MDPI

Article

"*Who Is the FAIRest of Them All*?" Authors, Entities, and Journals Regarding *FAIR Data Principles*

Luis Corujo

Centre for Classical Studies, School of Arts and Humanities, University of Lisbon, Alameda da Universidade, 1600-214 Lisbon, Portugal; luiscorujo@campus.ul.pt

Abstract: The perceived need to improve the infrastructure supporting the re-use of scholarly data since the second decade of the 21st century led to the design of a concise number of principles and metrics, named *FAIR Data Principles*. This paper, part of an extended study, intends to identify the main authors, entities, and scientific journals linked to research conducted within the *FAIR Data Principles*. The research was developed by means of a qualitative approach, using documentary research and a constant comparison method for codification and categorization of the sampled data. The sample studied showed that most authors were located in the Netherlands, with Europe accounting for more than 70% of the number of authors considered. Most of these are researchers and work in higher education institutions. These entities can be found in most of the territorial-administrative areas under consideration, with the USA being the country with more entities and Europe being the world region where they are more numerous. The journal with more texts in the used sample was *Insights*, with 2020 being the year when more texts were published. Two of the most prominent authors present in the sample texts were located in the Netherlands, while the other two were in France and Australia.

1. Introduction

Since the second decade of the 21st century, there has been a perceived need to improve the infrastructure supporting the re-use of scholarly data. As part of this, a set of stakeholders, including representatives from academia, industry, funding agencies, and academic publishers, have designed a concise number of principles and metrics, which they named *FAIR Data Principles* [1,2]. The acronym FAIR refers to the characteristics of Location (Findable), Accessibility (Accessible), Interoperability (Interoperable), and Reusability (Reusable). These principles specifically emphasize improving the ability of machines (in this context, interpreted as digital repositories) to find and use data automatically, as well as supporting its reuse by individuals [2] (p. 1).

Within this framework, the 15 principles presented appear divided by the identified categories. In order to be findable (Findable), (1) data (and/or metadata) are assigned a globally unique and persistent identifier; (2) data are described with enriched metadata (defined by principle 10); (3) metadata clearly and explicitly include the identifier of the data it describes; (4) data (and/or metadata) are recorded or indexed in a searchable resource. In order to be accessible (Accessible) (5), the data (and/or metadata) are retrievable via their identifier using a standardized communication protocol; (5.1) the protocol is open, free, and universally implementable; (5.2) the protocol allows an authentication and authorization procedure, where necessary; and (6) the metadata are accessible, even when the data are no longer available. In order to be interoperable (Interoperable), (7) the data (and/or metadata) use a formal, accessible, shared, and widely applicable language for knowledge representation; (8) the data (and/or metadata) use vocabulary that follow

FAIR principles; and (9) the data (and/or metadata) include qualified references to other data (and/or metadata). In order to be reusable (Reusable), (10) data (and/or metadata) are richly described with a plurality of accurate and relevant attributes; (10.1) data (and/or metadata) are communicated with a clear and accessible data use license; (10.2) data (and/or metadata) are associated with detailed provenance; (10.3) data (and/or metadata) meet standards relevant to the community domain.

By considering how to spread out the support for these principles, authors such as Mons et al. [3] considered that their interpretation also multiplied because (among other reasons) they allow a great deal of freedom regarding its implementation. For this reason, they stated that the *FAIR Data Principles* should not be considered: (1) a standard; (2) equal to RDF, linked data, or the semantic web; (3) limited to how humans are able to find, access, reformat, and reuse data; (4) equal to open; (5) a life sciences hobby [3] (p. 51–52). Instead, these authors consider the principles to be crucial for data and services used for evidence-based decisions, to expose malpractices in terms of intellectual property and research, and to allow the participation of citizens from any country in different scientific research projects [3].

The variety of interpretations of what it means to be FAIR, or how FAIR is an object, was also pointed out by Higman et al. [4]. While discussing the boundaries, intersections, and overlaps between the concepts of *FAIR*, Open Data, and Research Data Management, they considered that the first two are noble aspirations and a useful way for scientists to engage in good data practices by focusing on data sharing, in order to ensure that content is made available in ways that promote access and reutilization [4]. On the other hand, Research Data Management, since its conception, has been considered to be about data stewardship [4].

Recognizing the importance of *FAIR Data Principles* in raising awareness about adopting good practices and defining necessary characteristics of data objects to ensure that data are reusable by humans and machines, several stakeholders from various segments of the digital repository practice and research community argue that in order to keep data in accordance with these principles, while preserving them in the long term, requires trusted digital repositories with sustainable governance and organizational structures, along with reliable infrastructure and comprehensive policies that support the community-agreed practices [5] (p. 1). Within this framework, NESTOR—the network of expertise in the long-term storage of digital resources in Germany—criticizes *FAIR Data Principles* with regard to the lack of systematic attention toward digital preservation [6]. Indeed, trusted digital repositories support data curation and preservation of data collections with different levels of reuse. In situations of lower quality data, which cannot be reasonably improved or made more interoperable, but which may still retain high value for its user community, reliable stewardship is required.

Research regarding initiatives, concepts, and implementation of *FAIR Data Principles* has been presented in fields such as health data stewardship [7], bioinformatics [8], biological knowledge networks [9], or focused on heritage institutions such as libraries, archives, and Museums [10]. However, there is a perceived lack of studies on the authors and profiles that participate in research focused on *FAIR Data Principles*.

The current paper is part of developing research focused on identifying and analyzing, in a comparative way, the main programs and projects regarding or making use of *FAIR Data Principles* at a worldwide level; identifying the main actors and contrasting their perceptions and interpretations of the said principles; distinguishing the proposals and solutions that emerge from the analysis of perceptions and meaning related to those. It is also intended to, by means of a trend analysis of the specific scientific literature, identify projects, initiatives, and programs of international expression on the *FAIR Data Principles*. The intended future results of this research are a critical and trend-based theoretical construction of the examined literature. This might allow us to formulate recommendations for the use of *FAIR Data Principles*, besides showing possible consensuses and dissents, uncertainties, and certainties behind what is perceived of the said principles and their uses.

In this regard, within the scope of the cutting-edge international scientific literature and by means of document analysis of papers regarding *FAIR Data Principles*, dating from 2016 onward, collected in three scientific databases, the study presented in this paper specifically intended to identify the main authors, entities, and scientific journals linked to research emphasizing the *FAIR Data Principles* since its inception. For such aim, it defined questions concerning: the authors, their locations and professional occupation; the organizations/entities where the authors perform their professional and research activities, including the place where such entities are based; the journals where and when the authors' research was published; the most prominent authors in terms of authorship. The research questions can be summarized as follows:

1. Who is writing about *FAIR Data Principles*?
2. Where are they based?
3. What institutions and countries are leading the effort?
4. In which scientific journals are they publishing their research?

The study allowed us to establish the landscape in which research dedicated to, or adopting, the FAIR Principles takes place. This considers both theoretical research and exercises of the practical application of the said principles.

Within this scope, the study of the researchers' profiles revealed the original contexts of those who developed this type of research.

We established the year 2016 and later as the chronological boundaries for collecting the sample of scientific papers used for the document analysis since that was the year when the "FAIR Guiding Principles for scientific data management and stewardship" were published [2]. We chose the LISA—Library, Information Science Abstracts; LISTA—Library, Information Science & Technology Abstracts; and Scopus databases because we were given access to them. Only peer-reviewed documents with full text were considered so that they could be used in future expansions of this research. Future research will allow us to collect and analyze an expanded number of information sources to build a structured information *corpus* that will be further analyzed and interpreted.

Although delimited by the sample, this allows us to identify trends that, in further research, may be contrasted with data collected from a larger number of sources and with broader research criteria.

It should be noted that it is not the focus of this paper to make a thematic distinction between the texts in the sample or a typological delimitation as to the types of research presented in those texts.

2. Materials and Methods

We developed a study with a qualitative approach, using documentary research, which is considered a systematic procedure for reviewing or assessing documentary material with textual or image information recorded without human intervention. The researcher then uses the examination and interpretation of data to extract meaning, gain understanding and develop empirical knowledge [11]. This analytic procedure involves identifying, selecting, evaluating, and synthesizing the documents' data. Document analysis produces data—excerpts, quotes, or entire passages—which are then organized into main themes, categories, and case examples specifically through content analysis [11].

This documentary research provided the means to establish the contextualization necessary to outline the problem and define the *FAIR Data Principles* scenario, and, on the other hand, it helped to develop the approach to that issue within the international scientific literature.

We then developed a systematization proposal, delimited by the data collection, and elaborated upon the categories collected and identified through the constant comparison method (CCM). CCM is a qualitative analysis approach developed within the grounded theory methodology, which focuses on the comparison of and between all the data elements, which can be identified as a procedure for interpreting texts, through coding and

analysis, in order to develop theory [12,13] (p. 437). This method has four phases: (1) the comparison of incidents applicable to each category, (2) the integration of categories and their properties, (3) the delimitation of theory, and (4) the writing of theory [12] (p. 51–53) [13] (p.439–443). The analysis, which was carried out on the texts collected through the CCM, was developed by means of a back-and-forth process of coding, categorization, and saturation of the data on which this work is based, which constitutes a progressive spiral whose reflections allowed us to verify the main authors, entities, and scientific journals linked to the research conducted, with a special emphasis on the *FAIR Data Principles* since its inception. The use of the CCM will allow proceeding with the future analysis and interpretation of the information *corpus* resulting from expanded data gathering to be performed by the use of the grounded theory method.

The constructed model enabled the content analysis of the collected texts' main themes, originating a set of memoranda per category, which this study identifies. We sought, specifically, to reveal an invisible dimension, as explained by Bardin: "From the moment the content analysis decides to codify its material, it must produce a system of categories. The categorization has as its first objective (in the same way as the documentary analysis) to provide, by condensation, a simplified representation of the raw data. (…) Content analysis implicitly rests on the belief that categorization (the passage from raw data to organized data) does not introduce deviations (by excess or by refusal) in the material, but that it makes invisible indices known, at the level of the raw data" [14] (p. 148–149, our translation).

On 5 April 2022, research at LISA—Library, Information Science Abstracts; LISTA—Library, Information Science & Technology Abstracts; and Scopus databases were carried out in order to collect bibliographic references containing the terms "FAIR Data Principles". This delimitation was applied due to the acronym "FAIR" being easily confused for the word "fair". For that matter, it seems easy to find papers that use the concept of "fair principles" but are not concerned with "*FAIR Data Principles*". Nevertheless, future research will take that into account, especially the uses of the concept of "fair principles" that are especially concerned with *FAIR Data Principles*.

As stated before, the search strategy defined for collecting the sample of scientific papers used for the document analysis considered the year 2016 and later as the chronological boundaries since that was the year when the "FAIR Guiding Principles for scientific data management and stewardship" were published [2]. The three databases were chosen because the author has authorized access to them. Only peer-reviewed documents with full text were considered so that they could be used in future work expanding this research.

Table 1 identifies the following search strategies.

Table 1. Databases, search strategies, and number of results.

Database	Search Strategy	Number of Results
LISA	"FAIR Data Principles" AND PEER (yes) AND pd (>20160101)	24
LISTA	TX "FAIR Data Principles"; Expanders: Also search within the full text of the articles; Apply equivalent subjects; Limiters: Full Text; Peer Reviewed; Publication Date: 20160101–20221231; Source Types: All Results	5
SCOPUS	TX "FAIR Data Principles"; Expanders: Also search within the full text of the articles; Apply equivalent subjects; Limiters: Full Text; Peer Reviewed; Publication Date: 20160101–20221231; Source Types: All Results	12
	Total	41

A total of 41 articles were retrieved. The cleaning of duplicates (3) and the exclusion of articles that were not in English (5) were made. After reading the abstracts, one paper was excluded on account of the full text being behind a paywall. This resulted in 32 articles.

Table A1 in the Appendix A presents the papers that were retrieved and analyzed regarding the present research.

It is reiterated that it was not the focus of this particular paper to make a thematic distinction of the texts in the sample or a typological delimitation as to the types of research presented in them. Therefore, in the collection and selection of texts, as well as in the content analysis, there was no distinction or typological or thematic analysis of the research presented in the studied texts.

As mentioned before, the analysis was performed by means of CCM, a qualitative analysis approach. During this analysis, incidents in the data were compared for the creation of pre-textual codes, such as the names of authors, their professional occupations, the countries and institutions where they were based, the scientific journals that published the sampled texts, and the year in which they were published, in a back-and-forth process. These codes led to the identification of categories (which are more abstract codes) that led to further analysis of the collected texts to saturate the data regarding the professional occupation and the entity type.

Although this effort is presented very briefly, it constitutes the central task of this research, both for the time it took to be executed and for allowing the verification of which main authors, entities, and scientific journals were linked to the research conducted and for defining the organizational logic of the data that are reflected in the descriptions provided in the results and discussion section.

Since this paper is part of qualitative research, the results produced at the current stage must be confirmed or tested by further research, gathering data from other databases and information sources to produce sound theoretical assumptions and as a form of quality control. In order to help the validation and quality control of this research, memoranda will be produced during the future analysis and interpretation of the information *corpus* resulting from expanded data gathering as a methodological tool of the grounded theory method.

3. Results and Discussion

The papers from the sample were analyzed, and the data were collected and codified. The resulting codes from this process allowed for a more abstract codification, which permitted the development of categories. These categories revealed the need to saturate the properties of each of these categories, which led to the search for information in other sources (such as academic and professional platforms, personal or institutional, and journal websites). This information allowed us to establish information about the authors, entities, and journals and structure the presentation of what emerged from the analysis, as shown below.

3.1. Distribution of Authors by Country

The data on the provenance of the authors and their professional occupation were extracted from analyzed texts that were accompanied by a biographical note, from the authors' biographical notes that were available on the publications' websites and/or websites of the professional associations that publish these publications, from the curricula vitae available on the websites of the higher education institutions where they carry out their academic activity, and from the curricula vitae available on the ORCID and LinkedIn platforms. Unless otherwise stated, the information refers to the moment of authorship of the text(s). The provenance considers the location of the authors at the time of their publications.

From the sample of thirty-three texts used for this part of the research, 128 authors were identified.

