Intersections in the scientific production of library and information science and data science

Intersecciones en la producción científica de la ciencia de la información y la ciencia de datos / Interseções na produção científica da ciência da informação e ciência de dados

Maurício Torres de Matos
Master in Library and Information Science by the Universidade Federal do Pará (UFPA), Brazil.
mauriciotmatos@gmail.com

Marise Teles Condurú
PhD in Social and Environmental Development Sciences by the UFPA. Professor in the Faculty of Librarianship, in the master's program in Library and Information Science and the professional master's degree in Natural Resource Management and Local Development at the UFPA, Brazil.
marise@ufpa.br

Alegria Celia Benchimol
PhD in Library and Information Science by the agreement between Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) and Universidade Federal do Rio de Janeiro (UFRJ). Professor in the Library and Information Science graduate program at the UFPA and in the Sociocultural Diversity graduate program at the Museu Paraense Emílio Goeldi, Brazil.
alegria.benchimol@gmail.com

ABSTRACT

This article aims to analyze the relationship of scientific production between library and information science and data science using bibliometrics, with the analysis of 232 articles dated between 1996 and 2021. The correlation between them is highlighted by the growth of production, themes in common, significant production in open access, and wide dissemination of journals and institutions. It was concluded that the intersection between the disciplines seems to be a growing trend, which can be seen by the increase in interconnected production in recent years, allowing for greater dialogue.

Keywords: information science; data science; bibliometric study.

RESUMEN

Este artículo tiene como objetivo analizar la relación de producción científica entre la ciencia de la información y la ciencia de datos utilizando la bibliometría, con el análisis de 232 artículos entre 1996 y 2021. Se destaca la correlación entre ellos por el crecimiento de la producción, temas en común, producción significativa en acceso abierto y amplia difusión de revistas e instituciones. Se concluyó que la intersección entre las disciplinas parece ser una tendencia creciente, lo que se puede ver por el aumento de la producción interconectada en los últimos años, lo que permite un mayor diálogo.

Palabras clave: ciencia de la información; ciencia de datos; estudio bibliométrico.

RESUMO

Este artigo objetiva analisar a relação entre a produção científica da ciência da informação e da ciência de dados utilizando a bibliometria, com a análise de 232 artigos, entre 1996 e 2021. Destaca-se a correlação entre elas pelo crescimento da produção, temáticas em comum, significativa produção em acesso aberto e ampla difusão de periódicos e instituições. Conclui-se que a interseção entre as disciplinas parece ser uma tendência em crescimento, que pode ser percebida pelo aumento da produção interligada nos últimos anos, permitindo maior diálogo.

Palavras-chave: ciência da informação; ciência de dados; estudo bibliométrico.

Introduction

Data science (DS) is a discipline that emerged in the early 2000s and that focus on the analysis of digital data in a technological context where data exists in vast and diversified volumes and is produced at an accelerated pace. According to Coneglian, Gonçalves and Santarém Segundo (2017), this context has brought major changes in data analysis methods both in organizational and academic environments.

As its objective, data science aims to create techniques, technologies, and theories to utilize data from different and complex sources, to help with decision making and creating informational value (Cao, 2016). It is an interdisciplinary field which, while pursuing its objectives, concerns itself with data since its collection and representation, through its storage, security, analysis, and dissemination, having a strong technological component.

Considering Saracevic's (1999, p. 1.051) definition of library and information science (LIS) as "the science and practice dealing with the effective collection, storage, retrieval, and use of information", is possible to perceive similarities between its object of study and data science. The interdisciplinarity that characterizes LIS, especially with computer science, in addition to the need for LIS to keep up to date with technological advances, allows to consider a possible growth of the scientific production studying intersections between these areas, as well as possible actuation areas for information scientists in activities related to DS.

Thus, it was raised as a research question to know: how does the intersection between the scientific production of library and information science with data science occur, to identify possible common fields of study, as well as verify the production of open access related to the theme?

To answer this question, it was defined as general objective of this study to analyze the scientific production relating library and information science and data science, having as specific objectives: a) to identify institutions, countries, researchers, and open access scientific journals that most addressed the topic; b) verify study trends with intersections between LIS and DS.

