protocol; Section 4 outlines and analyzes the main findings of the bibliometric study; and Section 5 presents the conclusions.

## 2. Self-sovereign identity

SSI is a type of decentralized identity protocol in which all information is managed directly by its owner. This system encourages its users to generate a set of unique decentralized identifiers in order to better control, manage, and store their personal data. For instance, SSI provides its users with the possibility of adjusting the amount of identifying information that they are willing to provide to third parties in different contexts. Indeed, not all services require the same type nor amount of data to function: a financial platform does not need access to the user's medical history, for instance, nor does a social network need to know their financial information. With such a system in place, each service would only be granted targeted access

to its user's personal information. (62) define *identity* as the set of attributes of a person (which may be further divided into subsets) that allows them to identify themselves as having this or that role in the real world. Thus, with the aid of a SSI, each of the roles of a person may have a unique and single identifier that acts as a reference to one dimension of their real physical identity.

The ten principles of SSI may be grouped into three categories: (i) security: existence, protection, and persistence; (ii) self-sovereignty: control, consent, and minimization; and (iii) trust: interoperability, transparency, access, and portability. However, proper SSIs must be able to face the problem posed by Zooko's Triangle (40), *i.e.*, to create decentralized unique identifiers that are (i) readable by humans, (ii) secure, and (iii) decentralized. Take, for instance, the case of a hypothetical identity protocol based on a Blockchain address. Any identifier produced in this system would of course be secure, since data could not be modified once it entered the protocol; and it would also be decentralized, since the whole network would guarantee that the same identifier is not assigned to two or more users. However, such an identifier would not be easily readable by humans, given that each identifier would be represented as a long string of hexadecimal characters.

### 3. Search protocol

This study is based on a bibliometric analysis of the Scopus and Web of Science (WoS) scientific databases. Scopus is a high-quality peer-reviewed database containing more than 82 million documents, over seven thousand publishers, and more than 80 thousand institutions (20). WoS, on its part, contains information from seven indexed databases and four multidisciplinary repositories: (i) The Science Citation Index Expanded, with around 9.200 journals and more than 53 million documents; (2) The Social Sciences Citation Index, with about 3.400 journals and more than 9 million documents; (3) The Arts and Humanities Citation Index, with about 1.800 journals and more than 4,9 million papers; and (4) The Current Chemical Reactions and Index Chemicals (5).

The dataset of bibliographical references analyzed in this study was built by searching documents (articles, conference papers, or article reviews) written in English between January 1, 2017, and April 8, 2022. In order to arrive at a final document list, three iterations of the same process were carried out. In each iteration, a search string was introduced in the search engine of both databases, which resulted in a list of references in the BibTeX format. Thus, the metadata of each reference (author, title, volume, year, keywords, keywords-plus, abstract, publisher, DOI, affiliation, type of document, and references) were downloaded. Then, the entire dataset was analyzed with a reproducible approach, using a script designed with the bibliometrix 4.0 package (6) of the R programming language (version 4.1.3) in the RStudio 2022.02.1+46 editor. This package provides many functions for processing bibliographic references. For instance, it allows for the extraction of terms, the elimination of duplicates, the production of general descriptive data, and the visualization of the networks of interaction among different bibliometric metadata. Finally, the obtained bibliometric outputs were compared, and the search string was adjusted for the next iteration of the process. Table I shows the search string used in the last iteration of the

process.

**Table I.** Description of the search string used in the last iteration of the process for retrieving the publications on SSI to be analyzed in this study

<b>Scientific databases</b>	Scopus	132
	Web of Science	37
<b>Search terms by</b>	Title	“self-sovereign identity” OR (blockchain-based AND “identity management” OR “digital identity”)
	Keywords	(“Identity Management” OR idm) AND (ssi OR self-sovereign)
<b>Document types</b>	Article, conference paper or review article	
<b>Report date</b>	April 6th, 2022	
<b>Timespan</b>	2017 - 2022	
<b>Language</b>	English	

This process yielded 132 results in Scopus and 37 in WoS. After eliminating 26 duplicate documents, the final set to be analyzed had 143 publications. The BibTeX files downloaded from the databases, as well as their pre-processed version (bibliometrix) are available at (52).

Since the aim of this article is to carry out a bibliometric analysis (section 4), the following research questions will be answered: a) Who are the most productive authors in SSI and what is their distribution? b) What are the main or most relevant publications on SSI? c) Which keywords are the most widely used in SSI and which ones are lately rising in popularity? d) Which journals publish the most articles on SSI?

## 4. Results and discussion

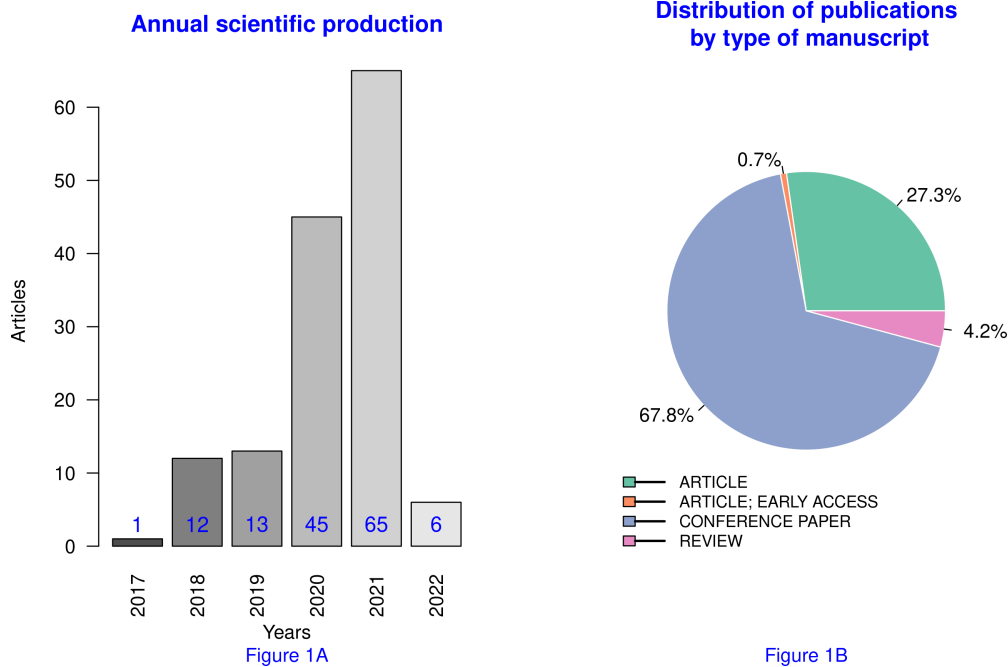
### 4.1. Productivity analysis

Table II contains the general bibliometric description of the set of publications obtained after the final iteration of the search process. The 143 documents found between 2017 and 2022 were distributed in 92 sources. However, three journals had a clear advantage in this area over the others: *Lecture notes in computer science* (10 publications), *Frontiers in blockchain* (8 publications), and *IEEE Access* (7 publications), covering 17,48% of the total documents. The results also showed a 43,1% annual growth rate in the publication of articles on SSI.

In addition, it was established that the total 143 articles were written by 366 authors, which shows that collaboration is a key feature of research in this area. Indeed, the results show that there are 0,391 articles per author and 2,56 authors per article; only 19 documents (13,29% of the total) were written by a single author. On the other hand, the analysis showed that the average number of citations per document during the analyzed period was 4,713, which means that each document had an average citation rate of 1,323 per year. Finally, 361 keywords and 639 keywords-plus were found in total.

**Table II.** General description of the obtained dataset

	Timespan	[2017,2022]
<b>Main information about data</b>	Sources	92
	Annual growth rate	43.1 %
	Total citations	4.495
	Average citations per document	4.713
	Average citations per year per document	1,323
<b>Documents (Total: 143)</b>	Articles	40
	conference papers	97
	Review articles	6
<b>Keywords</b>	Keyword-plus (ID)	639
	Keywords (DE)	361
<b>Authors</b>	All authors	366
	Author appearances	467
	Single-authored documents	19
	Documents per author	0.391
	Authors per documents	2.56



**Figure 1.** Annual scientific production (left) and distribution of publications by type of manuscript (right)

Fig. 1a illustrates the growing trend of publications on this topic, for the vast majority of scientific literature was produced in the second half of the analyzed period; 31, 47 % and 45, 45 % of the total articles published on SSI appeared during 2020 and 2021, respectively. Fig. 1b classifies the analyzed documents by type and shows that the majority of them were conference papers.

## 4.2. Bibliometric analysis of authors

Table III shows the top 15 authors analyzed in this study in terms of volume of publications.