The authors' distribution data by country at the time of authorship of the texts are shown in Figure 1 and Table A2 (in the Appendix A).
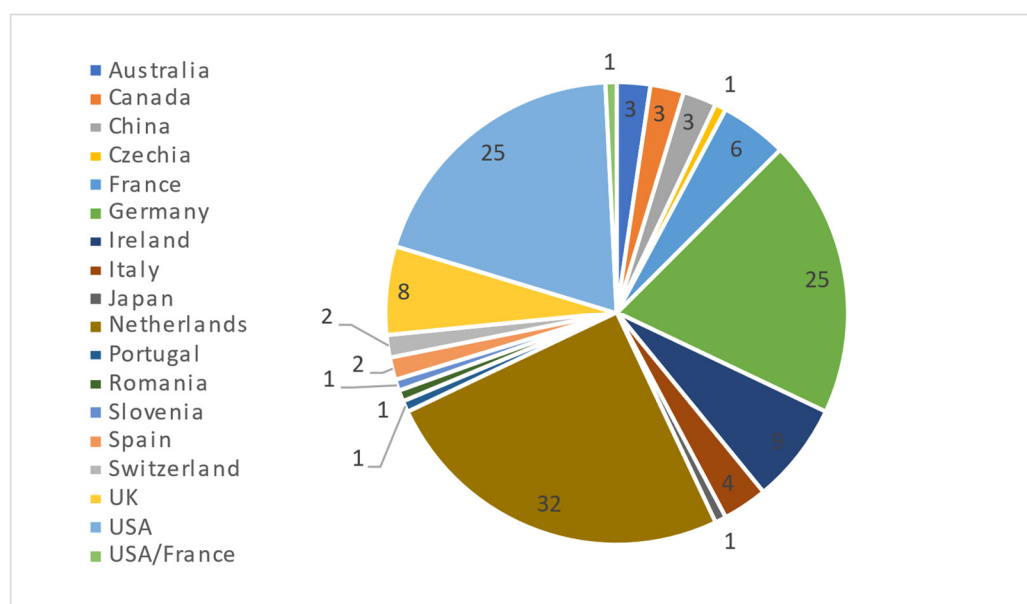
Figure 1. Distribution of Authors by Country.

It is noticeable that the majority of authors are from the Netherlands (25%, a quarter of the authors), followed *ex aequo* by Germany and the United States (19.53%, almost a fifth of the authors), Ireland (7.03%, with nine authors), and the United Kingdom (6.25%, with eight authors). The list of countries with more than one author also includes France, Italy, Australia, Canada, China, Spain, and Switzerland. It should be noted that an author from the Netherlands was at the time working in the VASCERN European Reference Centre, an International Reference Network (Leo Schultze Kool). In the specific situations of the countries where only one text was published by a single author, we indicate a Czech university lecturer (Lenka Kourimska), a Japanese researcher (Kai Nishikawa), a Portuguese researcher (Isabel Castanheira), a Romanian researcher (Nastasia Belc), a Slovenian university lecturer and researcher (Nives Ogrinc) and a PhD student at two higher education institutions in the United States and France (Coline Ferrant).

3.2. Professional Occupation

The coding of the authors' professional occupation resulted in a set of categories, which include CEO, independent consultant, information professionals, IT professionals, management, professors, researchers, and students. The independent consultant category includes senior analyst contractors and science publishers who work independently. Information professionals (library and information science—LIS) refers to employees, directors, and coordinators who work as librarians, data stewards, data management consultants, research data services professionals, data librarians, repository managers, and open access service managers. IT professionals refer to any computer or software engineer and/or specialist, technical data services professional, software developer, scientific and technical officer (STO)—data management, and data analyst/coordinator. Management refers to professionals who perform activities such as communication manager, director of partnerships, director of biosciences, and manager of the technology's programs, mainly in a corporate context. Professors exercise, in some way, teaching activities in higher education institutions, while researchers are those who develop professional activities in the field of research. Finally, students refer to those individuals who were undertaking their PhD.

We also found that some authors accumulate more than one professional role, which leads to the development of specific categories for those cases.

Despite being attached to an organization (in this case, the Marine Institute, an Irish Governmental Agency), authors such as Caoimhín Kelly defined themselves as contractors and were therefore coded as belonging to the category of independent consultants.

The data concerning the professional occupation of the authors can be found in Figure 2 and Table A3 (in the Appendix A).
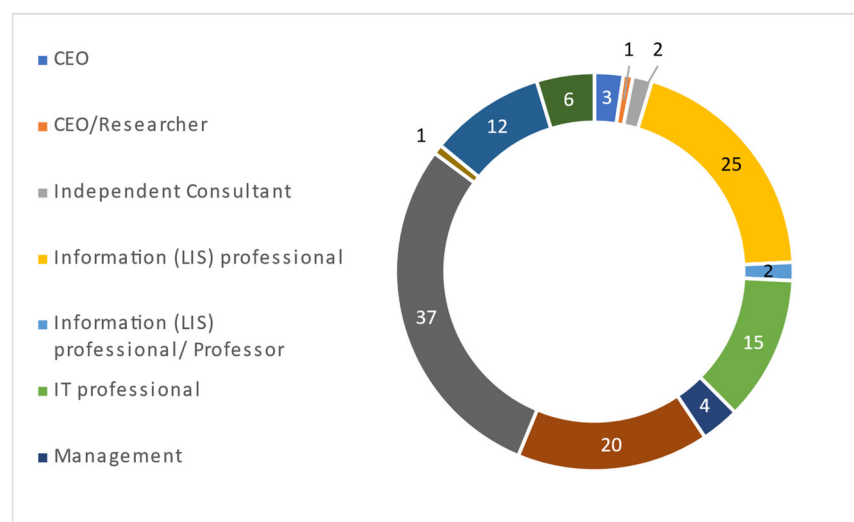

Figure 2. Authors' Professional Occupation.

These data allowed us to verify that 51 authors carried out research activities (39.85%), 12 of which were also professors (9.38%), one (Karl Presser) was a CEO, and another (Martijn Kersloot) a manager (0.78% each). However, within the academic universe, there were 34 higher education professors, corresponding to more than a quarter of the authors (26.56%), where, in addition to the previously indicated 12 authors who combined research activities (9.38% previously indicated), there were two who also worked as information professionals, Ayla Stein Kenfield and John J. Meier (1.56%). There were 20 full-time professors (15.62%). The existence of six students (4.69%) is also worth mentioning in this context.

More than a fifth of the authors (21.09%) were information (LIS) professionals. Apart from the two authors who combined such functions with university professorship (1.56%), 25 were exclusively involved in LIS activities (19.53%).

The fifteen IT professionals corresponded to more than a tenth of the authors (11.72%).

There were five authors with management activities (3.91%), four of whom—Heather Staines, Linda van den Berg, Maryann E Martone and Merlijn N. van Rijswijk—were full-time managers (3.13%). Likewise, the authors with CEO functions corresponded to four (3.13%): three in exclusive, Dominic Farace, Tiberius Ignat, and Wolfgang Colsman (2.34%); and Karl Presser, who already indicated that he also had a part in research activities (0.78%).

Finally, two independent consultants (1.56%), Caoimhín Kelly and Jan Velterop, were mentioned.

3.3. Distribution of Authors by Organizations and Entities

It is also important to know where the authors work and which type of organization. This information is provided in Table A4 in the Appendix A.

The coding process regarding the organizations where the authors carried out their activities required the identification of the parent entities where such organizations were integrated. There were many situations in which the entities were public, namely higher education organizations and research centers. However, it was decided to identify as governmental agencies only those organizations whose mandate specifically identified them

as such. Furthermore, it was deemed necessary to separate independent organizations from private organizations, given that the former refers to non-profit entities and the latter is more associated with the corporate world. Moreover, international reference networks are organizations created or funded by international entities to gather resources for research purposes (in this case, the VASCERN European Reference Centre).

Despite being attached to an organization (in this case, the Marine Institute, an Irish Governmental Agency), we reinforce that authors such as Caoimhín Kelly defined themselves as contractors and were therefore coded as belonging to the category of an independent consultant.

In the data presented, it should be noted that Barend Mons was a lecturer at three entities, namely Leiden University, Dutch Techcentre for Life Sciences, and The Netherlands eScience Centre; Coline Ferrant was a PhD student at Northwestern University (USA) and at the Paris Institute of Political Studies (Sciences Po, in France); Hélène Prost was an IT professional (Information engineer) at the University of Lille III and at the Centre National de la Recherche Scientifique, both in France; Jaap Heringa was a researcher at the Dutch Techcentre for Life Sciences and a professor at the Vrije Universiteit (VU) Amsterdam, both in the Netherlands; Karl Presser was a researcher at the Swiss Federal Institute of Technology and CEO of Premotec GmbH, both in Switzerland; Leo Schultze Kool was a professor at the VASCERN European Reference Centre and Radboud University (the Netherlands); Luiz Olavo Bonino da Silva Santos was a professor at the Vrije Universiteit (VU) Amsterdam and at the Dutch Techcentre for Life Sciences, both in the Netherlands; Marc Rittberger was a professor at the Darmstadt University of Applied Sciences and a researcher at the Leibniz Institute for Research and Information in Education, both in Germany; Marco Roos was a professor at Leiden University and researcher at the Dutch Techcentre for Life Sciences, both in the Netherlands; Martijn Kersloot was a researcher at Vrije Universiteit (VU) Amsterdam, and a Manager (Product Owner Data and Innovation) at Castor EDC, Amsterdam, both in the Netherlands; and Renaud Fabre was a professor at the University of Paris VIII and a researcher at the Centre National de la Recherche Scientifique.

Figure 3 provides information regarding the number of authors per type of entity.
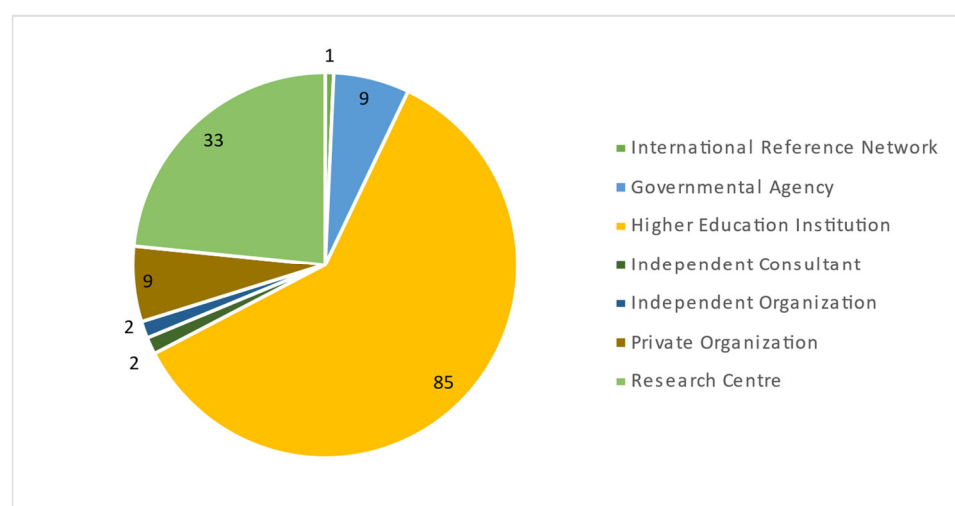
Figure 3. Number of Authors per Type of Entity.

Almost two-thirds of the authors were working in higher education institutions (60.28%), and slightly less than a quarter were working in research centers (23.40%). The entities with more referenced authors were, *ex aequo*, a higher education institution and a research center, respectively, the Delft University of Technology and the Dutch Techcentre for Life Sciences (6.38% of the referred authors each), both from the Netherlands. Government agencies and private organizations had nine authors each (6.38%), and

independent organizations had two authors (1.42%). Two authors identified themselves as an independent consultant (1.42%). Only one type of entity has only one author, corresponding to the international reference networks (VASCERN European Reference Centre).

The information on the 68 entities where the addressed authors develop their activities, plus two independent consultants (Jan Velterop and Caoimhín Kelly), can be found in Table A5 (in the Appendix A) and Figures 4 and 5. However, one must keep in mind that eleven authors developed professional or study activities in more than one entity (Barend Mons in three entities, and Coline Ferrant, Hélène Prost, Jaap Heringa, Karl Presser, Leo Schultze Kool, Luiz Olavo Bonino da Silva Santos, Marc Rittberger, Marco Roos, Martijn Kersloot, Renaud Fabre).

Figure 4 presents the number of organizations per type of entity.
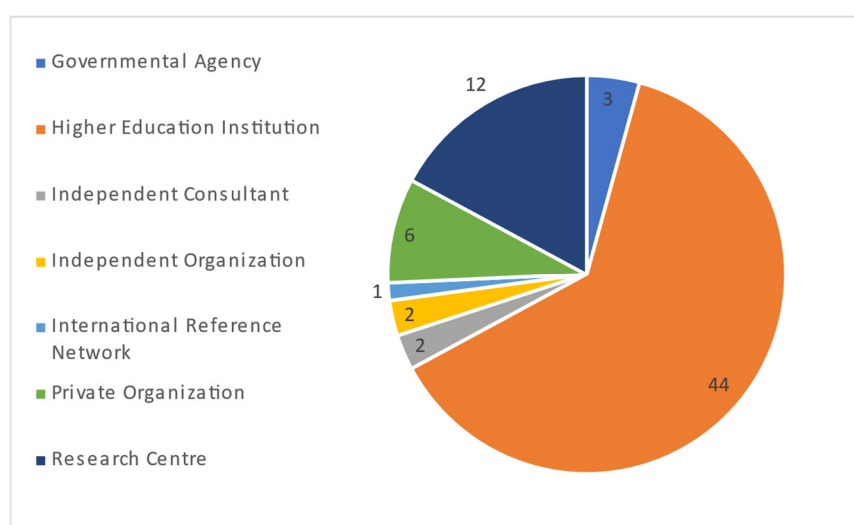
Figure 4. Number of Organizations per Type of Entity.
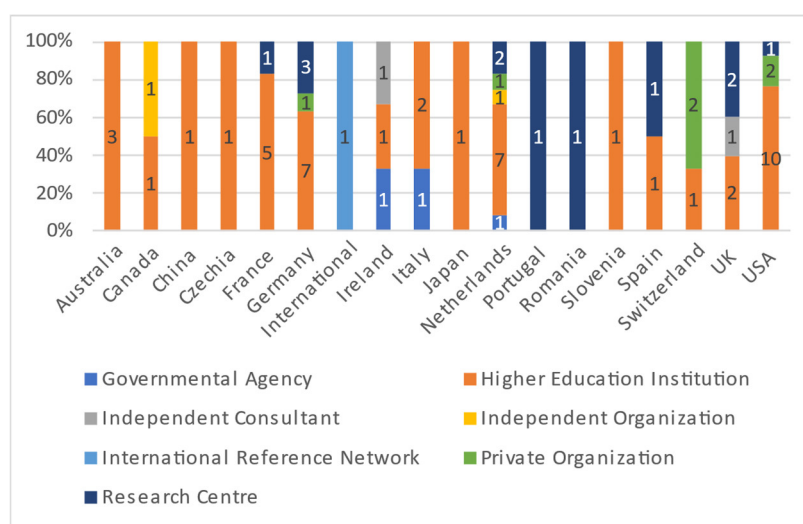
Figure 5. Distribution of Entity Types per territorial-administrative areas.