**Table III.** Top 15 most productive authors

Author	Frequency	Author	Frequency (fractional)
JENKINS P	7	JENKINS P	3,333333
NAIK N	7	NAIK N	3,333333
LIU Y	5	KULABUKHOVA N	2,250000
AN A	4	PANAIT A	2,166667
BANDARA E	4	LOCKWOOD M	2,000000
FOYTIK P	4	ANTONIDOSS A	1,500000
LIANG X	4	GIANNOPOULOU A	1,500000
LU Q	4	LIU Y	1,366667
MEINEL C	4	AN A	1,333333
NGUYEN U	4	NGUYEN U	1,333333
PANAIT A	4	SOLTANI R	1,333333
SHETTY S	4	ISHMAEV G	1,250000
SOLTANI R	4	KUBACH M	1,250000
ABRAHAM A	3	MEINEL C	1,166667
ANTONIDOSS A	3	OLIMID R	1,166667
<b>Table III-A</b>		<b>Table III-B</b>	

Despite them being a mere 4,10% of the total number of authors, they wrote 45,45% of the total articles. The frequency column of Table IIIa reports the total number of documents published by each author. The fractional frequency column of IIIb shows the distribution value, across all published documents, of each author. This value was normalized by dividing each document by the number of authors, which allowed assigning a numeric “contribution value” to each author for each of the documents that they published. Finally, the total fractional frequency for each author was calculated by adding all of their individual contribution values.

The relationship of authors per document shows an asymmetric distribution, is positive skewed with skewness values of 3,865335 for the 366 authors and 1,508548 for the top of 15 authors with the highest production, and leptokurtic with kurtosis values of 21,27048 for all authors and 4,325208 for the 15 most productive ones (Table IV). A high number of authors can be seen in a single publication, 84,43% (309 of 366), compared to 9,02% (33 of 366) with two documents and 6,56% (34 of 366) with three or more publications (Fig. 2a). Fig. 2b shows the distribution for the 15 most productive authors. Table IV presents the values of asymmetry, kurtosis, mean, median, variance, and standard deviation for both the 366 authors and the top 15.

Table V lists the citation metrics for the top 15 authors. The h-index measures productivity and impact by showing the number of articles (x) of a given author that have at least that number (x) of citations. This means that an author’s h-index cannot be higher than their total number of publications. The g-index, on the other hand, compares the accumulated citation frequency of



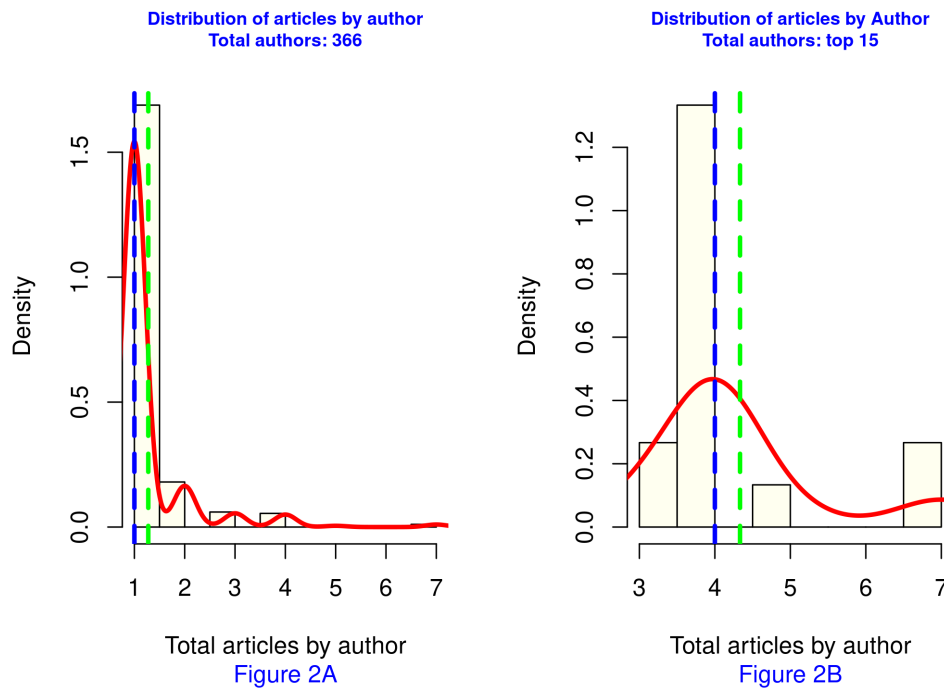


Figure 2. Distribution of documents by author. All authors (left) and top 15 authors (right)

Table IV. Values of the main statistic measurements for all authors and the top 15 ones

	Skewness	Kurtosis	Mean	Median	Variance	Standard Deviation
All Authors	3.863535	21.27048	1.275956	1	0.6222696	0.7888407
Top 15 Authors	1.508548	4.325208	4.333333	4	1.380952	1.175139

each author against their number of published articles squared. Finally, the m-index divides the h-index by the number of years during which a given author has been publishing.

The first two authors featured in Table V, Paul Jenkins and Nitin Naik, began publishing in 2020. They have co-authored 7 documents, 6 of which have been cited. They both have an h-index of 4, a g-index of 1, 333, and a total of 48 citations. These authors have published on SSI frameworks such as Sovrin (45) (46) and Uport (43); security risk analysis in SSI (41); access control and regulations such as GDPR (44); the principles of SSI (42); and digital wallets for SSI (42). Aijun An, Reza Soltani, and Uyen Trang Nguyen, positions 3, 5, and 6 in Table V, are co-authors of 4 articles, 3 of which have been cited. They have an h-index of 3, g-index of 4, and 27 citations. The topics of their articles are SSI ecosystem reviews (58), cryptographic key management (57), (56), and platform development for SSI (59). The author in the fourth position is Yang Liu, who has 5 published articles, an h-index of 3, a g-index of 5, and 42 citations. His research topics are review on identity management with blockchain (35) and platforms design



**Table V.** h-index, m-index and g-index of the top-15 authors

Author	h-index	g-index	m-index	Citations	Documents	Start year
JENKINS P	4	6	1.3333333	48	6	2020
NAIK N	4	6	1.3333333	48	6	2020
AN A	3	4	0.6000000	27	4	2018
LIU Y	3	5	0.7500000	42	5	2019
NGUYEN U	3	4	0.6000000	27	4	2018
SOLTANI R	3	4	0.6000000	27	4	2018
ABRAHAM A	2	2	0.6666667	7	2	2020
LU Q	2	4	0.6666667	16	4	2020
MEINEL C	2	2	0.4000000	15	2	2018
PANAIT A	2	2	0.6666667	9	4	2020
BANDARA E	1	1	0.5000000	2	2	2021
FOYTIK P	1	1	0.5000000	2	2	2021
LIANG X	1	1	0.5000000	2	2	2021
SHETTY S	1	1	0.5000000	2	2	2021

for SSI (36), (37), (35), (65). Qinghua Lu has an h-index of 2, a g-index of 4, and 16 citations. Furthermore, he co-authored 3 articles with Yang Liu. He also co-authored a paper on identity and the Internet of Things (IoT) (13). Eranga Bandara, Xueping Liang, Peter Foytik, and Sachin Shetty have co-authored 4 articles since 2021. They have an h-index of 1, a g-index of 1, and 2 citations in total. Their research topics are information exchange in SSI (7); development of platforms for SSI (9), (10); and privacy (8). Then comes Andreas Abraham, with an h-index of 2, a g-index of 2, and 17 citations. His research topics are digital wallets for SSI (3); verifiable credentials (2); and privacy (1). Christoph Meinel comes next, with an h-index of 2, a g-index of 2, and 15 citations. His articles are related to architecture for SSI (27), (26); SSI ecosystem reviews (54); and trust (25). Finally, the last author featured in Table V is Andreea-Elena Panait, with an h-index of 2, a g-index of 2, and 9 citations. Her research topics are security and privacy (51); zero-knowledge proofs (49); blockchain and digital identity (48); and frameworks for SSI (50). In summary, the top 15 authors on SSI can be grouped according to their collaborations. The authors that exclusively publish as a team are grouped in the first category: A1(Paul Jenkins and Nitin Naik), A2(Aijun An, Reza Soltani, and Uyen Trang Nguyen), and A3(Eranga Bandara, Xueping Liang, Peter Foytik, and Sachin Shetty). The authors that sometimes collaborate with each other are grouped in the second category: B1(Yang Liu and Qinghua Lu). Finally, the authors without co-authorships are grouped in the third category: Andreas Abraham, Christoph Meinel, and Andrea-Elena Panait.

Fig. 3a shows the participation of each top-15 author in the total 65 articles written by this upper section of researchers on SSI. This information may be used to prioritize the most important articles for a literature review on this topic. Fig. 3b shows the number of authors per document of the total 143 articles analyzed. Thus, an overview on cooperation among researchers on SSI was obtained.

The Most Productive Authors  
Total articles: 65

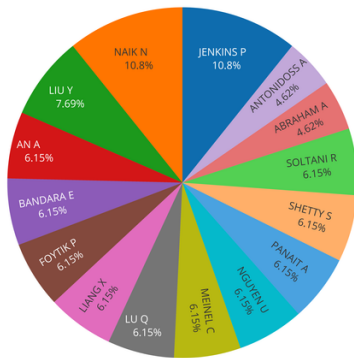


Figure 3A

Number of authors per article  
Total articles: 143

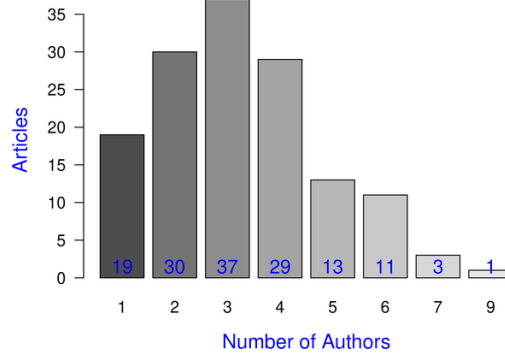


Figure 3B

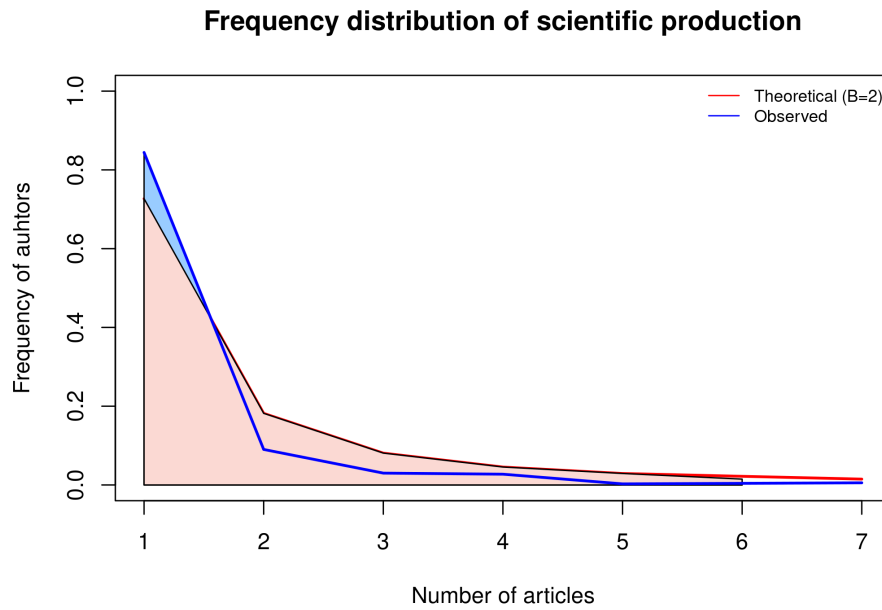
**Figure 3.** Productivity of the top 15 authors (left) and number of authors per document in the total articles on SSI (right)

Fig. 4 shows the publication timeline of the top 15 authors in SSI. The size of the circles in the graph is proportional to the number of documents published by each author, and the intensity of their color varies according to their citation index. Moreover, this figure shows which authors have had a constant output throughout the years. Considering all of these factors at once allows tracking the relevance of each author in the field. For instance, it could be noted that some authors had a high volume of publications in some years, but not so many citations, which indicates that some of their publications were not so influential.



**Figure 4.** Author-productivity timeline (top 15)

Finally, Fig. 5 and Table VI show how Lotka’s law applies to the articles analyzed in this study. This law states that there is always an unequal distribution of impact and productivity among



**Figure 5.** Author distribution (according to Lotka’s law)

the authors researching any given topic (38). Thus, the most relevant publications, regardless of the research area, will most likely be concentrated in a reduced number of authors. This common distribution makes it possible to quickly identify the most important authors in a given research field, but other factors must always be taken into account. Indeed, the quality of the top publications in different areas may vary, for instance, according to the popularity and nature of each research topic. Considering additional factors is always useful to compensate for the possible biases in some fields, which may lead to the exclusion of quality articles and authors from the (allegedly) most impactful group. To sum up, the application of Lotka’s law entails the creation of an inverse square model, where the number of authors  $A_n$  with  $n$  publications is inversely proportional to  $n^2$ .

**Table VI.** Author participation in academic production on SSI

Authored or co-authored documents	Authors	Frequency observed
1	309	0.8442623
2	33	0,0901639
3	11	0,0300546
4	10	0,0273224
5	1	0,0027322
7	2	0,0054645

Table VI shows the results of applying Lotka’s law to the dataset chosen for this study. Of the total 366 authors evaluated, 309 (84, 43 %) participated only in one publication, 33 (9, 02 %) in two, 11 (3, 01 %) in three, 10 (2, 73 %) in four, one (0, 27 %) in five, and finally two (0, 55 %) in seven.

in seven articles. Moreover, the actual beta coefficient was 2,8038, as opposed to the theoretical value of 2, with a constant of 0,7269 and a Kolmogorov-Smirnoff K-S goodness of fit of 0,910035 (p-value = 0,89278, independent samples t-test). Finally, the Kolmogorov-Smirnoff K-S test allowed concluding that the null hypothesis (H0), *i.e.*, the distribution follows the Lotka model, was most likely correct. Thus, given the p-value of the test, the null hypothesis was not rejected, with a confidence level of 85 %.

#### 4.2.1. Geographical distribution

Table VII groups the publications according to the country of origin of the authors and classifies them into two categories: single-country publications (SCP), publications with authors of a single nationality; and multiple-country publications (MCP), publications by authors of different nationalities.

**Table VII.** Geographical distribution the authors

Country	Documents	Citation frequency	SCP	MCP	MCP-rate
CHINA	12	0,1875	7	5	0,417
NETHERLANDS	5	0,0781	4	1	0,200
ROMANIA	5	0,0781	4	1	0,200
USA	5	0,0781	3	2	0,400
GERMANY	4	0,0625	4	0	0,000
INDIA	4	0,0625	4	0	0,000
ITALY	4	0,0625	4	0	0,000
BANGLADESH	3	0,0469	2	1	0,333
UNITED KINGDOM	3	0,0469	3	0	0,000
AUSTRALIA	2	0,0312	1	1	0,500

Moreover, Table VII shows that China has positioned itself as a relevant player in the blockchain ecosystem, since 18,75 % of the total analyzed citations were of Chinese authors (63). The Netherlands ranks second, followed by Romania and the United States. The latter had a 7,81 % frequency of citation and a lower level of international cooperation.

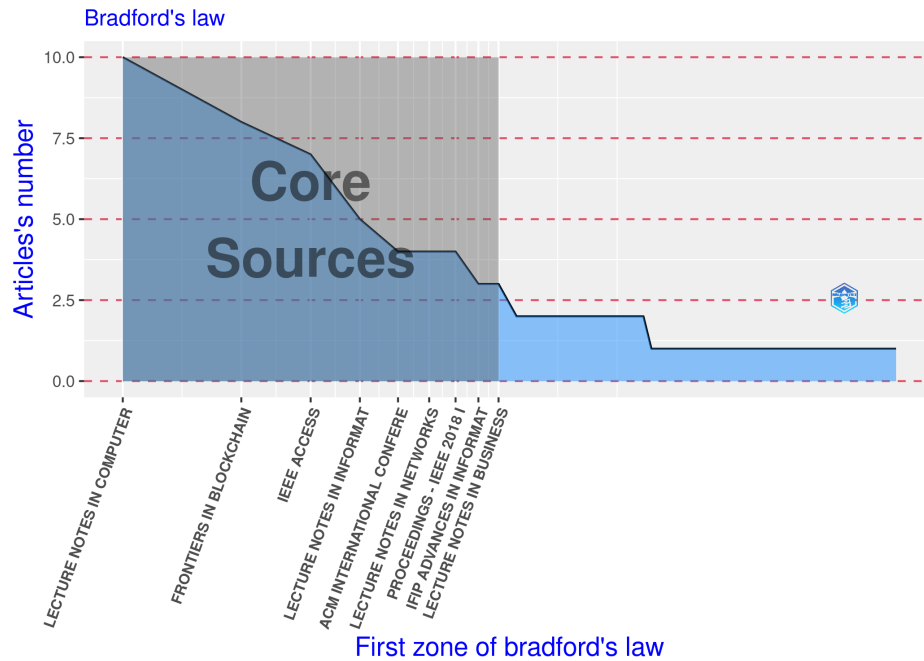
#### 4.3. Most relevant sources

Table VIII contains the 15 sources with the highest number of publications on SSI, *i.e.*, 41,96 % of the total evaluated documents appeared in one of said sources. The total number of published documents, the h-index, and the quartile is indicated when available in the Scimago index (32). It is worth noting that 11 of the top 15 sources are periodic conferences that publish the papers presented in them, which reflects a high positive impact on the dissemination of research results on SSI in this type of events. *IEEE Access* was the most influential source, according to its ranking (Q1) and its h-index (127), followed by *Frontiers in Blockchain*, a peer-reviewed and open-access journal with a multidisciplinary approach to blockchain.