These data show that almost two-thirds of the organizations are higher education institutions (62.86%), with research centers accounting for almost one-fifth of the organizations surveyed (17.14%). Private organizations represent less than one-tenth of the entities identified (8.57%). Government agencies include three organizations (4.29%). Two types of entities include two organizations or refer to two persons (the independent

organizations and the independent consultant), and one type of entity refers to one organization (international reference network).

Figure 5 presents the distribution of entity types per territorial-administrative area (countries and the European Union).

This shows us that higher education institutions are also those that cover most of the territorial-administrative areas under consideration and are only not represented in the sample at the international level and by countries such as Portugal and Romania. They are followed by the research centers, which limit their representation within the sample studied to France, Germany, the Netherlands, Portugal, Slovenia, the United Kingdom, and the United States.

Only one type of entity is represented in only one country or territorial-administrative unit: the International Reference Network in the European Union.

Most of the entities retrieved from the sample belong to the United States (18.57%, just under a fifth of the entities), followed by the Netherlands (17.14%), Germany (15.71%, just over a sixth), France (8.57%, less than a tenth), and the United Kingdom (7.14%). Four countries have three entities listed (Australia, Ireland, Italy, and Switzerland), two countries have two entities listed (Canada and Spain), and seven countries or territorial entities have only one of the entities identified (China, Czechia, European Union, Japan, Portugal, Romania, and Slovenia).

Almost one-sixth of the organizations surveyed were higher education institutions from the United States (14.29%), followed by German and Dutch higher education institutions, which each account for one-tenth of the organizations surveyed (10% each), French higher education institutions (7.14%), and Australian higher education institutions (4.29%), together with German research centers (also 4.29%). Overall, there is a preponderance of higher education institutions from the European Union, which represent more than a third (35.71%) of the institutions surveyed, while this type of institution from the English-speaking countries represents less than a quarter of the total number of institutions (24.3%).

The Netherlands had the broadest coverage in terms of entity types since it does not only include independent consultants and international reference networks. It was followed in this respect by Germany, Ireland, the United Kingdom, and the United States, each with three different types of entity. Canada, France, Italy, Spain, and Switzerland each had two types of entities. Finally, Australia, China, Czechia, the European Union, Japan, Portugal, Romania, and Slovenia had only one type of entity in the sample analyzed.

3.4. Distribution of Papers by Journals and Years

The information on the distribution of analyzed texts by scientific journals is presented in Figure 6 and Table A6 (in the Appendix A).
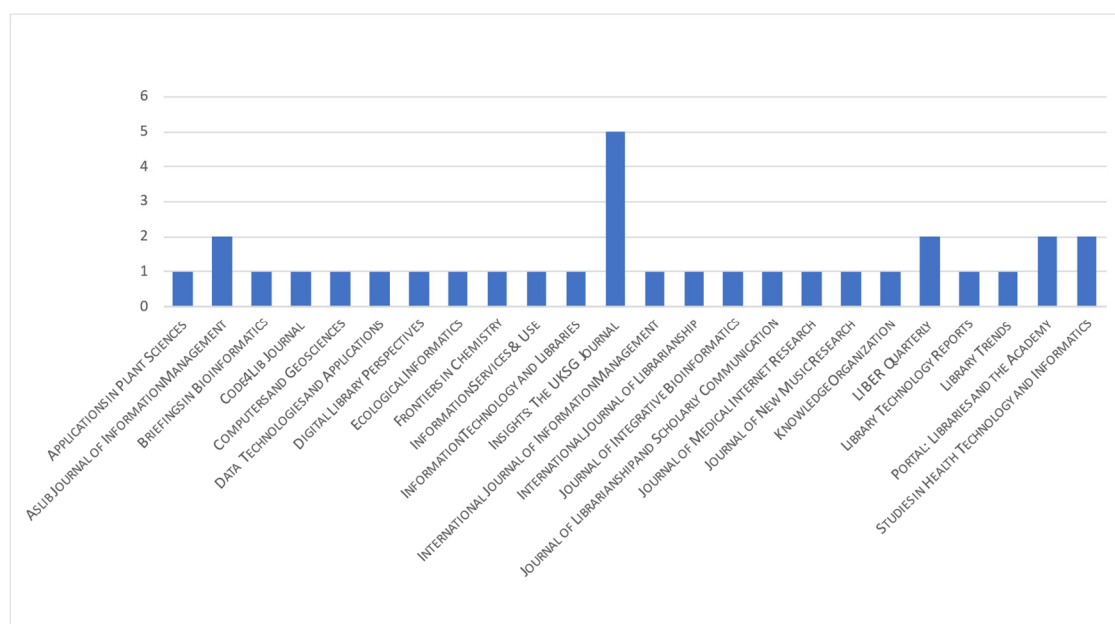
Figure 6. Distribution of Scientific Papers by Journals.

The data show that most of the scientific papers used in this study were published in *Insights: The UKSG Journal* (15.63%), with this scientific journal being followed *ex aequo* by the *Aslib Journal of Information Management*, *LIBER Quarterly*, *Portal: Libraries and the Academy*, and *Studies in Health Technology and Informatics* (6.25% each).

The journals from which only one article originates are as follows: *Applications in Plant Sciences; Briefings in Bioinformatics; Code4Lib Journal; Computers and Geosciences; Data Technologies and Applications; Digital Library Perspectives; Ecological Informatics; Frontiers in Chemistry; Information Services & Use; Information Technology and Libraries; International Journal of Information Management; International Journal of Librarianship; Journal of Integrative Bioinformatics; Journal of Librarianship and Scholarly Communication; Journal of Medical Internet Research; Journal of New Music Research; Knowledge Organization; Library Technology Reports;* and *Library Trends*.

Two journals account for a single author in the sample (*International Journal of Information Management and Library Technology Reports*).

The chronological distribution of the scientific texts in the sample is shown in Table 2 and Figure 7.

Table 2. Distribution of analyzed Papers by Year.

Years	Number of Texts	% Texts
2017	2	6.25%
2018	9	28.13%
2019	7	21.88%
2020	12	37.5%
2021	2	6.25%
Total	**32**	**100%**
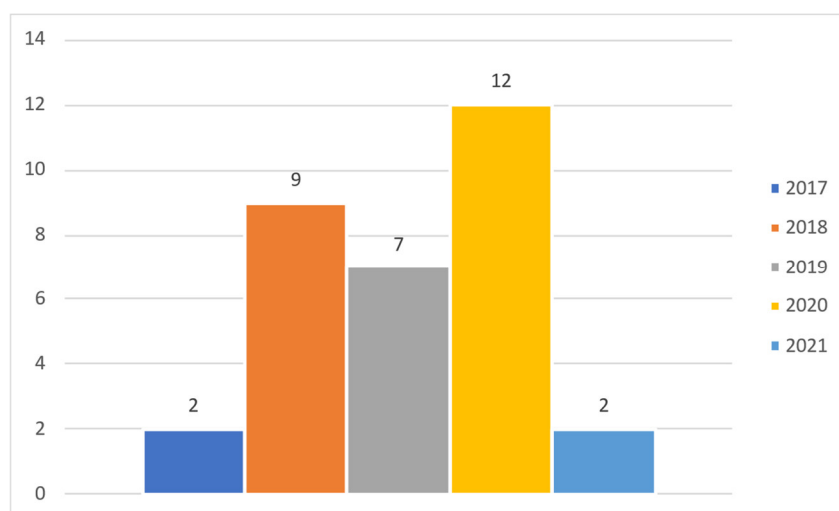
Figure 7. Distribution of Scientific Papers by Year.

This information allows us to note that, within the defined chronological delimitation, only in 2016 and 2022 (up to 5 April, when the research and collection of texts were carried out) were there no results. Thus, the years between 2017 and 2021 are represented by at least two of the sample texts.

2020 presents itself as the year in which a greater number of texts were published (37.5%, more than a third of the sample), followed by the number of texts from 2018 (28.13%, more than a quarter of the texts), followed by the number of texts from 2019 (21.88%, more than a fifth of the texts). The years 2017 and 2021 hold, *ex aequo*, fewer texts in this sample (only two each).

3.5. Distribution of Authors by Journal and Place of Origin

Figure 8 and Table A7 (in the Appendix A) show the data concerning the authors covered in the universe of publications used in this part of the study. Figure 8 is reproduced on a greater scale in Figure A1, in the Appendix B, for better visualization.
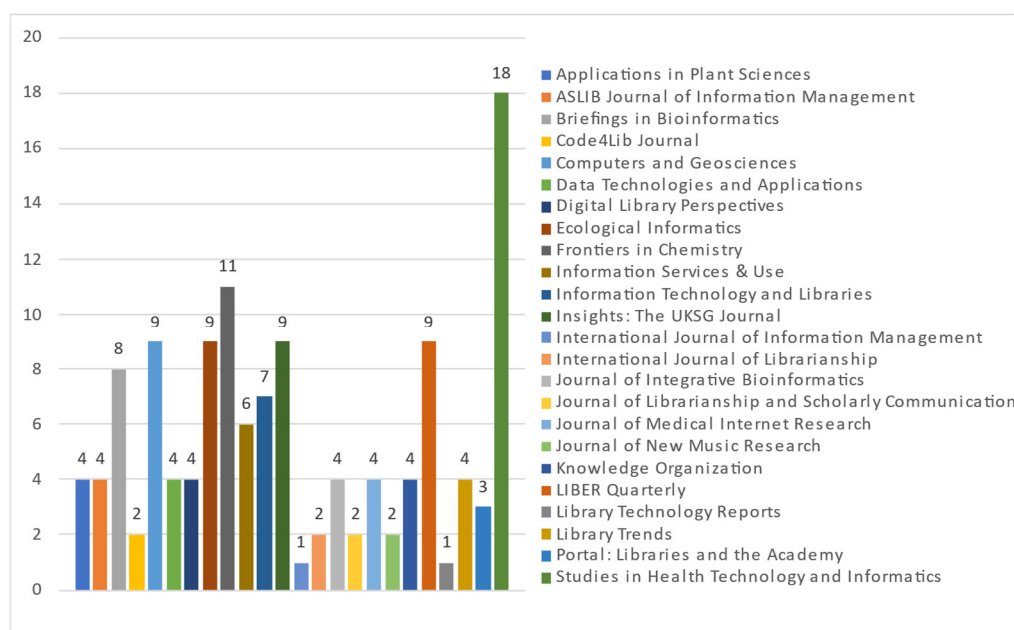
Figure 8. Number of Authors per Journal.

These data show that *Studies in Health Technology and Informatics* has the most authors among the publications used in this part of the study (13.74%), followed by *Frontiers in Chemistry* (8.4%), followed by *Computers and Geosciences, Ecological Informatics, Insights: The UKSG Journal*, and *LIBER Quarterly* (6.87%, each).

The publication *Briefings in Bioinformatics* results in 6.11%, *Information Technology and Libraries* in 5.34%, and *Information Services & Use* in 4.58%. Eight publications have four authors each: *Applications in Plant Sciences*; *ASLIB Journal of Information Management*; *Data Technologies and Applications*; *Digital Library Perspectives*; *Journal of Integrative Bioinformatics*; *Journal of Medical Internet Research*; *Knowledge Organization*; and *Library Trends*.

The publication *Portal: Libraries and the Academy* refers to 2.29% of authors, while *Code4Lib Journal, International Journal of Librarianship, Journal of Librarianship and Scholarly Communication*, and *Journal of New Music Research* have two authors.

The journals with a single author are the *International Journal of Information Management* (corresponding to Matthew I. Bellgard) and *Library Technology Reports* (Bohyun Kim).

These figures have to be weighted with the fact that some authors have publications, used in this part of the study, in more than one journal, such as Barend Mons, with publications in *Briefings in Bioinformatics* and *Information Services & Use*; Cameron Neylon, who has published in *Information Services & Use* and *Insights: The UKSG Journal*; and Joachim Schöpfel, with publications in *Knowledge Organization* and *Data Technologies and Applications*.

Information regarding journals by countries of origin of the authors of the articles is provided in Figure 9 and Table A8 in the Appendix A. Figure 9 is reproduced on a greater scale in Figure A2, in the Appendix B, for better visualization. We underlined that the provenance/geographical distribution of the authors considers the location of the authors at the time of their publications. The study does not assume their nationality.
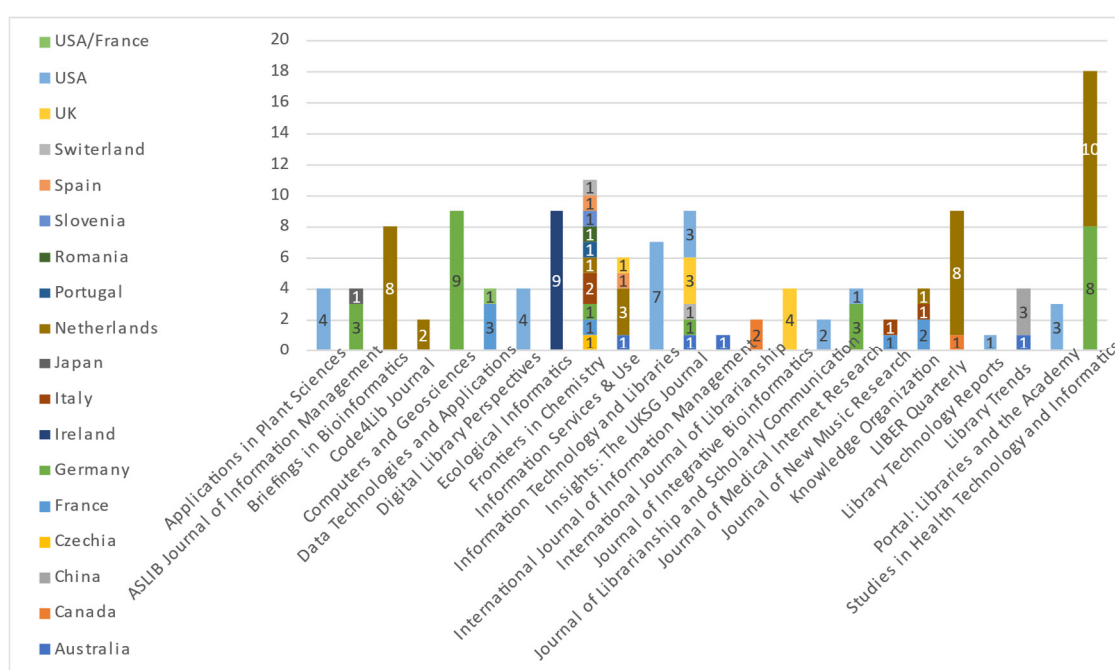
Figure 9. Geographical Distribution of Authors by Journal.