**Table VIII.** Most relevant sources (top 15)

Sources	Documents	Type	h-index	Rank
Lecture Notes in Computer Science	10	Conference Proceedings, Book Series	400	Q3
	<a href="https://www.scimagojr.com/journalsearch.php?q=25674&amp;tip=sid&amp;clean=0">https://www.scimagojr.com/journalsearch.php?q=25674&amp;tip=sid&amp;clean=0</a>			
Frontiers in Blockchain	8	Journal	-	-
	<a href="https://www.frontiersin.org/journals/blockchain">https://www.frontiersin.org/journals/blockchain</a>			
IEEE Access	7	Journal	127	Q1
	<a href="https://www.scimagojr.com/journalsearch.php?q=21100374601&amp;tip=sid&amp;clean=0">https://www.scimagojr.com/journalsearch.php?q=21100374601&amp;tip=sid&amp;clean=0</a>			
Lecture Notes in Informatics (Lni) Proceedings	5	Conference Proceedings	25	-
	<a href="https://www.scimagojr.com/journalsearch.php?q=21100229106&amp;tip=sid&amp;clean=0">https://www.scimagojr.com/journalsearch.php?q=21100229106&amp;tip=sid&amp;clean=0</a>			
ACM International Conference Proceeding Series	4	Conference Proceedings	123	-
	<a href="https://www.scimagojr.com/journalsearch.php?q=11600154611&amp;tip=sid&amp;clean=0">https://www.scimagojr.com/journalsearch.php?q=11600154611&amp;tip=sid&amp;clean=0</a>			
Lecture Notes in Networks and Systems	4	Conference Proceedings	14	Q4
	<a href="https://www.scimagojr.com/journalsearch.php?q=21100901469&amp;tip=sid&amp;clean=0">https://www.scimagojr.com/journalsearch.php?q=21100901469&amp;tip=sid&amp;clean=0</a>			
2018 IEEE International Conference on Internet of Things and IEEE Green Computing and Communications (GreenCom)	4	Conference Proceedings	-	-
	<a href="https://www.computer.org/csdl/proceedings/1800308">https://www.computer.org/csdl/proceedings/1800308</a>			
Ifip Advances in Information and Communication Technology	3	Conference Proceedings	53	Q3
	<a href="https://www.scimagojr.com/journalsearch.php?q=19400157163&amp;tip=sid&amp;clean=0">https://www.scimagojr.com/journalsearch.php?q=19400157163&amp;tip=sid&amp;clean=0</a>			
Lecture Notes in Business Information Processing	3	Conference Proceedings, Book Series	49	Q3
	<a href="https://www.scimagojr.com/journalsearch.php?q=17500155101&amp;tip=sid&amp;clean=0">https://www.scimagojr.com/journalsearch.php?q=17500155101&amp;tip=sid&amp;clean=0</a>			
2ND Conference On Blockchain Research & Applications for Innovative Networks And Services	2	Conference Proceedings	-	-
	<a href="https://brains.dnac.org/2020/">https://brains.dnac.org/2020/</a>			
2020 International Symposium On Networks Computers And Communications Isncc 2020	2	Conference Proceedings	-	-
	<a href="https://ieeexplore.ieee.org/xpl/conhome/9297148/proceeding">https://ieeexplore.ieee.org/xpl/conhome/9297148/proceeding</a>			
Advances In Intelligent Systems And Computing	2	Conference Proceedings, Book Series	41	-
	<a href="https://www.scimagojr.com/journalsearch.php?q=5100152904&amp;tip=sid&amp;clean=0">https://www.scimagojr.com/journalsearch.php?q=5100152904&amp;tip=sid&amp;clean=0</a>			
Applied Sciences	2	Journal	52	Q2
	<a href="https://www.scimagojr.com/journalsearch.php?q=21100829268&amp;tip=sid&amp;clean=0">https://www.scimagojr.com/journalsearch.php?q=21100829268&amp;tip=sid&amp;clean=0</a>			
Future Internet	2	Journal	28	Q2
	<a href="https://www.scimagojr.com/journalsearch.php?q=21100409311&amp;tip=sid&amp;clean=0">https://www.scimagojr.com/journalsearch.php?q=21100409311&amp;tip=sid&amp;clean=0</a>			
Global Internet of Things Summit (GIoTS)	2	Conference Proceedings	-	-
	<a href="https://ieeexplore.ieee.org/xpl/conhome/9116578/proceedin">https://ieeexplore.ieee.org/xpl/conhome/9116578/proceedin</a>			

Bradford’s law divides a set of sources into three zones, each one of which contains approximately one third of the total number of articles on a given topic. The first zone is the core zone, as it groups the documents of the journals that have published the most on a given topic (14). The documents grouped in the second zone are those of journals with a relatively low impact. A greater number of these low-impact journals would be required to match the impact of any of the journals grouped in the core zone. Finally, the third zone groups the articles that have an even lower impact than those in the second zone. The impact of the articles decreases exponentially with each step from one zone to the next. Bradford’s law identifies the distribution



**Figure 6.** Bradford's law applied to the set of publications on SSI

of the sources in a given research area, and, for this reason, it can be used as a selection/exclusion criterion in a systematic literature review. Considering journals outside of the core zone would entail diminishing returns, increasingly, as one moves further away from it. Fig. 6 shows the results of applying Bradford's law to our research topic: the core zone, in this case, is made up of 48 articles (33, 57% of the total) and 9 sources (9, 78% of the total); the second zone comprises 48 articles (33, 57%) and 36 sources (39, 13%); the third zone contains 47 articles (32, 86%) and 47 sources (51, 09%). Thus, a linear regression analysis shows a strong negative correlation between the number of journals grouped in each zone and the impact of the documents of said journals. Indeed, a correlation coefficient (Pearson's  $r$ ) of  $-0,72341$  and a coefficient of determination ( $R^2$ ) of  $0,523322$  were obtained.

#### 4.4. Keywords

Fig. 7 depicts, through a word cloud, the keyword frequency (Fig. 7a) and the keywords-plus frequency (Fig. 7b) of the set of publications analyzed.

It is clear that technological elements such as distributed ledger technology, zero-knowledge proofs, and frameworks for SSI are crucial for research on this topic. Additionally, the principles and characteristics of SSI, such as decentralized identifiers, authentication, verifiable credentials, privacy, security, trust, access control, sovereignty, decentralization, and data protection (GDPR, EIDAS), are often given much attention. On the other hand, the analysis of keywords-plus (*i.e.*, a keyword list created by an algorithm that scans certain parts of the contents of each article) showed the importance of the IoT, data privacy, design for privacy, digital identity management,

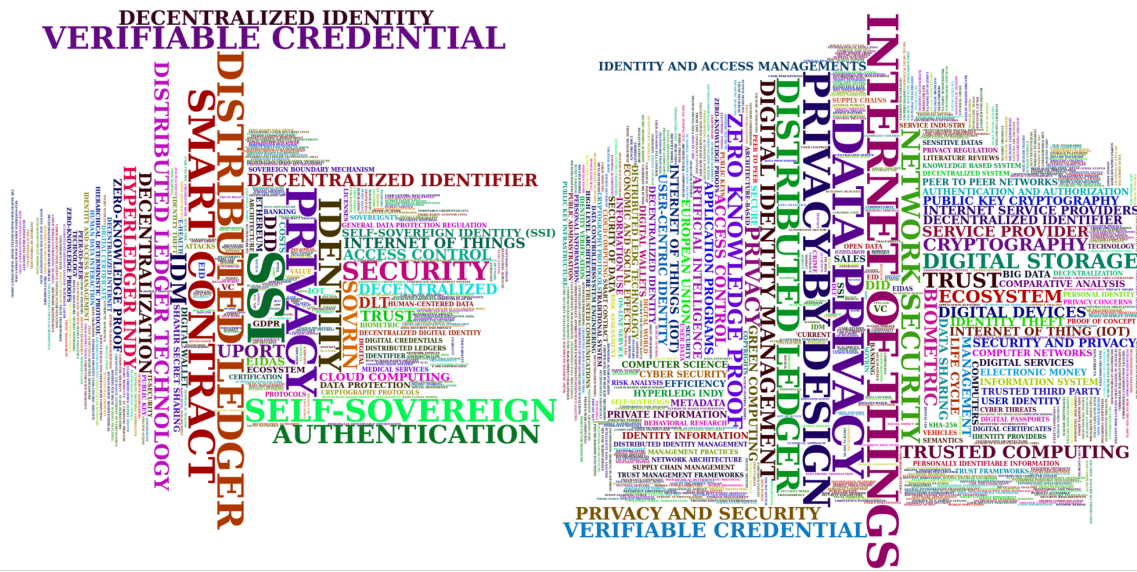


Figure 7A - Keywords

Figure 7B - Keywords-plus

Figure 7. Keyword and keyword-plus cloud

and network security for researchers on SSI. Moreover, the 361 keywords featured in the articles of the dataset had a cumulative frequency of 811. The 15 keywords with the highest frequency were blockchain (11,47%), self-sovereign identity (8,88%), identity management (4,81%), digital identity (2,84%), SSI (2,22%), privacy (1,85%), distributed ledger (1,6%), identity management system (1,6%), smart contract (1,6%), identity (1,36%), self-sovereign (1,36%), verifiable credential (1,36%), authentication (1,11%), security (1,11%), and IDM (0,99%). On the other hand, the 639 terms identified as keywords-plus had a cumulative frequency of 1,255. The 15 keywords-plus with the highest frequency were blockchain (7,25%), authentication (3,51%), identity management (3,51%), electronic document identification system (2,71%), digital identity (2,55%), self-sovereign identity (2,07%), identity management system (1,99%), data privacy (1,20%), Internet of Things (1,20%), information management (1,04%), privacy by design (1,04%), decentralized (0,88%), distributed ledger (0,88%), network security (0,72%), and privacy preserving (0,72%).

Table IX groups keywords and keywords-plus into four different categories: identity (C1), SSI (C2), attributes and principles (C3), and technological elements (C4). Fig. 8 presents the keyword-interconnection map (Fig. 8a) and the keyword-plus-interconnection map (Fig. 8b) obtained by determining the relationships between the terms collected and organized in Table IX.

The keyword-interconnection map comprises two main sections (red and blue), while the keyword-plus-interconnection map has three (red, green, and purple) surrounding a core concept: blockchain. Fig. 8a shows that authors have a tendency, in their keyword selection, to relate sovereign identity with security, trust, regulations for personal digital identification process —such as the Digital Identification Regulation in Europe (EIDAS)—, privacy, access



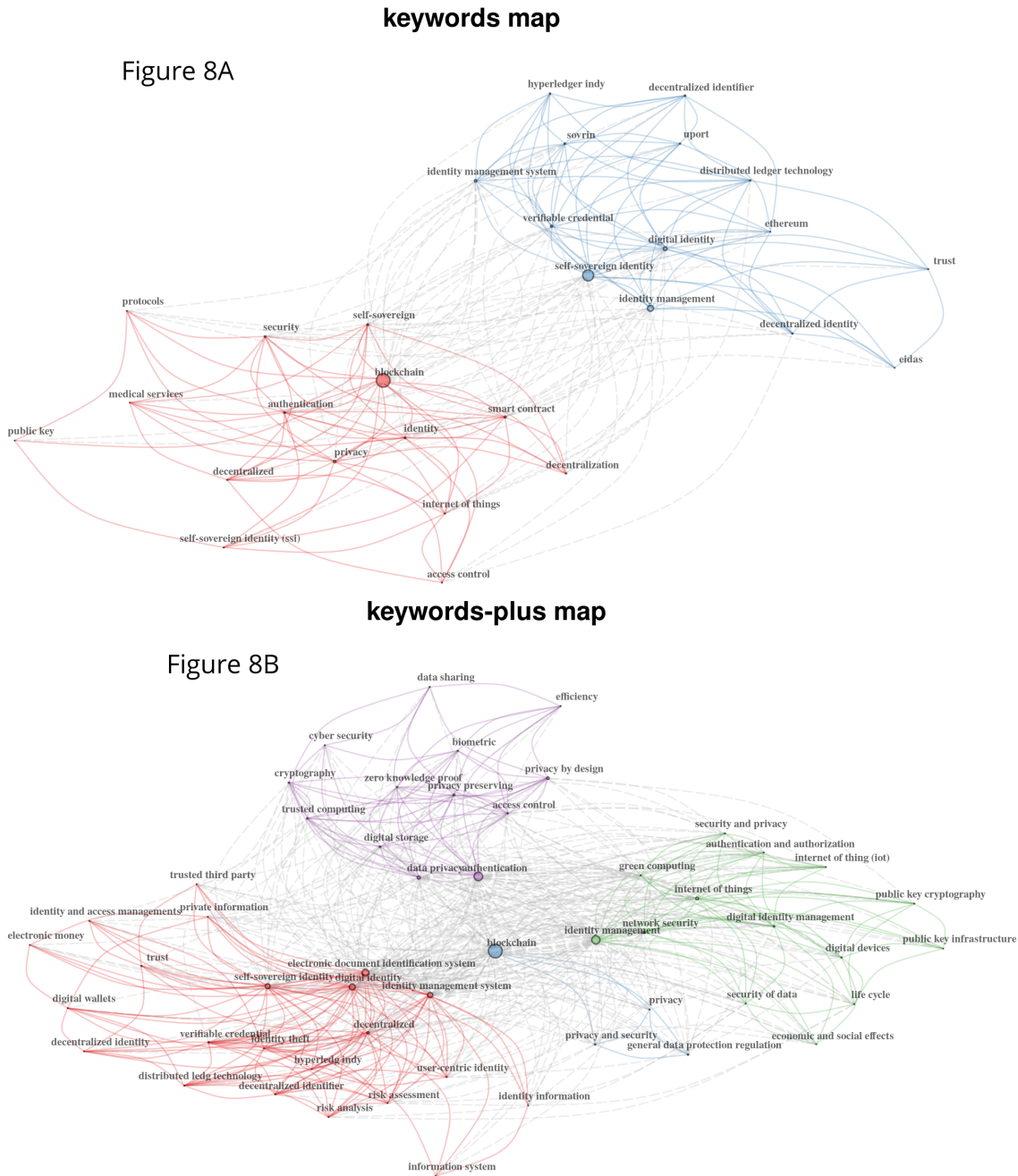
**Table IX.** Most frequent keywords and keywords-plus and their use by the top 15 authors

Category	Keywords	Keywords (361)	Keywords-plus (639)
C1 - Identity	DIGITAL IDENTITY	23	32
	IDENTITY MANAGEMENT	39	44
	IDENTITY MANAGEMENT SYSTEM	13	25
	IDM	8	
	IDENTITY	11	
C2 - SSI	SELF-SOVEREIGN IDENTITY	72	26
	SSI	18	
	PRIVACY	15	
C3 - Attributes, principles	DATA PRIVACY		15
	PRIVACY BY DESIGN		13
	PRIVACY PRESERVING		9
	AUTHENTICATION	9	44
	SECURITY	9	
	VERIFIABLE CREDENTIAL	11	
	SELF-SOVEREIGN	11	
	DECENTRALIZED		11
C4 - Technological elements	DISTRIBUTED LEDGER	13	11
	BLOCKCHAIN	93	91
	SMART CONTRACT	13	
	ELECTRONIC DOCUMENT IDENTIFICATION SYSTEM		34
	INTERNET OF THINGS		15
	INFORMATION MANAGEMENT		13
	NETWORK SECURITY		11

control, and authentication. They also tend to associate the concept of identity management with decentralization and self-sovereignty, distributed ledger technology, the Ethereum blockchain, decentralized identifiers, verifiable credentials, and frameworks for SSI such as Uport, Hyperledger Indy, and Sovrin. The blue section of the map roughly corresponds to the keywords in the C1 and C4 categories described in Table IX, while the red section corresponds to the C2 and C3 categories.

In Fig. 8b, on the other hand, the purple section of the map covers some of the keywords-plus grouped in the C3 category, in addition to some other important terms such as cryptography, zero-knowledge proofs, reliable computing, and cybersecurity. The green section of the map is focused on digital identity, linking identity management with authorization, authentication, security, privacy, and public keys. The terms covered in this section mostly belong to the C1 and C3 categories. Finally, the red section of the map relates concepts from all four categories, plus two additional concepts that were previously uncategorized: (i) user-centric identity with verifiable and decentralized credentials and (ii) identity-theft analyses and risk assessments.

To conclude this section, Fig. 9 shows the accumulated frequency of keyword use between 2018 and 2022. The notable increase in the appearance of blockchain among the keywords of articles on self-sovereign identity highlights the ever-growing importance of this disruptive technology in that area. On the other hand, the growing interest in identity management becomes clear by looking at the increase in the use of the following keywords: identity management,



**Figure 8.** Keyword-interconnection and keyword-plus-interconnection maps

electronic document identification systems, digital identity, identity management systems, and self-sovereign identity. Finally, the data shows that research has focused on authentication, privacy by design, and data regulations such as GDPR.

Finally, Fig. 10 shows the keyword use tendencies, as a percentage of the total number of articles published in a given year, for the top 15 keywords in the 2018-2022 period. The fact that

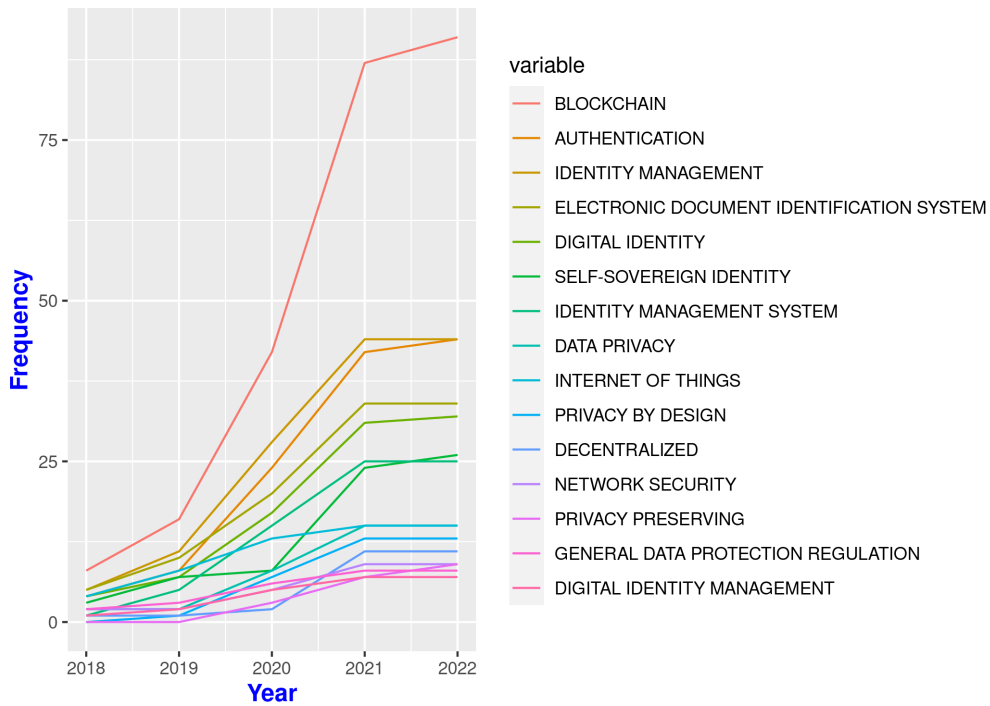


Figure 9. Cumulative frequency of the top 15 keywords (2018-2022)