This information shows that most of the articles from *Studies in Health Technology and Informatics* in the studied sample were written by authors located in the Netherlands (7.63%), including one article whose authorship was shared between eight authors located in Germany (Christian-Alexander Behrendt, Dennis Kadioglu, Fatlume Sadiku, Frank Übert, Holger Storf, Jannik Shaaf, Jens Goebel, and Thomas O.F. Wagner) and two from the Netherlands (David van Enckevort and Marco Roos).

The journal *Frontiers in Chemistry* appears in the sample with only one text whose authorship is shared by the greatest number of authors in the sample, in addition to being the one with the greatest number of authors from different geographical provenance. Thus, it has two authors located in Italy (Claudia Zoani and Giovanna Zappa) (1.53%), in addition to one author located in each of these countries: Czechia (Lenka Kourimska); France (Olivier F.X. Donard); Germany (Michael Rychlik); The Netherlands (Marga C. Ocké); Portugal (Isabel Castanheira); Romania (Nastasia Belc); Slovenia (Nives Ogrinc); Spain (Larraitz Añorga); and Switzerland (Karl Presser).

The journals *Computers and Geosciences* and *Ecological Informatics* share the largest number of authors from the same country (6.87% each), considering that each one is represented in the studied sample by a scientific paper. The authors who published in the first journal were located in Germany (Carsten Hoffmann, Kristian Senkler, M.A. Muqit Zoarder, Markus Stecker, Nikolai Svoboda, Philipp Gärtner, Udo Einspanier, Uwe Heinrich, and Xenia Specka), while those of the second journal was located in Ireland (Adam Leadbetter, Andrew Conway, Caoimhín Kelly, Deirdre Brophy, Elizabeth Tray, Elvira de Eyto, Niall Ó Maoiléidigh, Siobhan Moran, Will Meaney).

The journal *LIBER Quarterly* is represented in the sample by two papers, one authored by eight authors (6.11%) from the Netherlands (Esther Plomp, Heather Andrews Mancilla, Jasper van Dijck, Kees den Heijer, Marta Teperek, Robbert Eggermont, Shalini Kurapati, and Yasemin Turkyilmaz-van der Velden), and the other by an author residing in Canada (David Wilcox).

Another journal that corresponds to nine authors who published texts from the sample is *Insights: The UKSG Journal*. From this journal, the sample presents two papers published each by an author, where one comes from the United States (William H Walters) and another from Australia (Cameron Neylon). It also presents two articles, each published by two authors, where one is authored by two authors from the United States (Heather Staines and Maryann E Martone), and the other has shared authorship by authors located in the United Kingdom (Paul Ayris) and Switzerland (Tiberius Ignat). It also presents another paper that shares authorship between two authors from the United Kingdom (Rosie Higman and Sarah Jones) and one from Germany (Daniel Bangert).

Regarding the other journals in the sample, seven other journals were identified as sharing authors from different countries. Thus, *Information Services & Use* has only one text written by three authors located in the Netherlands (Barend Mons, Luiz Olavo Bonino da Silva Santos, Michel Dumontier), one from Australia (Cameron Neylon), one resident in Spain (Mark D. Wilkinson), and another from the United Kingdom (Jan Velterop). *Knowledge Organization* presents a text by two authors from France (Hélène Prost and Joachim Schöpfel), one from Italy (Antonella Zane), and one from the Netherlands (Dominic Farace). The *ASLIB Journal of Information Management* presents a text by three authors located in Germany (Christoph Schindler, Julian Hocker and Marc Rittberger) and another text by an author located in Japan (Kai Nishikawa). From *Data Technologies and Applications* comes a text published by three authors living in France (Francis Andre, Joachim Schöpfel, and Renaud Fabre) and a PhD student who shares his location between France and the United States (Coline Ferrant). The *Journal of Medical Internet Research* is represented by one text from three authors located in Germany (Atinkut Alamirrew Zeleke, Dagmar Waltemath, and Esther Thea Inau) and one from the United States (Jean Sack). *Library Trends* also features a text shared by three authors from China (Jie Hu, Jilong Zhang, and Shenqin Yin) and one located in Australia (Menghao Jia). Moreover, the *Journal of New Music Research* presents a text by an author located in France (Francesca Frontini) and another one in Italy (Silvia Calamai).

In addition to these, there are journals with texts in the sample shared only by authors located in the same country. The United States presents the largest number of scientific journals that, in the sample, appear with texts published by authors located in that country. Within this scope, *Information Technology and Libraries* present a text by seven authors located in the same country (Guillaume Viger, Joseph P. Ferrie, Kristi Holmes, Lisa

O'Keefe, Matthew B. Carson, Norrina B. Allen, and Sara Gonzales); *Applications in Plant Sciences*, with a text by four authors (Anna K. Monfils, Edward E. Gilbert, Michael W. Belitz, Rachel A. Hackett); *Digital Library Perspectives*, with another text by four authors (Ben Welker, Brian Rennick, Dennis Della Corte, and Wolfgang Colsman); *Portal: Libraries and the Academy*, with a text by two authors (Bradley Wade Bishop and Rose M. Borden) and another by one author (Ayla Stein Kenfield); the *Journal of Librarianship and Scholarly Communications*, with a text by two authors (Elise Gowen and John J. Meier); and *Library Technology Reports* with a text by Bohyun Kim.

Next are the two journals with texts from the sample whose authorship is attributed to authors located in the Netherlands: *Briefings in Bioinformatics* presents a text with eight authors (Barend Mons, Celia W. G. van Gelder, Jaap Heringa, Linda van den Berg, Marcel Reinders, Merlijn N. van Rijswijk, Rob W. W. Hooft, and Ruben G. Kok); and *Code4Lib Journal*, with a text by Lukas Koster and Saskia Woutersen-Windhouwer.

Finally, the *Journal of Integrative Bioinformatics* features a text by four authors located in the United Kingdom (Ajit Singh, Christopher Rawlings, Keywan Hassani-Pak, and Marco Brandizi), the *International Journal of Librarianship*, with a text by two authors located in Canada (Guoying Liu and Kristi Thompson), and the *International Journal of Information Management*, with a text by an author from Australia (Matthew I. Bellgard).

We reiterate that these figures have to be weighted into consideration the fact that some authors have publications used in this part of the study in more than one journal, as is the case of Barend Mons, with publications in *Briefings in Bioinformatics* and *Information Services & Use*; Cameron Neylon, who has publications in *Information Services & Use* and *Insights: The UKSG Journal*; and Joachim Schöpfel, with publications in *Knowledge Organization* and *Data Technologies and Applications*.

3.6. Most Frequent Authors: Production and Profile

The information about the authors with more than one scientific paper in the used sample in this study is shown in Table 3.

Table 3. Authors with more than one paper in the used sample.

Author	Paper Name	Year	Journal Name
Barend Mons	*Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud*	2017	*Information Services & Use*
	Bioinformatics in the Netherlands: The value of a nationwide community	2019	*Briefings in Bioinformatics*
Cameron Neylon	*Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud*	2017	*Information Services & Use*
	Social infrastructures in research communication: a personal view of the FORCE11 story	2018	*Insights: The UKSG Journal*
Joachim Schöpfel	*Research data management in the French National Research Center (CNRS)*	2018	*Data Technologies and Applications*
	Data Papers as a New Form of Knowledge Organization in the Field of Research Data	2019	*Knowledge Organization*
Marco Roos	*OSSE Goes FAIR-Implementation of the FAIR Data Principles for an Open-Source Registry for Rare Diseases*	2018	*Studies in Health Technology and Informatics*
	Applying the FAIR data principles to the Registry of vascular anomalies (VASCA)	2020	

The data allow us to perceive that the authors with more scientific papers in the sample used in this study are Barend Mons, Cameron Neylon, Joachim Schöpfel, and Marco Roos, who appear with two articles each. From this group of authors, it can be noted that

Barend Mons and Cameron Neylon share the same article, which has a total of six authors. The other paper with Barend Mons' participation has a total of eight authors, while Cameron Neylon is the sole author of her second paper. Each of Joachim Schöpfel's articles has a total of four authors, while Marco Roos accounts for one article written by ten authors and another by nine authors. Moreover, Marco Roos' papers were published in the same journal (*Studies in Health Technology and Informatics*).

Figure 10 presents the number of scientific papers from the sample, distributed by the number of authors.
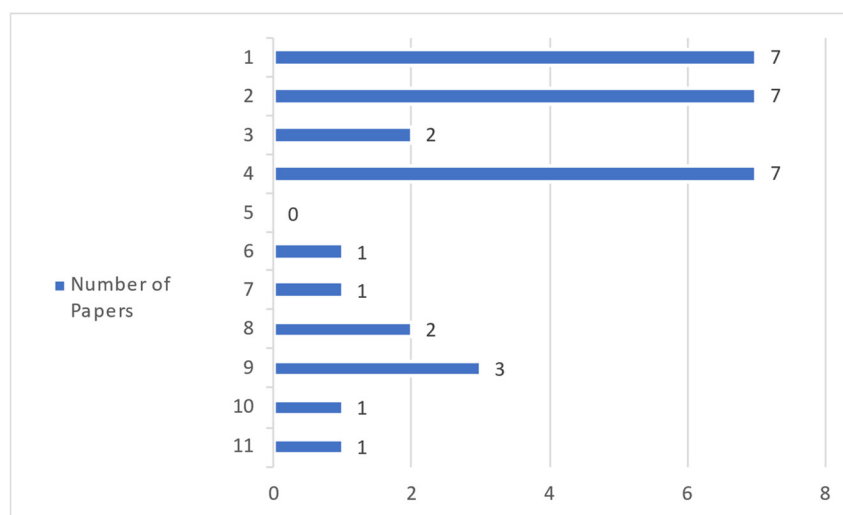
Figure 10. Number of scientific papers by number of authors.

Overall, the sample shows one paper with eleven authors [15], one with ten authors [16], three with nine authors [17–19], two with eight authors [8,20], one with seven authors [21], one with six authors [3], seven with four authors [7,9,22–26], two with three authors [4,27], seven with two authors [10,28–33], and seven with one author (referring to the following authors: Cameron Neylon [1]; Bohyun Kim [34]; William H Walters [35]; Matthew I. Bellgard [36]; Kai Nishikawa [37]; Ayla Stein Kenfield [38]; and David Wilcox [39]).

The research resulted in Barend Mons [40] obtaining a PhD in molecular biology from Leiden University in 1986. His research focuses on malaria, in close collaboration with endemic countries, and computer-assisted knowledge discovery. He was part (as an expert) of the INCO-DC European Commission program (1993–1996) and the Netherlands Organisation for Scientific Research (NWO 1966–1999). The author also co-founded several spin-off companies, such as the Biosemantics group. Currently, he is a professor in biosemantics at the Leiden University Medical Center. He was also Head of ELIXIR-NL at the Dutch Techcentre for Life Sciences (until 2015), Integrator Life Sciences at the Netherlands eScience Center, and board member of the Leiden Centre of Data Science. He was one of the authors that, in 2014, initiated the FAIR data initiative and, in the following year, was appointed Chair of the European Commission's High-Level Expert Group for the "European Open Science Cloud" until 2016. Currently, Barend is an ambassador of GO FAIR and co-founder of the GO FAIR initiative and was elected President of the Executive Committee of CODATA.

Cameron Neylon's [41] earlier research focused on structural biology and biophysics and on researchers' culture, the political economy of research institutions, and how these interact, and collide with, the changing technological environments. He is currently a professor of research communication at Curtin University, where he co-leads the Curtin Open Knowledge Initiative, a project examining the future of universities in a networked world. He is also a director of KU research and an advocate of open research practice who has worked in research and support areas, including chemistry, advocacy, policy, technology, publishing, political economy, and cultural studies. He was a contributor to the Panton

Principles for Open Data, the Principles for Open Scholarly Infrastructure, the altmetrics manifesto, founding board member, and past president of FORCE11 and served on the boards and advisory boards of organizations such as Impact Story, Crossref, altmetric.com, OpenAIRE, the LSE Impact Blog, and various editorial boards. His previous positions include Advocacy Director at PLOS, Senior Scientist (Biological Sciences) at the STFC, and tenured faculty at the University of Southampton.

Joachim Schöpfel [42] holds a PhD in psychology from the University of Hamburg and is a lecturer in information and communication sciences at the University of Lille and a member of the GERiiCO laboratory. He is interested in scientific communication, in particular in open science and grey literature, and in the evolution of the functions, professions, and institutions of scientific and technical information. His current projects focus on the use of digital resources in different contexts, on the link between informational practices and scientific production, on the evolution of scientific information systems and the link with research infrastructures and systems, on the legal aspects of scientific communication, and on the development of libraries and documentary services. He directed the UFR IDIST from 2009 to 2012, was director of the Atelier National de la Reproduction des Thèses from 2012 to 2018, and is responsible for the first year of the Master Information Documentation in the SID department. He is an independent consultant and partner of Ourouk, Paris.

Marco Roos [43–45] is an advocate of FAIR Data Principles and Linked Data to create a powerful substrate and a robust worldwide infrastructure for knowledge discovery across heterogeneous data distributed over institutes and countries. His earlier scientific interest was in biology, regarding the role of chromatin in the functioning of the cell, to bridge between genotype and phenotype using data linking techniques and data science. After including computer science subjects in his MSc in molecular biology, Marco worked as a multidisciplinary researcher in research groups of life science and computer science. Currently, his research focuses on state-of-the-art computer science applied to enhancing biomedical research, particularly for rare human diseases, and with knowledge discovery and data linking techniques. As group leader of the Biosemantics research group of Prof. Dr. Barend Mons, LUMC, he leads the research, development, and application of knowledge discovery methods for human genetics research. He co-leads the rare disease community of the European life science data infrastructure ELIXIR, FAIR Data Principles at source' activities in the European Joint Program Rare Diseases, and initiated the Rare Diseases Global Open FAIR implementation network.