### Keyword use tendencies (2018-2022)

(as a % of the total published articles)

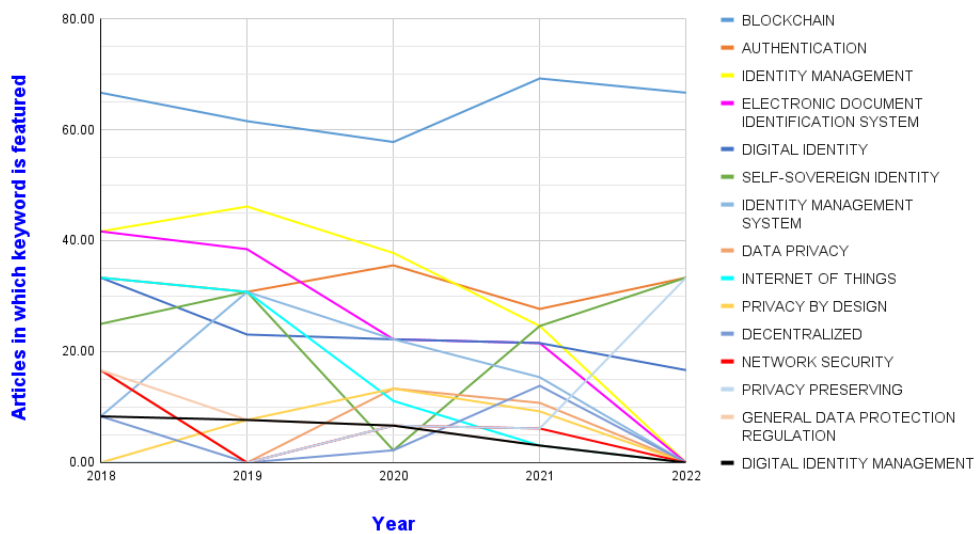


Figure 10. Keyword use tendencies (2018-2022)

blockchain has been included as a keyword in more than half of the articles on SSI throughout the whole timeframe under study, even taking into account the overall increase in academic production on that topic, also goes to show the importance of this technology for the development of new approaches to personal identity management.

## 4.5. Citations analysis

### 4.5.1. Most cited articles

A high citation index in an article is often telling of its quality and relevance. Moreover, highly cited articles are useful for analyzing the evolution of and the interactions within a line of research. Table X lists the 15 articles with the highest number of citations, along with their average number of citations per year.

**Table X.** Top 15 most cited articles

Ref.	Article title	Year	Citations	Average per year
(40)	A survey on essential components of a self-sovereign identity	2018	60	12,000000
(61)	Deployment of a blockchain-based self-sovereign identity	2018	35	7,000000
(34)	Blockchain-based identity management systems: A review	2020	27	9,000000
(33)	A new transitively closed undirected graph authentication scheme for blockchain-based identity management systems	2018	27	5,400000
(21)	BPDIMS:A blockchain-based personal data and identity management system	2019	25	6,250000
(23)	Blockchain-based identity management with mobile device	2018	25	5,000000
(22)	In search of self-sovereign identity leveraging Blockchain Technology	2019	23	5,750000
(47)	The horcrux protocol: A method for decentralized biometric-based self-sovereign identity	2018	23	4,600000
(31)	Blockchain-based identity management: A survey from the enterprise and ecosystem perspective	2020	22	7,333333
(11)	WiP: A novel blockchain-based trust model for cloud identity management	2018	21	4,200000
(66)	A survey on blockchain-based identity management systems for the internet of things	2018	20	4,000000
(29)	Analysis of identity management systems using blockchain technology	2019	18	4,500000
(4)	DNS-IdM: A blockchain identity management system to secure personal data sharing in a network	2019	17	4,250000
(62)	Self-sovereign identity in a globalized World: Credentials-based identity systems as a driver for economic inclusion	2020	17	5,666667
(30)	Self-sovereign identity on public blockchains and the GDPR	2020	17	5,666667

The articles published during 2018 focused on two main issues. On the one hand, they dealt with the organization and documentation of the scientific production on SSI, as well as the identity management for devices in the IoT. The article written by (40), with 60 citations and an average of 12 references per year, is particularly noteworthy in this area. Indeed, not only do they use the principles proposed by Christopher Allen (60) in order to contextualize SSI, but they present a general architecture for SSI, describe the authentication and identification operations necessary for it to work, explain the importance of verifiable credentials, and outline some data-storage challenges that would arise in any attempt to implement SSI. On the other hand, the

articles published in 2018 dealt with the design of decentralized systems for the implementation of SSI. The work of Quiinten Stokkink and Johan Pouwelse (61) stands out in particular. They analyze the problems of designing SSI and focus particularly in portability. Moreover, they explore minimization of disclosure, interoperability, and data protection (four out of Christopher Allen's ten principles). Finally, they include systems like TrustChain or Tangle in their design in order to store credentials and their metadata, as well as to verify them in the future.

## 5. Conclusions

This article presents a comprehensive bibliometric survey of the literature on self-sovereign identity published between 2017 and 2022 and indexed in WoS and Scopus. The analysis offers an overview of the current state of the research on SSI and its changes over the years. The set of evaluated publications is made up of 143 articles, out of which 67,8% are conference papers, 28% are articles and 4,2% are review articles. These articles were written by 366 authors from 274 countries, were published in 92 journals, and were cited 4,479 times. China is currently dominating the academic production in this area (44,44% of the total articles were written by authors of Chinese origin), followed by the Netherlands, Romania, and the USA. Furthermore, there was an annual publication growth rate of 43,09% between 2018 and 2022. Keyword examination considered 361 terms included directly by authors in their articles and 639 keyword-plus terms. A word cloud was constructed to offer an easily-readable depiction of the influence of each term, and the analysis was supplemented with a keyword-interconnection map. This analysis highlighted the impact of crucial terms for SSI, such as blockchain, identity management, electronic document identification system, and digital identity. Regarding the authors, Lotka's law allowed distinguishing specialized authors from transitory ones. Indeed, it could be observed that 24 (6,56%) authors out of a total 366 authored or co-authored 92 (64,34%) of the total 143 publications. Jenkins P., Naik N. (with seven published papers and an observed frequency of 0,0054645), and Liu Y. (with five documents and an observed frequency of 0,0027322) were the most specialized authors of the entire group. Very much in the same vein, Bradford's law helped to determine the main sources where most of the knowledge on SSI is being published: *Lecture Notes in Computer Science*, *Frontiers in Blockchain*, *IEEE Access*, *Lecture Notes in Informatics Proceedings*, *ACM International Conference Proceeding Series*, *Lecture Notes in Networks and Systems*, *2018 IEEE International Conference on Internet of Things and IEEE Green Computing and Communications*, *Advances in Information and Communication Technology*, and *Lecture Notes in Business Information Processing*.

## 6. Author contributions

All authors contributed equally to the research.

## References

- [1] A. Abraham, K. Koch, S. More, S. Ramacher, and M. Stopar, "Privacy-preserving eID derivation to self-sovereign identity systems with offline revocation," in *TrustCom*, pp. 506-513, 2021. <https://doi.org/10.1109/TrustCom53373.2021.00080> ↑9
- [2] A. Abraham, S. More, C. Rabensteiner, and F. Hörandner, "Revocable and offline-verifiable self-sovereign identities," in *TrustCom*, pp. 1020-1027, 2020. <https://doi.org/10.1109/TrustCom50675.2020.00136> ↑9
- [3] A. Abraham, C. Schinnerl, and S. More, "SSI strong authentication using a mobile-phone based identity wallet reaching a high level of assurance," in *Secrypt 2021*, pp. 137-148, 2021. <https://doi.org/10.5220/0010542801370148> ↑9
- [4] J. A. Kassem, S. Sayeed, H. Marco-Gisbert, Z. Pervez, and K. Dahal, "DNS-IdM: A blockchain identity management system to secure personal data sharing in a network," *Appl. Sci.*, vol. 9, no. 15, art. 2953, Jul. 2019. <https://doi.org/10.3390/app9152953> ↑19
- [5] Clarivate Analytics, "Web of science", 2022. [Online]. Available: <https://clarivate.com/> ↑4
- [6] M. Aria and C. Cuccurullo, "bibliometrix: An R-tool for comprehensive science mapping analysis," *J. Informetr.*, vol. 11, no. 4, pp. 959-975, 2017. <https://doi.org/10.1016/j.joi.2017.08.007> ↑3,4
- [7] E. Bandara, X. Liang, P. Foytik, and S. Shetty, "Blockchain and self-sovereign identity empowered cyber threat information sharing platform," in *Smartcomp*, 2021, pp. 258-263. <https://doi.org/10.1109/SMARTCOMP52413.2021.00057> ↑9
- [8] E. Bandara et al., "A blockchain empowered and privacy preserving digital contact tracing platform," *Inf. Process. Manag.*, vol. 58, no. 4, art. 102572. 2021. <https://doi.org/10.1016/j.ipm.2021.102572> ↑9
- [9] E. Bandara, X. Liang, P. Foytik, S. Shetty, and K. De Zoysa, "A blockchain and self-sovereign identity empowered digital identity platform," In *2021 ICCCN*, 2021, pp. 1-7. <https://doi.org/10.1109/ICCCN52240.2021.9522184> ↑9
- [10] E. Bandara et al., "Promize - blockchain and self-sovereign identity empowered mobile atm platform," in *Intelligent Computing. Lecture Notes in Networks and Systems.*, K. Arai, Ed., Berlin, Germany: Springer, 2021, vol. 284, pp. 891-911. [https://doi.org/10.1007/978-3-030-80126-7\\_63](https://doi.org/10.1007/978-3-030-80126-7_63) ↑9
- [11] K. Bendiab, N. Kolokotronis, S. Shiaeles, and S. Boucherkha, "Wip: A novel blockchain-based trust model for cloud identity management," in *2018 IEEE 16th Intl. Conf Dependable, Autonomic Sec. Comp., 16th Int. Conf. Perov. Intell. Comp., 4th Int. Conf. Big Data Intell. Comp. and Cyber Sci. Tech. Congress (DASC/PiCom/DataCom/CyberSciTech)*, 2018, pp. 724-729. <https://doi.org/10.1109/DASC/PiCom/DataCom/CyberSciTec.2018.00126> ↑19
- [12] J. Bernal-Bernabe, J. L. Canovas, J. L. Hernández-Ramos, R. Torres-Moreno, and A. Skarmeta, "Privacy-preserving solutions for blockchain: Review and challenges," *IEEE Access*, vol. 4, pp. 1-33, Nov. 2019. <https://doi.org/10.1109/ACCESS.2019.2950872> ↑3