This allowed us to perceive that the most prominent authors have a career as professors in higher education and that the original scientific area of most of them is Biology, except for Joachim Schöpfel, whose scientific area of origin is Psychology.

As stated earlier, since this paper is part of qualitative research, the results presented at the current stage of research must be confirmed or tested by further research, gathering data from other databases and information sources, to produce sound theoretical assumptions, and as a form of quality control.

This research takes into consideration the dynamic dimension of the phenomenon under study. The need for continued research to ensure that the developments concerning this phenomenon are captured is evident. This implies the periodic repetition of the same research, which will allow comparison with the current dataset and its updating.

4. Conclusions

This paper is part of developing research focused on identifying and analyzing, in a comparative way, the main programs and projects regarding, or making use, of *FAIR Data Principles* at a worldwide level; identifying the main actors and contrasting their perceptions and meanings about the said principles; distinguishing the proposals and solutions that will emerge from the analysis of perceptions and meaning related to the said principles. The intended future results of this research are a critical and trend-based theoretical construction of the examined literature. This might allow us to formulate

recommendations for the use of *FAIR Data Principles*, in addition to showing possible consensuses and dissents, uncertainties, and certainties behind what is perceived of the said principles and their uses.

Our intention was, within the scope of cutting-edge international scientific literature, to identify the main authors, entities, and scientific journals linked to research conducted, with a special emphasis on the *FAIR Data Principles* since its inception. This allowed the establishment of the general scenario in which research dedicated to or adopting *FAIR Principles* takes place. Within this scope, the study of the researchers' profiles affords better awareness of the original contexts of those who develop this type of research. For such purpose, we defined questions regarding the authors, their locations, and professional occupation; the organizations/entities where the authors perform their professional and research activities, including the place where such entities are based; the journals where and when the authors' research was published; and the most prominent authors in terms of authorship.

This research was developed by means of a qualitative approach, using documentary research and a constant comparison method for codification and categorization of the sampled data extracted from a final set of 32 documents.

In conclusion, it can be stated that, with regard to the authors in the sample, the majority are located in the Netherlands, that the European continent (including the United Kingdom) accounts for more than 70% of the authors, and that the English-speaking countries (including Ireland) comprise just over one-third of the authors discussed in this paper. Four Asian authors are also noted (three from China and one from Japan). The only authors located in the Southern Hemisphere are based in Australia.

Most of these authors are researchers, followed by information (LIS) professionals, and thirdly, the professors in higher education institutions (a quarter of the authors in the sample). Less than five percent of the authors are students.

There are also four CEOs, two independent consultants, and five authors that perform management roles.

More than half of the authors in the sample work in a higher education institution (either professionally or as students), with the institutions with more authors in this sample being, *ex aequo*, the Delft University of Technology and the Dutch Techcentre for Life Sciences. In addition to the research centers, private organizations, and governmental agencies, there are types of entities represented by two authors (Independent Organisations) and one author (International Reference Network, being the VASCERN Europe-a Reference Centre). There are also two independent consultants (Jan Velterop and Caoimhín Kelly).

Since more than half of the organizations with which the authors of the sample are associated are higher education institutions, these cover most of the territorial-administrative areas under consideration. Research centers account for more than one-sixth of the countries in the sample. All types of entities exist in more than one country, including the International Reference Network, as it is an international entity.

The preponderance of the United States, countrywide, can be seen in the number of entities to which this sample refers, with 70% of the entities being from the European world region (58.57% from the European Union) and the English-speaking countries (including Ireland) accounting for just over a third (37.14%). In the case of the higher education institutions considered in the sample, the United States is predominant, and the European Union countries represent more than a third (35.71%) of this type of institution. In the case of the English-speaking countries, there is less than a quarter (24.3%) of the organizations in the sample. The Netherlands appears as the country with the broadest range of entity types in the sample.

At the publication level, it can be seen that the journal where more texts of the sample were published was *Insights: The UKSG Journal*, and most of the authors in the sample were published in *Studies in Health Technology and Informatics*. Most of the articles in the sample published in this journal originate from authors from the Netherlands, followed by authors located in Germany. In addition to *Studies in Health Technology and Informatics*,

only *Frontiers in Chemistry* has articles written by more than ten authors in the sample. This last journal appears in the sample with only one text whose authorship is shared by the greatest number of authors in the sample, in addition to being the one with the greatest number of authors from different geographical provenance (Czechia, France, Germany, Italy, The Netherlands, Portugal, Romania, Slovenia, Spain, and Switzerland). The journals *Information Technology and Libraries*, *Applications in Plant Sciences*, *Digital Library Perspectives*, *Portal: Libraries and the Academy*, *Journal of Librarianship and Scholarly Communications*, *Library Technology Reports*, *Briefings in Bioinformatics*, *International Journal of Librarianship*, and *International Journal of Information Management* are, along with the journals with a single author in the sample (*International Journal of Information Management* and *Library Technology Reports*), the journals that only have articles in the sample from a single country. Most of the articles in the sample were published in 2020.

Finally, the most prominent authors in terms of authorship in the sample texts are Barend Mons and Marco Roos, both from the Netherlands, Cameron Neylon (Australia), and Joachim Schöpfel (France). Although delimited by the sample, this allows us to identify trends that, in further research, may be contrasted with data collected from a larger number of sources and with broader research criteria. In this case, it was perceived that the most prominent authors have a university professor career and that the scientific area of origin of most of them is Biology, except Joachim Schöpfel, whose scientific area of origin is Psychology.

This research takes into consideration the dynamic dimension of the phenomenon under study. The need for continued research to ensure that the developments concerning this phenomenon are captured is evident. This implies the periodic repetition of the same research, which will allow comparison with the current dataset and its updating.

In future work, we will consider expanding the data gathering from other academic journals, conference proceedings, reports, and theses from other databases and collections as sources to verify and compare with the results obtained in the present study. It is also intended to, by means of a trend analysis of the specific scientific literature, identify projects, initiatives, and programs of international expression on the *FAIR Data Principles*.

Moreover, it is intended to proceed to identify the authors and their thoughts, addressing the discussions, perceptions, and meanings that are carried out in and the around this phenomenon. This will allow us to identify and analyze, in a comparative way, the main programs and projects regarding, or making use, of *FAIR Data Principles* at a worldwide level; identify the main actors (also as a way of validation of the results brought out by this paper) and contrast their perceptions and meanings about the said principles; and distinguish the proposals and solutions that will emerge from the analysis of perceptions and meaning related to the said principles. Since the current paper makes use of the CCM, the future analysis and interpretation of the information *corpus* resulting from expanded data gathering will be performed by the use of the Grounded Theory Method. Following the particular nature of this methodology, the intended results are a critical and trend-based theoretical construction of the examined literature. This might allow the formulation of recommendations for the use of *FAIR Data Principles*, besides showing possible consensuses and dissents, uncertainties, and certainties behind what is perceived as the said principles and their uses.

We will also consider guiding the documentary research into developing an analysis that will allow coding and categorization to discern thematic or typological distinctions regarding the research presented by the studied texts. This will also take into account the uses of the concept of "fair principles" that are especially concerned with *FAIR Data Principles*.

Furthermore, future research will look into the scientific areas of origin of the researchers (those already found in this paper and others that will be presented with the expanded data gathering). This will allow us to check the hypothesis that most of the authors who engage in this type of research have biology as their scientific area of origin, as it was perceived by the analysis of the researchers' profiles of the most prominent authors in terms of authorship in the sample texts.

The main limitation of this study is concerned with the amount of data retrieved and the time needed for a deeper analysis, as this theme is already well documented in the scientific literature. Nevertheless, it should be addressed that there were constraints with the collection of the full texts of scientific papers due to the fact that most periodical publications are not freely accessible and are not part of the publishers' contractual packages with the institutions to which we belong. This matter is intrinsically linked to the question of open science and affects the way research can be conducted.

Data Availability Statement: Data are available in Appendices A and B and on request from the authors.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A—Tables

Table A1. List of papers analyzed.

Title	Authors	Journal	#	Pages	Year
A data management workflow of biodiversity data from the field to data users	Rachel A. Hackett Michael W. Belitz Edward E. Gilbert Anna K. Monfils	*Applications in Plant Sciences*	7 (12)		2019
An open-source database model and collections management system for fish scale and otolith archives	Elizabeth Tray Adam Leadbetter Will Meaney Andrew Conway Caoimhín Kelly Niall Ó Maoiléidigh Elvira de Eyto Siobhan Moran Deirdre Brophy	*Ecological Informatics*	59		2020
Applying the FAIR data principles to the Registry of vascular anomalies (VASCA)	Bruna dos Santos Vieira Karlijn Groenen P.A.C. 'T Hoen Annika Jacobsen Marco Roos Rajaram Kaliyaperumal Martijn Kersloot Ronald Cornet Leo Schultze Kool	*Studies in Health Technology and Informatics*	271		2020
Bioinformatics in the Netherlands: The value of a nationwide community	Celia W. G. van Gelder Rob W. W. Hooft Merlijn N. van Rijswijk Linda van den Berg Ruben G. Kok Marcel Reinders Barend Mons Jaap Heringa	*Briefings in Bioinformatics*	20 (2)	375–383	2019
Built to last! Embedding open science principles and practice into European universities	Tiberius Ignat Paul Ayris	*Insights*	33 (1)	1–19	2020
Chapter 2. The Digital Meets the Physical and the Biological	Bohyun Kim	*Library Technology Reports*	56 (2)	8–17	2020

Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud	Barend Mons Cameron Neylon Jan Velterop Michel Dumontier Luiz Olavo Bonino da Silva Santos Mark D. Wilkinson	*Information Services & Use*	37 (1)	49–56	2017
Community feedback on scholarly content: why it is important and why it should be preserved	Heather Staines Maryann E Martone	*Insights*	31	13	2018
Data journals: incentivizing data access and documentation within the scholarly communication system	William H Walters	*Insights*	33 (1)	18	2020
Data Papers as a New Form of Knowledge Organization in the Field of Research Data	Joachim Schöpfel Dominic Farace Hélène Prost Antonella Zane	*Knowledge Organization*	46 (8)	622–638	2019
Ensuring food integrity by metrology and FAIR data principles	Michael Rychlik Giovanna Zappa Larraitz Añorga Nastasia Belc Isabel Castanheira Olivier F.X. Donard Lenka Kourimska Nives Ogrinc Marga C.Ocké Karl Presser Claudia Zoani	*Frontiers in Chemistry*	6	49	2018
ERDMAS: An exemplar-driven institutional research data management and analysis strategy	Matthew I. Bellgard	*International Journal of Information Management*	50	337	2020
FAIR data principles and their application to speech and oral archives	Silvia Calamai Francesca Fontini	*Journal of New Music Research*	47 (4)	339–354	2018
FAIR Principles for Library, Archive and Museum Collections: A proposal for standards for reusable collections	Lukas Koster Saskia Woutersen-Windhouwer	*Code4Lib Journal*	40	1	2018
How are research data governed at Japanese repositories? A knowledge commons perspective	Kai Nishikawa	*ASLIB Journal of Information Management*	72 (4)	671–685	2020
How to Evaluate and Select a Data Repository for Humanities and Social Science: A Case Study of Fudan University Data Repository for Humanities and Social Science	Shenqin Yin Jilong Zhang Menghao Jia Jie Hu	*Library Trends*	69 (1)	125–137	2020
Initiatives, Concepts, and Implementation Practices of FAIR (Findable, Accessible, Interoperable, and Reusable) Data Principles in Health Data Stewardship Practice: Protocol for a Scoping Review	Esther Thea Inau Jean Sack Dagmar Waltemath Atinkut Alamirrew Zeleke	*Journal of Medical Internet Research*	23 (2)		2021
Library eArchiving with ZONTAL Space and the Allotrope Data Format	Dennis Della Corte Wolfgang Colsman Ben Welker Brian Rennick	*Digital Library Perspectives*	36 (1)	69–77	2020
Lives in Data: Prominent Data Librarians, Archivists and Educators Share Their Thoughts	Kristi Thompson Guoying Liu	*International Journal of Librarianship*	2 (1)	66–72	2017

Metadata Documentation Practices at ARL Institutional Repositories	Ayla Stein Kenfield	*Portal: Libraries and the Academy*	19 (4)	667–699	2019
On a Quest for Cultural Change-Surveying Research Data Management Practices at Delft University of Technology	Heather Andrews Mancilla Marta Teperek Jasper van Dijck Kees den Heijer Robbert Eggermont Esther Plomp Yasemin Turkyilmaz-van der Velden Shalini Kurapati	*LIBER Quarterly*	29 (1)	1	2019
OSSE Goes FAIR-Implementation of the FAIR Data Principles for an Open-Source Registry for Rare Diseases	Jannik Shaaf Dennis Kadioglu Jens Goebel Christian-Alexander Behrendt Marco Roos David van Enckevort Frank Ückert Fatlume Sadiku Thomas O.F. Wagner Holger Storf	*Studies in Health Technology and Informatics*	253	209–213	2018
Participatory design for ontologies: a case study of an open science ontology for qualitative coding schemas	Julian Hocker Christoph Schindler Marc Rittberger	*ASLIB Journal of Information Management*	72 (4)	671–685	2020
Research data management in the French National Research Center (CNRS)	Joachim Schöpfel Coline Ferrant Francis Andre Renaud Fabre	*Data Technologies and Applications*	52 (2)	248–265	2018
Research Data Management Services and Strategic Planning in Libraries Today: A Longitudinal Study	Elise Gowen John J. Meier	*Journal of Librarianship and Scholarly Communication*	8 (1)	eP2336	2020
Scientists' Research Data Management Questions: Lessons Learned at a Data Help Desk	Bradley Wade Bishop Rose M. Borden	*Portal: Libraries and the Academy*	20 (4)	677–692	2020
Social infrastructures in research communication: a personal view of the FORCE11 story	Cameron Neylon	*Insights*	31	8	2018
Supporting FAIR Data Principles with Fedora	David Wilcox	*LIBER Quarterly*	28 (1)		2018
The BonaRes metadata schema for geospatial soil-agricultural research data–Merging INSPIRE and DataCite metadata schemes	Xenia Specka Philipp Gärtner Carsten Hoffmann Nikolai Svoboda Markus Stecker Udo Einspanier Kristian Senkler M.A. Muqit Zoarder Uwe Heinrich	*Computers and Geosciences*	132	33–41	2019
Three camps, one destination: the intersections of research data management, FAIR and Open	Rosie Higman Daniel Bangert Sarah Jones	*Insights*	32 (1)	18	2019
Towards FAIRer Biological Knowledge Networks Using a Hybrid Linked Data and Graph Database Approach	Marco Brandizi Ajit Singh Christopher Rawlings Keywan Hassani-Pak	*Journal of Integrative Bioinformatics*	15 (3)		2018

| | Sara Gonzales | | | | |
| User Testing with Microinteractions | Matthew B. Carson Guillaume Viger Lisa O'Keefe Norrina B. Allen Joseph P. Ferrie Kristi Holmes | *Information Technology and Libraries* | 40 (1) | 1–16 | 2021 |

Volume and/or Issue Number of Journal where it was published.

Table A2. Distribution of authors by country.

Country	Authors/Country	Number of Authors/Country	% Authors/Country
Australia	Cameron Neylon Matthew I. Bellgard Menghao Jia	3	2.34%
Canada	David Wilcox Guoying Liu Kristi Thompson	3	2.34%
China	Jie Hu Jilong Zhang Shenqin Yin	3	2.34%
Czechia	Lenka Kourimska	1	0.78%
France	Francesca Frontini Francis Andre Hélène Prost Joachim Schöpfel Olivier F.X. Donard Renaud Fabre	6	4.69%
Germany	Atinkut Alamirrew Zeleke Carsten Hoffmann Christian-Alexander Behrendt Christoph Schindler Dagmar Waltemath Daniel Bangert Dennis Kadioglu Esther Thea Inau Fatlume Sadiku Frank Ückert Holger Storf Jannik Shaaf Jens Goebel Julian Hocker Kristian Senkler M.A. Muqit Zoarder Marc Rittberger Markus Stecker Michael Rychlik Nikolai Svoboda Philipp Gärtner Thomas O.F. Wagner Udo Einspanier Uwe Heinrich Xenia Specka	25	19.53%
Ireland	Adam Leadbetter	9	7.03%

	Andrew Conway		
	Caoimhín Kelly		
	Deirdre Brophy		
	Elizabeth Tray		
	Elvira de Eyto		
	Niall Ó Maoiléidigh		
	Siobhan Moran		
	Will Meaney		
Italy	Antonella Zane	4	3.13%
	Claudia Zoani		
	Giovanna Zappa		
	Silvia Calamai		
Japan	Kai Nishikawa	1	0.78%
The Netherlands	Annika Jacobsen	32	25%
	Barend Mons		
	Bruna dos Santos Vieira		
	Celia W. G. van Gelder		
	David van Enckevort		
	Dominic Farace		
	Esther Plomp		
	Heather Andrews Mancilla		
	Jaap Heringa		
	Jasper van Dijck		
	Karlijn Groenen		
	Kees den Heijer		
	Leo Schultze Kool		
	Linda van den Berg		
	Luiz Olavo Bonino da Silva Santos		
	Lukas Koster		
	Marcel Reinders		
	Marco Roos		
	Marga C. Océ		
	Marta Teperek		
	Martijn Kersloot		
	Merlijn N. van Rijswijk		
	Michel Dumontier		
	P.A.C. (Peter-Bram) 'T Hoen		
	Rajaram Kaliyaperumal		
	Rob W. W. Hooft		
	Robbert Eggermont		
	Ronald Cornet		
	Ruben G. Kok		
	Saskia Woutersen-Wind-houwer		
	Shalini Kurapati		
	Yasemin Turkyilmaz-van der Velden		
Portugal	Isabel Castanheira	1	0.78%
Romania	Nastasia Belc	1	0.78%
Slovenia	Nives Ogrinc	1	0.78%
Spain	Larraitz Añorga	2	1.56%
	Mark D. Wilkinson		
Switzerland	Karl Presser	2	1.56%
	Tiberius Ignat		

		Number of	
UK	Ajit Singh Christopher Rawlings Jan Velterop Keywan Hassani-Pak Marco Brandizi Paul Ayris Rosie Higman Sarah Jones	8	6.25%
USA	Anna K. Monfils Ayla Stein Kenfield Ben Welker Bohyun Kim Bradley Wade Bishop Brian Rennick Dennis Della Corte Edward E. Gilbert Elise Gowen Guillaume Viger Heather Staines Jean Sack John J. Meier Joseph P. Ferrie Kristi Holmes Lisa O'Keefe Maryann E Martone Matthew B. Carson Michael W. Belitz Norrina B. Allen Rachel A. Hackett Rose M. Borden Sara Gonzales William H Walters Wolfgang Colsman	25	19.53%
USA/France	Coline Ferrant	1	0.78%
Total		**128**	**100%**

Table A3. Authors' professional occupation.

Professional Occupation	Authors Name	Number of Authors	Authors %
CEO	Dominic Farace Tiberius Ignat Wolfgang Colsman	3	2.34%
CEO/Researcher	Karl Presser	1	0.78%
Independent consultant	Caoimhín Kelly Jan Velterop	2	1.56%
Information (LIS) professional	Antonella Zane Bohyun Kim Brian Rennick Bruna dos Santos Vieira Elise Gowen Esther Plomp Guoying Liu Heather Andrews Mancilla Jasper van Dijck	25	19.53%

	Jean Sack		
	Jie Hu		
	Jilong Zhang		
	Karlijn Groenen		
	Kees den Heijer		
	Kristi Thompson		
	Lisa O'Keefe		
	Lukas Koster		
	Marta Teperek		
	Paul Ayris		
	Rosie Higman		
	Sara Gonzales		
	Sarah Jones		
	Saskia Woutersen-Windhouwer		
	William H Walters		
	Yasemin Turkyilmaz-van der Velden		
Information (LIS) professional/Professor	Ayla Stein Kenfield	2	1.56%
	John J. Meier		
	Ajit Singh		
	Andrew Conway		
	David Wilcox		
	Guillaume Viger		
	Hélène Prost		
	Jens Goebel		
	Kristian Senkler		
IT professional	Marco Brandizi	15	11.72%
	Markus Stecker		
	Rob W. W. Hooft		
	Robbert Eggermont		
	Rose M. Borden		
	Siobhan Moran		
	Udo Einspanier		
	Will Meaney		
	Heather Staines		
Management	Linda van den Berg	4	3.13%
	Maryann E Martone		
	Merlijn N. van Rijswijk		
	Anna K. Monfils		
	Barend Mons		
	Bradley Wade Bishop		
	Cameron Neylon		
	Dagmar Waltemath		
	Dennis Della Corte		
	Francesca Frontini		
	Holger Storf		
Professor	Joachim Schöpfel	20	15.62%
	Joseph P. Ferrie		
	Kristi Holmes		
	Lenka Kourimska		
	Leo Schultze Kool		
	Luiz Olavo Bonino da Silva Santos		
	Marcel Reinders		
	Michael Rychlik		
	Norrina B. Allen		
	P.A.C. (Peter-Bram) 'T Hoen		

	Silvia Calamai		
	Michel Dumontier		
Researcher	Adam Leadbetter	37	28.91%
	Annika Jacobsen		
	Atinkut Alamirrew Zeleke		
	Carsten Hoffmann		
	Celia W. G. van Gelder		
	Christian-Alexander Behrendt		
	Christoph Schindler		
	Claudia Zoani		
	Daniel Bangert		
	David van Enckevort		
	Edward E. Gilbert		
	Elizabeth Tray		
	Elvira de Eyto		
	Fatlume Sadiku		
	Francis Andre		
	Giovanna Zappa		
	Isabel Castanheira		
	Jannik Shaaf		
	Julian Hocker		
	Kai Nishikawa		
	Larraitz Añorga		
	M.A. Muqit Zoarder		
	Marga C.Ocké		
	Mark D. Wilkinson		
	Matthew B. Carson		
	Matthew I. Bellgard		
	Nastasia Belc		
	Niall Ó Maoiléidigh		
	Nikolai Svoboda		
	Olivier F.X. Donard		
	Philipp Gärtner		
	Rajaram Kaliyaperumal		
	Shalini Kurapati		
	Shenqin Yin		
	Thomas O.F. Wagner		
	Uwe Heinrich		
	Xenia Specka		
Researcher/ Management	Martijn Kersloot	1	0.78%
Researcher/Professor	Christopher Rawlings	12	9.38%
	Deirdre Brophy		
	Dennis Kadioglu		
	Frank Ückert		
	Jaap Heringa		
	Keywan Hassani-Pak		
	Marc Rittberger		
	Marco Roos		
	Nives Ogrinc		
	Renaud Fabre		
	Ronald Cornet		
	Ruben G. Kok		
Student	Ben Welker	6	4.69%
	Coline Ferrant		
	Esther Thea Inau		

	Menghao Jia			
	Michael W. Belitz			
	Rachel A. Hackett			
Total			**128**	**100%**

Table A4. Distribution of authors per organization, by type of entity.

Institutions	Authors	Country	Number of Authors	% Authors
International Reference Network			**1**	**0.71%**
VASCERN European Reference Centre	Leo Schultze Kool	European Union	1	0.71%
Governmental Agency			**9**	**6.38%**
Marine Institute	Adam Leadbetter Andrew Conway Elvira de Eyto Niall Ó Maoiléidigh Siobhan Moran Will Meaney	Ireland	6	4.26%
National Agency for New Technologies, Energy and Sustainable Economic Development	Claudia Zoani Giovanna Zappa	Italy	2	1.42%
National Institute for Public Health and the Environment	Marga C. Océ	The Netherlands	1	0.71%
Higher Education Institution			**85**	**60.28%**
Amsterdam University of Applied Sciences	Lukas Koster Saskia Woutersen-Windhouwer	The Netherlands	2	1.42%
Arizona State University	Edward E. Gilbert	USA	1	0.71%
Brigham Young University	Ben Welker Brian Rennick Dennis Della Corte	USA	3	2.13%
Central Michigan University	Anna K. Monfils Michael W. Belitz Rachel A. Hackett	USA	3	2.13%
Curtin University	Cameron Neylon	Australia	1	0.71%
Czech University of Life Sciences	Lenka Kourimska	Czechia	1	0.71%
Darmstadt University of Applied Sciences	Marc Rittberger	Germany	1	0.71%
Delft University of Technology	Esther Plomp Heather Andrews Mancilla Jasper van Dijck Kees den Heijer Marcel Reinders Marta Teperek Robbert Eggermont Shalini Kurapati Yasemin Turkyilmaz-van der Velden	The Netherlands	9	6.38%
Fudan University	Jie Hu Jilong Zhang Shenqin Yin	China	3	2.13%
Galway-Mayo Institute of Technology	Deirdre Brophy Elizabeth Tray	Ireland	2	1.42%
Goethe University Frankfurt	Holger Storf Jannik Shaaf Jens Goebel Thomas O.F. Wagner	Germany	4	2.84%
Greifswald Medical School	Atinkut Alamirrew Zeleke Dagmar Waltemath Esther Thea Inau	Germany	3	2.13%

Johns Hopkins Bloomberg School of Public Health	Jean Sack	USA	1	0.71%
Jožef Stefan Institute	Nives Ogrinc	Slovenia	1	0.71%
Leiden University	Annika Jacobsen Barend Mons Marco Roos Rajaram Kaliyaperumal	The Netherlands	4	2.84%
Maastricht University	Michel Dumontier	The Netherlands	1	0.71%
Manhattan College	William H Walters	USA	1	0.71%
Northwestern University	Joseph P. Ferrie Kristi Holmes Norrina B. Allen Coline Ferrant Guillaume Viger Lisa O'Keefe Matthew B. Carson Sara Gonzales	USA	8	5.67%
Paris Institute of Political Studies	Coline Ferrant	France	1	0.71%
Penn State University	Elise Gowen John J. Meier	USA	2	1.42%
Queensland University of Technology	Matthew I. Bellgard	Australia	1	0.71%
Radboud University	Bruna dos Santos Vieira Karlijn Groenen Leo Schultze Kool P.A.C. (Peter-Bram) 'T Hoen	The Netherlands	4	2.84%
Swiss Federal Institute of Technology	Karl Presser	Switzerland	1	0.71%
Technical University of Munich	Michael Rychlik	Germany	1	0.71%
Universidad Politécnica de Madrid	Mark D. Wilkinson	Spain	1	0.71%
Università degli Studi di Siena	Silvia Calamai	Italy	1	0.71%
Université de Pau et des Pays de l'Adour	Olivier F.X. Donard	France	1	0.71%
Université Paul-Valéry Montpellier 3	Francesca Frontini	France	1	0.71%
University College London	Paul Ayris	UK	1	0.71%
University of Göttingen	Daniel Bangert	Germany	1	0.71%
University of Groningen	David van Enckevort	The Netherlands	1	0.71%
University of Hamburg	Christian-Alexander Behrendt Dennis Kadioglu	Germany	2	1.42%
University of Hildesheim	Julian Hocker	Germany	1	0.71%
University of Illinois	Ayla Stein Kenfield	USA	1	0.71%
University of Lille III	Hélène Prost Joachim Schöpfel	France	2	1.42%
University of Manchester	Rosie Higman	UK	1	0.71%
University of Michigan	Bohyun Kim	USA	1	0.71%
University of Padova	Antonella Zane	Italy	1	0.71%
University of Paris VIII	Renaud Fabre	France	1	0.71%
University of South Australia	Menghao Jia	Australia	1	0.71%
University of Tennessee	Bradley Wade Bishop	USA	1	0.71%
University of Tsukuba	Kai Nishikawa	Japan	1	0.71%
University of Windsor	Guoying Liu Kristi Thompson	Canada	2	1.42%
Vrije Universiteit (VU) Amsterdam	Jaap Heringa Luiz Olavo Bonino da Silva Santos Martijn Kersloot Ronald Cornet	The Netherlands	4	2.84%
Independent Consultant			**2**	**1.42%**
Consultant	Jan Velterop	UK	1	0.71%
Consultant	Caoimhín Kelly	Ireland	1	0.71%
Independent Organization			**2**	**1.42%**
DuraSpace	David Wilcox	Canada	1	0.71%

GreyNet	Dominic Farace	The Netherlands	1	0.71%
Private Organization			**9**	**6.38%**
Castor EDC, Amsterdam	Martijn Kersloot	The Netherlands	1	0.71%
con Terra GmbH	Kristian Senkler Markus Stecker Udo Einspanier	Germany	3	2.13%
Hypothesis	Heather Staines Maryann E Martone	USA	2	1.42%
Premotec GmbH	Karl Presser	Switzerland	1	0.71%
Scientific Knowledge Services	Tiberius Ignat	Switzerland	1	0.71%
Zontal, Inc	Wolfgang Colsman	USA	1	0.71%
Research Center			**33**	**23.40%**
Centre for Electrochemical Technologies	Larraitz Añorga	Spain	1	0.71%
Centre National de la Recherche Scientifique	Francis Andre Hélène Prost Renaud Fabre	France	3	2.13%
Digital Curation Centre	Sarah Jones	UK	1	0.71%
Dutch Techcentre for Life Sciences	Barend Mons Celia W. G. van Gelder Jaap Heringa Linda van den Berg Luiz Olavo Bonino da Silva Santos Marco Roos Merlijn N. van Rijswijk Rob W. W. Hooft Ruben G. Kok	The Netherlands	9	6.38%
German Cancer Research Center	Fatlume Sadiku Frank Ückert	Germany	2	1.42%
Instituto Nacional de Saúde Doutor Ricardo Jorge	Isabel Castanheira	Portugal	1	0.71%
Leibniz Centre for Agricultural Landscape Research	Carsten Hoffmann M.A. Muqit Zoarder Nikolai Svoboda Philipp Gärtner Uwe Heinrich Xenia Specka	Germany	6	4.26%
Leibniz Institute for Research and Information in Education	Christoph Schindler Julian Hocker Marc Rittberger	Germany	3	2.13%
National R&D Institute for Food Bioresources	Nastasia Belc	Romania	1	0.71%
The Netherlands eScience Centre	Barend Mons	The Netherlands	1	0.71%
Rothamsted Research	Ajit Singh Christopher Rawlings Keywan Hassani-Pak Marco Brandizi	UK	4	2.84%
Sandia National Laboratories	Rose M. Borden	USA	1	0.71%
Total			141	100%

Table A5. Entities by type and country.