- [13] M. A. Bouras, Q. Lu, S. Dhelim, and H. Ning, "A lightweight blockchain-based IoT identity management approach", *Fut. Int.*, vol. 13, no. 2, art. 24. 2021. <https://doi.org/10.3390/fi13020024> ↑9
- [14] B.C. Brookes, "'Sources of information on specific subjects" by S.C. Bradford," *J. Inf. Sci. Eng.*, vol. 10, no. 4, pp. 173-175, 1985. <https://doi.org/10.1177/016555158501000406> ↑13
- [15] Civic Technologies, "Civic whitepaper," 2017. [Online]. Available: <https://tokensale.civic.com/CivicTokenSaleWhitePaper.pdf> ↑3
- [16] ConsenSys, "uPort – Tools for Decentralized Identity and Trusted Data," 2020. [Online]. Available: <https://www.uport.me/> ↑3
- [17] Y. A. De Montjoye, E. Shmueli, S. S. Wang, and A. S. Pentland, "OpenPDS: Protecting the privacy of metadata through SafeAnswers," *PLoS One*, vol. 9, no. 7, art. 9, 2014. <https://doi.org/10.1371/journal.pone.0098790> ↑3
- [18] DIF Foundation, "DIF - decentralized identity foundation," 2020. [Online]. Available: <https://identity.foundation/> ↑3
- [19] P. Dunphy and F. A. P. Petitcolas, "A first look at identity management schemes on the blockchain," *IEEE Secur. Priv.*, vol. 16, no. 4, pp. 20-29, 2018. <https://doi.org/10.1109/MSP.2018.3111247> ↑3
- [20] Elsevier, "Scopus," 2022. [Online]. Available: <https://www.elsevier.com/solutions/scopus> ↑4
- [21] B. Faber, G. Michelet, N. Weidmann, R. R. Mukkamala, and R. Vatrappu, "BPDIMS: A blockchain-based personal data and identity management system," in *52nd Hawaii Int. Conf. Sys. Sci.*, 2019, pp. 6855-6864. <https://doi.org/10.24251/HICSS.2019.821> ↑19
- [22] Md. S. Ferdous, F. Chowdhury, and M. O. Alassafi, "In search of self-sovereign identity leveraging blockchain technology," *IEEE Access*, vol. 7, pp. 103059-103079, 2019. <https://doi.org/10.1109/ACCESS.2019.2931173> ↑19
- [23] Z. Gao et al., "Blockchain-based identity management with mobile device," in *CryBlock'18*, 2018, pp. 66-70. <https://doi.org/10.1145/3211933.3211945> ↑19
- [24] G. Goodell and T. Aste, "A decentralized digital identity architecture," *Front. Blockchain*, vol. 2, pp. 1-19, 2019. <https://doi.org/10.3389/fbloc.2019.00017> ↑3
- [25] A. Grüner, A. Mühle, T. Gayvoronskaya, and C. Meinel, "A quantifiable trust model for blockchain-based identity management," in *iThings/ GreenCom/ CPSCoM/ SmartData/ Blockchain/ CIT 2018*, 2018, pp. 1475-1482. [https://doi.org/10.1109/Cybermatics\\_2018.2018.00250](https://doi.org/10.1109/Cybermatics_2018.2018.00250) ↑9
- [26] A. Grüner, A. Mühle, and C. Meinel, "An integration architecture to enable service providers for self-sovereign identity," in *2019 IEEE 18th Int. Symp. Network Comp. App. (NCA)*, 2019, pp. 1-5. <https://doi.org/10.1109/NCA.2019.8935015> ↑9
- [27] A. Grüner, A. Mühle, and C. Meinel, "ATIB: Design and evaluation of an architecture for brokered self-sovereign identity integration and trust-enhancing attribute aggregation for service provider," *IEEE Access*, vol. 9, pp. 138553-138570, 2021. <https://doi.org/10.1109/ACCESS.2021.3116095> ↑9



- [28] O. J. Gstrein and D. Kochenov, "Digital identity and distributed ledger technology: Paving the way to a neo-feudal brave new world?" *Front. Blockchain*, vol. 3, art. 00010, 2020. <https://doi.org/10.3389/fbloc.2020.00010> ↑3
- [29] S. El Haddouti and M. Dafir Ech-Cherif El Kettani, "Analysis of identity management systems using blockchain technology" in *CommNet*, 2019, pp. 1-7. <https://doi.org/10.1109/COMMNET.2019.8742375> ↑19
- [30] G. Kondova and J. Erbguth, "Self-sovereign identity on public blockchains and the GDPR," in *SAC'20*, 2020, pp. 342-345. <https://doi.org/10.1145/3341105.3374066> ↑19
- [31] M. Kuperberg, "Blockchain-based identity management: A survey from the enterprise and ecosystem perspective," *IEEE Trans. Eng. Manag.*, vol. 67, no. 4, pp. 1008-1027, 2020. <https://doi.org/10.1109/TEM.2019.2926471> ↑19
- [32] Scimago Lab "Scimago journal and country rank," 2022. [Online]. Available: <https://www.scimagojr.com/index.php> ↑12
- [33] C. Lin, D. He, X. Huang, M. K. Khan, and K. Choo, "A new transitively closed undirected graph authentication scheme for blockchain-based identity management systems," *IEEE Access*, vol. 6, pp. 28203-28212, 2018. <https://doi.org/10.1109/ACCESS.2018.2837650> ↑19
- [34] Y. Liu, D. He, M. S. Obaidat, N. Kumar, M. K. Khan, and K. Choo, "Blockchain-based identity management systems: A review," *J. Network Comp. Appl.*, vol. 166, 2020. <https://doi.org/10.1016/j.jnca.2020.102731> ↑19
- [35] Y. Liu, Q. Lu, H.-Y. Paik, and X. Xiwei, "Design patterns for blockchain-based self-sovereign identity," in *EuroPLoP 2020*, 2020, art. 16. <https://doi.org/10.1145/3424771.3424802> ↑8,9
- [36] Y. Liu, Q. Lu, H.-Y. Paik, X. Xu, S. Chen, and L. Zhu, "Design pattern as a service for blockchain-based self-sovereign identity." *IEEE Soft.*, vol. 37, no. 5, pp. 30-36, 2020. <https://doi.org/10.1109/MS.2020.2992783> ↑9
- [37] Y. Liu, Q. Lu, C. Zhu, and Q. Yu, "A blockchain-based platform architecture for multimedia data management," *Multimedia Tools Appl.*, vol. 80, no. 20, pp. 30707-30723, 2021. <https://doi.org/10.1007/s11042-021-10558-z> ↑9
- [38] A. J. Lotka, "The frequency distribution of scientific productivity," *J. Wash. Acad. Sci.*, vol. 16, no.12, 317-323, 1926. ↑11
- [39] T. Lyons, L. Courcelas, and K. Timsit, "Blockchain and digital identity," 2020. [Online]. Available: <https://www.eublockchainforum.eu/reports/blockchain-and-digital-identity> ↑3
- [40] A. Mühle, A. Grüner, T. Gayvoronskaya, and C. Meinel, "A survey on essential components of a self-sovereign identity," *Comput. Sci. Rev.*, vol. 30, pp. 80-86, 2018. <https://doi.org/10.1016/j.cosrev.2018.10.002> ↑4,19
- [41] N. Naik, P. Grace, and P. Jenkins, "An attack tree based risk analysis method for investigating attacks and facilitating their mitigations in self-sovereign identity," in *2021 IEEE Symp. Series*