		Australia	Canada	China	Czechia	France	Germany	International	Ireland	Italy	Japan	The The Netherlands	Portugal	Romania	Slovenia	Spain	Switzerland	UK	USA	Total
Governmental Agency	(a) *								1	1		1								3
	(b) *								1.43%	1.43%		1.43%								4.29%
	(c) *								33.33%	33.33%		8.33%								4.29%
	(d) *								33.33%	33.33%		33.33%								100%
Higher Education Institution	(a) *	3	1	1	1	5	7		1	2	1	7			1	1	1	2	10	44
	(b) *	4.29%	1.43%	1.43%	1.43%	7.14%	10%		1.43%	2.86%	1.43%	10%			1.43%	1.43%	1.43%	2.86%	14.29%	62.86%
	(c) *	100%	50%	100%	100%	83.33%	63.64%		33.33%	66.67%	100%	58.33%			100%	50%	33.33%	40%	76.92%	62.86%
	(d) *	6.82%	2.27%	2.27%	2.27%	11.36%	15.91%		2.27%	4.55%	2.27%	15.91%			2.27%	2.27%	2.27%	4.55%	22.73%	100%
Independent Consultant	(a) *								1									1		2
	(b) *								1.43%									1.43%		2.86%
	(c) *								33.33%									20%		2.86%
	(d) *								50%									50%		100%
Independent Organization	(a) *		1									1								2
	(b) *		1.43%									1.43%								2.86%
	(c) *		50%									8.33%								2.86%
	(d) *		50%									50%								100%
International Reference Network	(a) *							1												1
	(b) *							1.43%												1.43%
	(c) *							100%												1.43%
	(d) *							100%												100%
Private Organization	(a) *						1					1					2		2	6
	(b) *						1.43%					1.43%					2.86%		2.86%	8.57%
	(c) *						9.09%					8.33%					66.67%		15.38%	8.57%
	(d) *						16.67%					16.67%					33.33%		33.33%	100%
Research Centre	(a) *					1	3					2	1	1		1		2	1	12
	(b) *					1.43%	4.29%					2.86%	1.43%	1.43%		1.43%		2.86%	1.43%	17.14%
	(c) *					16.67%	27.27%					16.67%	100%	100%		50%		40%	7.69%	17.14%
	(d) *					8.33%	25%					16.67%	8.33%	8.33%		8.33%		16.67%	8.33%	100%
(e) *		3	2	1	1	6	11	1	3	3	1	12	1	1	1	2	3	5	13	70
(f) *		4.29%	2.86%	1.43%	1.43%	8.57%	15.71%	1.43%	4.29%	4.29%	1.43%	17.14%	1.43%	1.43%	1.43%	2.86%	4.29%	7.14%	18.57%	100%

* (a) Organizations; (b) % Organizations; (c) % Organizations per Country; (d) Organizations per type; (e) Organizations per Country Total; (f) Organizations Percentage Total.

Table A6. Distribution of analyzed papers by scientific journals.

Journals	ISSN	Country	# Texts	% Texts
Applications in Plant Sciences	2168-0450	USA	1	3.13%
Aslib Journal of Information Management	2050-3814	UK	2	6.25%
Briefings in Bioinformatics	1467-5463	UK	1	3.13%
Code4Lib Journal	1940-5758	USA	1	3.13%
Computers and Geosciences	0098-3004	UK	1	3.13%
Data Technologies and Applications	2514-9288	UK	1	3.13%
Digital Library Perspectives	2059-5824	UK	1	3.13%
Ecological Informatics	1878-0512	The Netherlands	1	3.13%
Frontiers in Chemistry	2296-2646	Switzerland	1	3.13%
Information Services & Use	1875-8789	The Netherlands	1	3.13%
Information Technology and Libraries	2163-5226	USA	1	3.13%
Insights: The UKSG Journal	2048-7754	UK	5	15.63%
International Journal of Information Management	0268-4012	UK	1	3.13%
International Journal of Librarianship	2474-3542	USA	1	3.13%
Journal of Integrative Bioinformatics	1613-4516	Germany	1	3.13%
Journal of Librarianship and Scholarly Communication	2162-3309	USA	1	3.13%
Journal of Medical Internet Research	1438-8871	Canada	1	3.13%
Journal of New Music Research	0929-8215	The Netherlands	1	3.13%
Knowledge Organization	0943-7444	Germany	1	3.13%
LIBER Quarterly	1435-5205	Germany	2	6.25%
Library Technology Reports	1945-4538	USA	1	3.13%
Library Trends	0024-2594	USA	1	3.13%
Portal: Libraries and the Academy	1530-7131	USA	2	6.25%
Studies in Health Technology and Informatics	0926-9630	European Union	2	6.25%
		Total	**32**	**100%**

Table A7. Number of authors per journal.

Journal	Number of Authors	% Authors
Applications in Plant Sciences	4	3.05%
ASLIB Journal of Information Management	4	3.05%
Briefings in Bioinformatics	8	6.11%
Code4Lib Journal	2	1.53%
Computers and Geosciences	9	6.87%
Data Technologies and Applications	4	3.05%
Digital Library Perspectives	4	3.05%
Ecological Informatics	9	6.87%
Frontiers in Chemistry	11	8.4%
Information Services & Use	6	4.58%
Information Technology and Libraries	7	5.34%
Insights: The UKSG Journal	9	6.87%
International Journal of Information Management	1	0.76%
International Journal of Librarianship	2	1.53%
Journal of Integrative Bioinformatics	4	3.05%
Journal of Librarianship and Scholarly Communication	2	1.53%
Journal of Medical Internet Research	4	3.05%
Journal of New Music Research	2	1.53%
Knowledge Organization	4	3.05%
LIBER Quarterly	9	6.87%
Library Technology Reports	1	0.76%

Library Trends	4	3.05%
Portal: Libraries and the Academy	3	2.29%
Studies in Health Technology and Informatics	18	13.74%
Total	**131**	**100%**

Table A8. Geographical distribution of authors by journal.

Journals/Provenance of Authors	Number of Authors	% Author
Applications in Plant Sciences	**4**	**3.05%**
USA	4	3.05%
ASLIB Journal of Information Management	**4**	**3.05%**
Germany	3	2.29%
Japan	1	0.76%
Briefings in Bioinformatics	**8**	**6.11%**
The Netherlands	8	6.11%
Code4Lib Journal	**2**	**1.53%**
The Netherlands	2	1.53%
Computers and Geosciences	**9**	**6.87%**
Germany	9	6.87%
Data Technologies and Applications	**4**	**3.05%**
France	3	2.29%
USA/France	1	0.76%
Digital Library Perspectives	**4**	**3.05%**
USA	4	3.05%
Ecological Informatics	**9**	**6.87%**
Ireland	9	6.87%
Frontiers in Chemistry	**11**	**8.40%**
Czechia	1	0.76%
France	1	0.76%
Germany	1	0.76%
Italy	2	1.53%
The Netherlands	1	0.76%
Portugal	1	0.76%
Romania	1	0.76%
Slovenia	1	0.76%
Spain	1	0.76%
Switzerland	1	0.76%
Information Services & Use	**6**	**4.58%**
Australia	1	0.76%
The Netherlands	3	2.29%
Spain	1	0.76%
UK	1	0.76%
Information Technology and Libraries	**7**	**5.34%**
USA	7	5.34%
Insights: The UKSG Journal	**9**	**6.87%**
Australia	1	0.76%
Germany	1	0.76%
Switzerland	1	0.76%
UK	3	2.29%
USA	3	2.29%
International Journal of Information Management	**1**	**0.76%**
Australia	1	0.76%
International Journal of Librarianship	**2**	**1.53%**
Canada	2	1.53%
Journal of Integrative Bioinformatics	**4**	**3.05%**

|			
UK		4	3.05%
Journal of Librarianship and Scholarly Communication		**2**	**1.53%**
USA		2	1.53%
Journal of Medical Internet Research		**4**	**3.05%**
Germany		3	2.29%
USA		1	0.76%
Journal of New Music Research		**2**	**1.53%**
France		1	0.76%
Italy		1	0.76%
Knowledge Organization		**4**	**3.05%**
France		2	1.53%
Italy		1	0.76%
The Netherlands		1	0.76%
LIBER Quarterly		**9**	**6.87%**
Canada		1	0.76%
The Netherlands		8	6.11%
Library Technology Reports		**1**	**0.76%**
USA		1	0.76%
Library Trends		**4**	**3.05%**
Australia		1	0.76%
China		3	2.29%
Portal: Libraries and the Academy		**3**	**2.29%**
USA		3	2.29%
Studies in Health Technology and Informatics		**18**	**13.74%**
Germany		8	6.11%
The Netherlands		10	7.63%
	Total	**131**	**100%**

Appendix B—Figures
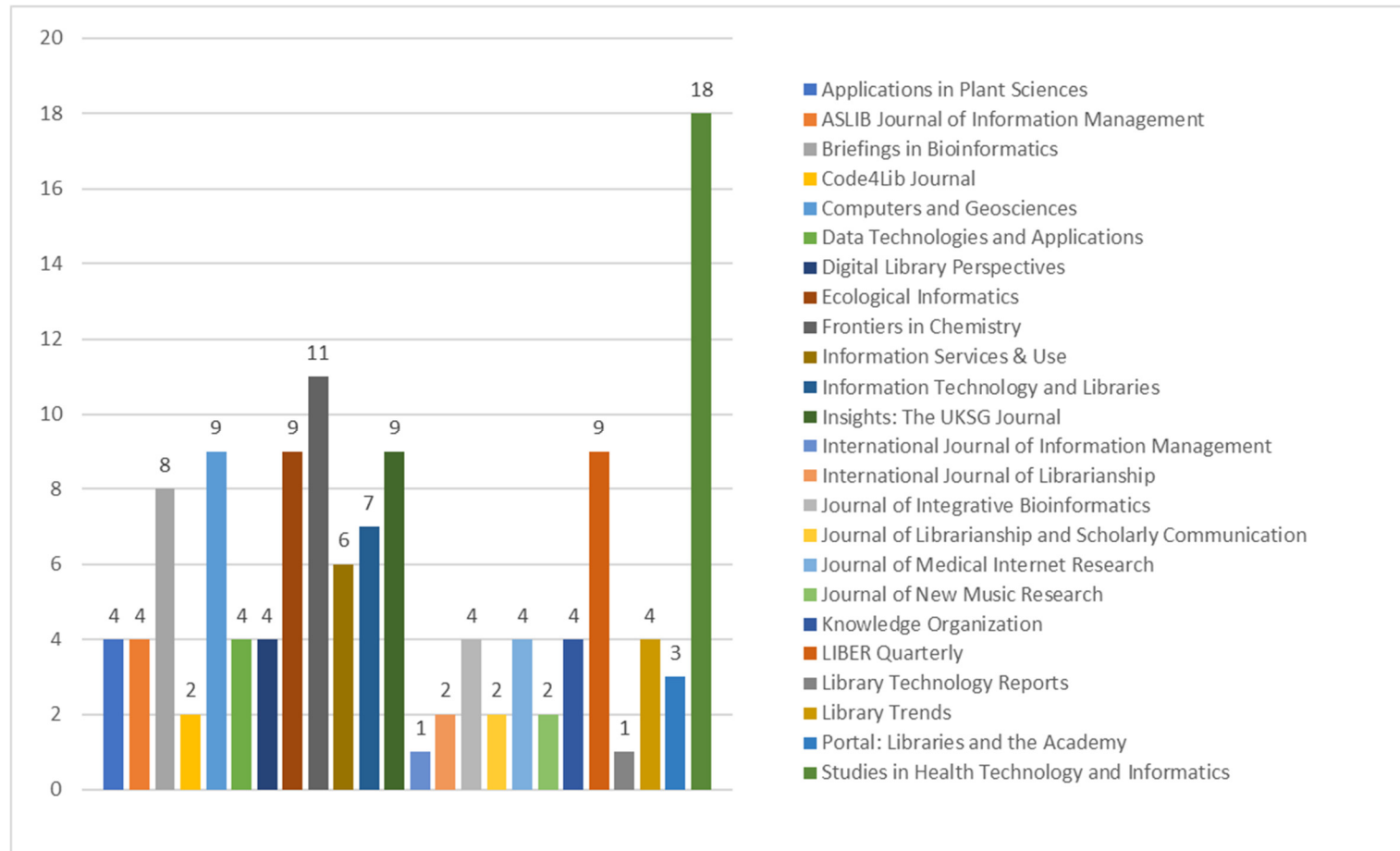
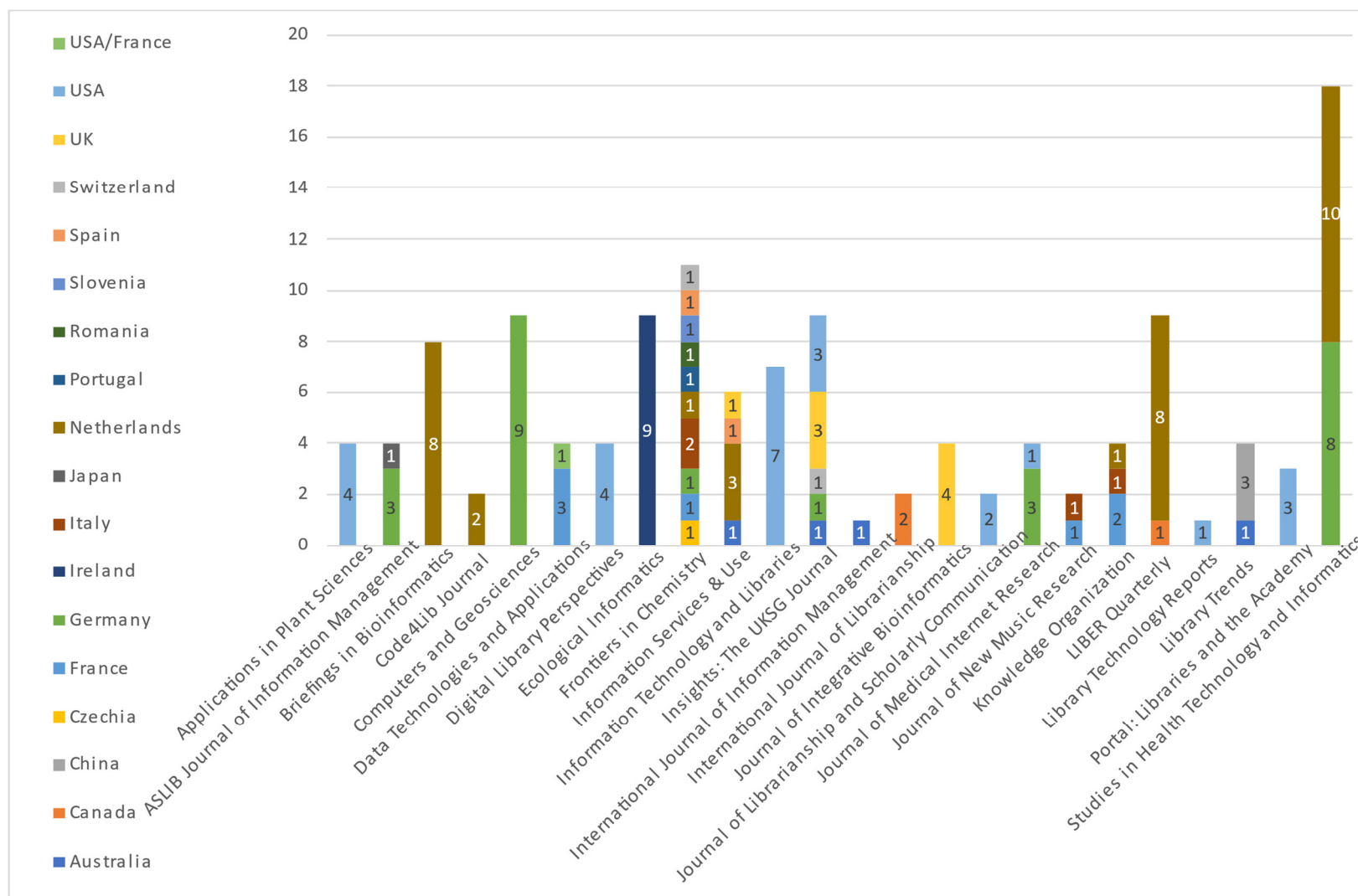**Figure A1.** Authors per journal.