- Comp. Int. (SSCI)*, 2021, pp. 1-8. <https://doi.org/10.1109/SSCI50451.2021.9659929> ↑8
- [42] N. Naik and P. Jenkins, "Governing principles of self-sovereign identity applied to blockchain enabled privacy preserving identity management systems", in *ISSE 2020*, 2020, pp. 1-6. <https://doi.org/10.1109/ISSE49799.2020.9272212> ↑8
- [43] N. Naik and P. Jenkins. "üPort open-source identity management system: An assessment of self-sovereign identity and user-centric data platform built on blockchain," in *ISSE 2020*, 2020, pp. 1-7. <https://doi.org/10.1109/ISSE49799.2020.9272223> ↑8
- [44] N. Naik and P. Jenkins, "Your identity is yours: Take back control of your identity using GDPR compatible self-sovereign identity," in *2020 7th Int. Conf. Behaviour. Soc. Comp. (BESC)*, 2020, pp. 1-6. <https://doi.org/10.1109/BESC51023.2020.9348298> ↑8
- [45] N. Naik and P. Jenkins, "Does Sovrin network offer sovereign identity?" in *ISSE 2021*, 2021, pp. 1-6. <https://doi.org/10.1109/ISSE51541.2021.9582472> ↑8
- [46] N. Naik and P. Jenkins, "Sovrin network for decentralized digital identity: Analysing a self-sovereign identity system based on distributed ledger technology," in *ISSE 2021*, 2021, pp. 1-7. <https://doi.org/10.1109/ISSE51541.2021.9582551> ↑8
- [47] A. Othman and J. Callahan, "The horcrux protocol: A method for decentralized biometric-based self-sovereign identity," in *IJCNN 2018*, 2018, pp. 1-7. <https://doi.org/10.1109/IJCNN.2018.8489316> ↑19
- [48] A.-E. Panait, "Is the user identity perception influenced by the blockchain technology?, in *2020 IEEE Int. Conf. Intell. Sec. Inform. (ISI)*, 2020, pp. 1-3. ↑9
- [49] A.-E. Panait and R. F. Olimid, "On using zk-SNARKs and zk-STARKs in blockchain-based identity management," in *Innovative Security Solutions for Information Technology and Communications. SecITC 2020. Lecture Notes in Computer Science*, D. Maimut, A.-G. Oprina, and D. Sauveron, Eds., Cham, Germany: Springer International Publishing, 2021, vol. 12596, pp. 130-145. [https://doi.org/10.1007/978-3-030-69255-1\\_9](https://doi.org/10.1007/978-3-030-69255-1_9) ↑9
- [50] A.-E. Panait, R. F. Olimid, and A. Stefanescu, "Analysis of uPort Open, an identity management blockchain-based solution," in *Trust, Privacy and Security in Digital Business. TrustBus 2020. Lecture Notes in Computer Science*, S. Gritzalis, E. R. Weippl, G. Kotsis, A. Min Tjoa, and I. Khalil, Eds., Cham, Germany: Springer International Publishing, 2020, vol. 12395, pp. 3-13. [https://doi.org/10.1007/978-3-030-58986-8\\_1](https://doi.org/10.1007/978-3-030-58986-8_1) ↑9
- [51] A.-E. Panait, R. F. Olimid, and A. Stefanescu, "Identity management on blockchain-privacy and security aspects," *Proc. Rom. Acad. Ser. A Math. Phys. Tech. Sci. Inf. Sci.*, vol. 21, no. 1, pp. 45-52, 2020. ↑9
- [52] L. Niño-Vásquez, R. Pava-Díaz, R. Páez-Mendez, and D. López-Sarmiento, "Preprocesamiento de publicaciones acerca de identidad digital descentralizada y autogobernada," 2022. [Online]. Available: <https://doi.org/10.6084/m9.figshare.19579492> ↑3, 5
- [53] R Core Team, R: A language and environment for statistical computing,"2020. [Online]. Available: <https://www.R-project.org/> ↑3

- 
- [54] K. Schmidt, A. Mühle, A. Grüner, and C. Meinel, "Clear the fog: Towards a taxonomy of self-sovereign identity ecosystem members," in *2021 18th Int. Conf. Priv. Sec. Trust (PST)*, 2021, pp. 1-7. <https://doi.org/10.1109/PST52912.2021.9647797> ↑9
- [55] ShoCard Inc., "Identity management verified using the blockchain," 2017. [Online]. Available: <https://shocard.com/wp-content/uploads/2019/02/ShoCard-Whitepaper-2019.pdf> ↑3
- [56] R. Soltani, U. T. Nguyen, and A. An, "Practical key recovery model for self-sovereign identity based digital wallets," in *2019 DASC-PiCom-CBDCom-CyberSciTech*, 2019, pp. 320-325. <https://doi.org/10.1109/DASC/PiCom/CBDCom/CyberSciTech.2019.00066> ↑8
- [57] R. Soltani, U. T. Nguyen, and A. An, "Decentralized and privacy-preserving key management model," in *ISNCC 2020*, 2020, pp. 1-7. <https://doi.org/10.1109/ISNCC49221.2020.9297294> ↑8
- [58] R. Soltani, U. T. Nguyen, and A. An, "A survey of self-sovereign identity ecosystem," *Sec. Commun. Networks.*, vol. 2021, art. 8873429, 2021. <https://doi.org/10.1155/2021/8873429> ↑8
- [59] R. Soltani, U. T. Nguyen, and A. An, "A new approach to client onboarding using self-sovereign identity and distributed ledger," in *2018 iThings/GreenCom/CPSCo/SmartData/Blockchain/CIT*, 2018, pp. 1129-1136. [https://doi.org/10.1109/Cybermatics\\_2018.2018.00205](https://doi.org/10.1109/Cybermatics_2018.2018.00205) ↑8
- [60] Sovrin Foundation, "Sovrin governance framework – Sovrin," 2019. [Online]. Available: <https://sovrin.org/library/sovrin-governance-framework/> ↑3, 19
- [61] Q. Stokkink and J. Pouwelse, "Deployment of a blockchain-based self-sovereign identity," in *2018 iThings/GreenCom/CPSCo/SmartData/Blockchain/CIT*, 2018, pp. 1336-1342. ↑19, 20
- [62] F. Wang and P. De Filippi, "Self-sovereign identity in a globalized world: Credentials-based identity systems as a driver for economic inclusion," *Front. Blockchain*, vol. 2, art. 00028, 2020. <https://doi.org/10.3389/fbloc.2019.00028> ↑4, 19
- [63] Q. Wang, M. Su, and R. Li, "Is China the world's blockchain leader? Evidence, evolution and outlook of China's blockchain research," *J. Clean. Prod.*, vol. 264, art. 121742. <https://doi.org/10.1016/j.jclepro.2020.121742> ↑12
- [64] P. Windley and D. Reed, "Sovrin: A protocol and token for self-sovereign identity and decentralized trust," 2018. [Online]. Available: <https://sovrin.org/wp-content/uploads/Sovrin-Protocol-and-Token-White-Paper.pdf> ↑3
- [65] Z. Zhao and Y. Liu, "A blockchain based identity management system considering reputation," in *2019 2nd Int. Conf. Info. Syst. Comp. Aided Ed. (ICISCAE)*, 2019, pp. 32-36. <https://doi.org/10.1109/ICISCAE48440.2019.221582> ↑9
- [66] X. Zhu and Y. Badr, "A survey on blockchain-based identity management systems for the Internet of Things," in *2018 iThings/GreenCom/CPSCo/SmartData/Blockchain/CIT*, 2018, pp. 1568-1573. ↑19
-

## Roberto Albeiro Pava Díaz

Magister en Ingeniería - Ingeniería de sistemas y computación, Universidad Nacional de Colombia; Ingeniero de Sistemas, Universidad Nacional de Colombia; Docente en Universidad Distrital Francisco José de Caldas. Facultad de Ingeniería. Carrera 7 # 40B-53 Código postal: 11021-110231588. Bogotá, Colombia.

**Email:** [rapavad@udistrital.edu.co](mailto:rapavad@udistrital.edu.co)

## Rafael Vicente Páez Méndez

Doctor en Ingeniería Telemática, Universidad Politécnica de Cataluña - España; Especialista en Seguridad de Redes Informáticas, Universidad Católica de Colombia; Ingeniero de Sistemas y Computación - Universidad Católica de Colombia - Colombia. Docente en Pontificia Universidad Javeriana. Facultad de Ingeniería – Depto. de Ingeniería de Sistemas. Carrera 7 # 40-62 Código postal: 111711. Bogotá, Colombia.

**Email:** [paez-r@javeriana.edu.co](mailto:paez-r@javeriana.edu.co)

## Luis Fernando Niño V.

Ph.D Computer Science, The University of Memphis, Estados Unidos; M. Sc. Computer Science, The University of Memphis, Estados Unidos; Magister en Matemáticas, Universidad Nacional de Colombia; Ingeniero de Sistemas, Universidad Nacional de Colombia. Docente en Universidad Nacional de Colombia - Sede Bogotá. Facultad de Ingeniería - Depto. de Ingeniería de Sistemas e Industrial. Carrera 45 # 26-85 Código postal: 11001. Bogotá, Colombia.

**Email:** [lfninov@unal.edu.co](mailto:lfninov@unal.edu.co)