Figure A2. Geographical distribution of authors by journal.

References

1. Neylon, C. Social infrastructures in research communication: A personal view of the FORCE11 story. *Insights* **2018**, *31*, 8. https://doi.org/10.1629/uksg.404.
2. Wilkinson, M.D.; Dumontier, M.; Aalbersberg, I.J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.-W.; da Silva Santos, L.B.; Bourne, P.E.; et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **2016**, *3*, 160018. https://doi.org/10.1038/sdata.2016.18.
3. Mons, B.; Neylon, C.; Velterop, J.; Dumontier, M.; da Silva Santos, L.O.B.; Wilkinson, M.D. Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud. *Inf. Serv. Use* **2017**, *37*, 49–56. https://doi.org/10.3233/ISU-170824.
4. Higman, R.; Bangert, D.; Jones, S. Three camps, one destination: The intersections of research data management, FAIR and Open. *Insights* **2019**, *32*, 18. https://doi.org/10.1629/uksg.468.
5. Lin, D.; Crabtree, J.; Dillo, I.; Downs, R.R.; Edmunds, R.; Giaretta, D.; De Giusti, M.; L'Hours, H.; Hugo, W.; Jenkyns, R.; et al. The TRUST Principles for digital repositories. *Sci. Data* **2020**, *7*, 144. https://doi.org/10.1038/s41597-020-0486-7.
6. NESTOR. *NESTOR Endorsement of TRUST Principles*; NESTOR: Frankfurt am Main, Germany, 2020. Available online: https://www.langzeitarchivierung.de/Webs/nestor/SharedDocs/Downloads/EN/2020EndorsementOfTrustPrinci-ples.html?nn=182510 (accessed on 25 May 2022).
7. Inau, E.T.; Sack, J.; Waltemath, D.; Zeleke, A.A. Initiatives, Concepts, and Implementation Practices of FAIR (Findable, Accessible, Interoperable, and Reusable) Data Principles in Health Data Stewardship Practice: Protocol for a Scoping Review. *J. Med. Internet Res.* **2021**, *10*, e22505. https://doi.org/10.2196/22505.
8. Van Gelder, C.W.G.; Hooft, R.W.W.; Van Rijswijk, M.N.; Van Den Berg, L.; Kok, R.G.; Heringa, J.; Reinders, M.; Mons, B. Bioinformatics in the Netherlands: The value of a nationwide community. *Brief. Bioinform.* **2019**, *20*, 375–383. https://doi.org/10.1093/bib/bbx087.
9. Brandizi, M.; Singh, A.; Rawlings, C.; Hassani-Pak, K. Towards FAIRer Biological Knowledge Networks Using a Hybrid Linked Data and Graph Database Approach. *J. Integr. Bioinform.* **2018**, *15*, 20180023. https://doi.org/10.1515/jib-2018-0023.
10. Koster, L.; Woutersen-Windhouwer, S. FAIR Principles for Library, Archive and Museum Collections: A proposal for standards for reusable collections. *Code4Lib J.* **2018**, *40*, 1. Available online: https://journal.code4lib.org/articles/13427 (accessed on 30 June 2022).
11. Silva, C. Investigação Documental. In *Manual de Investigação Qualitativa*; Gonçalves, S., Gonçalves, P.E., Marques, C., Eds.; Pactor: Lisbon, Portugal, 2021; pp. 103–123.
12. Flick, U. *Doing Grounded Theory*, 2nd ed.; SAGE: Los Angeles, CA, USA, 2018.
13. Glaser, B.G. The Constant Comparative Method of Qualitative Analysis. *Soc. Probl.* **1965**, *12*, 436–445.
14. Bardin, L. *Análise de Conteúdo*; Edições 70: São Paulo, Brasil, 2011.
15. Rychlik, M.; Zappa, G.; Anorga, L.; Belc, N.; Castanheira, I.; Donard, O.; Kourimska, L.; Ogrinc, N.; Ocke, M.; Presser, K.; et al. Ensuring Food Integrity by Metrology and FAIR Data Principles. *Front. Chem.* **2018**, *6*, 49. https://doi.org/10.3389/fchem.2018.00049.
16. Schaaf, J.; Goebel, J.; Storf, H.; Kadioglu, D.; Behrendt, C.-A.; Roos, M.; van Enckevort, D.; Ückert, F.; Sadiku, F.; Wagner, T.O.F. OSSE Goes FAIR—Implementation of the FAIR Data Principles for an Open-Source Registry for Rare Diseases. *Stud. Health Technol. Inform.* **2018**, *253*, 209–213.
17. Tray, E.; Brophy, D.; Leadbetter, A.; Meaney, W.; Conway, A.; Kelly, C.; Moran, S.; Maoiléidigh, N.Ó.; de Eyto, E. An open-source database model and collections management system for fish scale and otolith archives. *Ecol. Inform.* **2020**, *59*, 101115. https://doi.org/10.1016/j.ecoinf.2020.101115.
18. Vieira, B.; Groenen, K.; Hoen, P.A.C.T.; Jacobsen, A.; Roos, M.; Kaliyaperumal, R.; Kersloot, M.; Cornet, R.; Koola, L.S. Applying the FAIR data principles to the Registry of vascular anomalies (VASCA). *Stud. Health Technol. Inform.* **2020**, *271*, 115–116. https://doi.org/10.3233/SHTI200085.
19. Specka, X.; Gärtner, P.; Hoffmann, C.; Svoboda, N.; Zoarder, M.A.M.; Heinrich, U.; Stecker, M.; Einspanier, U.; Senkler, K. The BonaRes metadata schema for geospatial soil-agricultural research data—Merging INSPIRE and DataCite metadata schemes. *Comput. Geosci.* **2019**, *132*, 33–41.
20. Mancilla, H.A.; Teperek, M.; van Dijck, J.; den Heijer, K.; Eggermont, R.; Plomp, E.; Turkyilmaz-van der Velden, Y.; Kurapati, S. On a Quest for Cultural Change—Surveying Research Data Management Practices at Delft University of Technology. *LIBER Q.* **2019**, *29*, 1.
21. Gonzales, S.; Carson, M.B.; Viger, G.; O'Keefe, L.; Allen, N.B.; Ferrie, J.P.; Holmes, K. User Testing with Microinteractions. *Inf. Technol. Libr.* **2021**, *40*, 1–16.
22. Hackett, R.A.; Belitz, M.W.; Monfils, A.K.; Gilbert, E.E. A data management workflow of biodiversity data from the field to data users. *Appl. Plant Sci.* **2019**, *7*, e11310. https://doi.org/10.1002/aps3.11310.
23. Schöpfel, J.; Farace, D.; Prost, H.; Zane, A. Data Papers as a New Form of Knowledge Organization in the Field of Research Data. *Knowl. Organ.* **2019**, *46*, 622–638.
24. Yin, S.; Zhang, J.; Jia, M.; Hu, J. How to Evaluate and Select a Data Repository for Humanities and Social Science: A Case Study of Fudan University Data Repository for Humanities and Social Science. *Libr. Trends* **2020**, *69*, 125–137.
25. Dennis, D.C.; Colsman, W.; Welker, B.; Rennick, B. Library eArchiving with ZONTAL Space and the Allotrope Data Format. *Digit. Libr. Perspect.* **2020**, *36*, 69–77.

26. Schöpfel, J.; Ferrant, C.; Francis, A.; Fabre, R. Research data management in the French National Research Center (CNRS). *Data Technol. Appl.* **2018**, *52*, 248–265.
27. Hocker, J.; Schindler, C.; Rittberger, M. Participatory design for ontologies: A case study of an open science ontology for qualitative coding schemas. *Aslib J. Inf. Manag.* **2020**, *72*, 671–685.
28. Ignat, T.; Ayris, P. Built to last! Embedding open science principles and practice into European universities. *Insights* **2020**, *33*, 9.
29. Staines, H.; Martone, M.E. Community feedback on scholarly content: Why it is important and why it should be preserved. *Insights* **2018**, *31*, 13.
30. Calamai, S.; Frontini, F. FAIR data principles and their application to speech and oral archives. *J. New Music Res.* **2018**, *47*, 339–354.
31. Thompson, K.; Liu, G. Lives in Data: Prominent Data Librarians, Archivists and Educators Share Their Thoughts. *Int. J. Librariansh.* **2017**, *2*, 66–72.
32. Gowen, E.; Meier, J.J. Research Data Management Services and Strategic Planning in Libraries Today: A Longitudinal Study. *J. Librariansh. Sch. Commun.* **2020**, *8*, eP2336. https://doi.org/10.7710/2162-3309.2336.
33. Bishop, W.B.; Borden, R.M. Scientists' Research Data Management Questions: Lessons Learned at a Data Help Desk. *Portal Libr. Acad.* **2020**, *20*, 677–692.
34. Kim, B. Chapter 2. The Digital Meets the Physical and the Biological. *Libr. Technol. Rep.* **2020**, *56*, 8–17.
35. Walters, W.H. Data journals: Incentivizing data access and documentation within the scholarly communication system. *Insights* **2020**, *33*, 18.
36. Bellgard, M.I. ERDMAS: An exemplar-driven institutional research data management and analysis strategy: SSIS. *Int. J. Inf. Manag.* **2020**, *50*, 337.
37. Nishikawa, K. How are research data governed at Japanese repositories? A knowledge commons perspective. *Aslib J. Inf. Manag.* **2020**, *72*, 837–852.
38. Ayla, S.K. Metadata Documentation Practices at ARL Institutional Repositories. *Portal Libr. Acad.* **2019**, *19*, 667–699.
39. Wilcox, D. Supporting FAIR Data Principles with FEDORA. *LIBER Q.* **2018**, *28*, 1–8. https://doi.org/10.18352/lq.10247.
40. CODATA Officers and Executive Commitee. Available online: https://codata.org/about-codata/executive-committee/ (accessed on 30 June 2022).
41. Science in the Open—The Online Home of Cameron Neylon. Available online: https://cameronneylon.net/about/biographies/ (accessed on 30 June 2022).
42. Enseignants-Chercheurs, Chercheurs, Enseignants de la Université de Lille. Available online: https://pro.univ-lille.fr/joachim-schopfel/ (accessed on 30 June 2022).
43. Homepage of Marco Roos. Available online: https://www.nanopubmed.org/?content=MyShortCV (accessed on 30 June 2022).
44. Leiden University Medical Center—Marco Roos Assistant Professor. Available online: https://www.lumc.nl/org/humane-genetica/medewerkers/marco-roos (accessed on 30 June 2022).
45. Leiden University Medical Center—Marco Roos (PhD). Available online: https://www.lumc.nl/org/bioinformatica/medewerkers/909290026392525 (accessed on 30 June 2022).