The research was carried out founded on basic bibliometric patterns of scientific publications that present a correlation between library and information science and data science, available in the Web of Science and Scopus databases, to identify and measure advances and trends, based on the most frequent topics, volume of publications, most prolific authors, open access scientific journals, scientific institutions, and countries with the highest scientific

production. Temporal limitations were not used in the research, to include the largest possible number of results to compose the research sample and so the results covered the period from 1996 to 2021. The research also sought to analyze the international scenario of the subject studied, intending to better identify trends and then compare it with the Brazilian production, as will be discussed in the results.

To support the discussion regarding the results, theoretical aspects of LIS and DS are discussed below.

About data science and library and information science

The last few decades have seen technological advances at an unprecedented pace in human history, especially in information and communication technologies (ICT). These advances generated new habits and ways that people interact with technology, resulting in the production of a vast volume and variety of information at an accelerated pace that, although it has made access to information more democratic and simpler in many ways, has created challenges for the use, mediation, organization, and retrieval of information, especially in distinguishing between information relevant to the user and expendable information.

Because of that context, around the early 2000s the first works appeared in a new discipline, called data science. This science, characterized by its interdisciplinarity, focuses on obtaining meaning from data by using processes of knowledge seeking through said data, to obtain feasible information for decision-making (Cao, 2016).

There are different views and definitions of data science. Dhar (2013), Gelman et al. (2014 apud Lin Wang, 2018), Provost and Fawcett (2013), for example, associate DS with data analysis and the extraction of knowledge and information from data. Cao (2018) points out that several terms, such as "analytics", "big data", and data science, are extremely interconnected and are often confused with each other. Such a definition, as corroborated by Frické (2015) and Lowrie (2017), conceptualize DS more as a professional field than as a science, since it lacks a precise curricular definition, its own theoretical systems, and disciplinary structures, depending largely on methods and tools defined by other sciences.

In a similar view, Cleveland (2001) highlights that DS tends to focus only on computational aspects of data, which limits and possibly compromises its results. For the author, there is potential for the involvement of more areas, as data analysis involves more than just technology, so he suggests

that works in the area should be dedicated to multidisciplinary research, creating models and methods for data and theory building.

Library and information science, on the other hand, is a discipline with consolidated practical and theoretical aspects, whose object of study (information) is directly related to the concept of data and which, as seen in the results of this research, studies a significant number of topics commonly related to DS. So, it is believed that a more complex interdisciplinary relationship between LIS and DS is possible, considering that the main objective of data science is the the production of information and knowledge by data analysis (Provost; Fawcett, 2013). This possibility is also addressed by Virkus and Garoufallou (2020).

The changes in LIS due to technological changes and the advent of data science can already be felt. In the United States, for instance, most library and information science graduate programs were impacted by advances in information technology, changing the courses offered and with greater focus on providing students with technological knowledge, including data science related courses, to meet the skills and knowledge needed by information professionals to act in accordance with the current needs of libraries and information agencies (Hu, 2013).

This tendency is also perceived by Lin Wang (2018), with the caveat that, for library and information science to seize on data science opportunities, it is necessary to elucidate the relationship between the two disciplines so that data science can bring greater benefits to LIS. As the author points out, the other way around is also true, since theories, models, techniques and processes of LIS, a more consolidated science than DS, can also generate great contributions to the development of the latter.

As Pinheiro (2006) highlights, library and information science, due to its interdisciplinary nature, is in constant transformation in terms of its interdisciplinary relationships, resulting in new epistemological relationships that modify its boundaries. The focus of LIS on information and knowledge representation processes and their records, given the technological innovations in these processes and records, seems to lead the area to a new paradigm shift. Technologies such as mobile networks, smartphones, tablets, cloud storage, etc. – and the immense masses of data resulting from them – have changed information cycles, information supports and the relationships with information records (Souza et al., 2013), making it a necessity for LIS to adapt to this new reality.

Added to that, the scenario of technological advances together with changes in the perception of science and in knowledge production regimes – facts

that originated in the open science movement (Smart et al., 2019) – led to discussion about openness and sharing of scientific data to provide the creation of new knowledge, greater scientific participation, and possibilities for innovation. Data science, as a technological area dependent of digital data (especially open data) as its research objects, relates with the goals of open science and open data due to the importance of data reproducibility and transparency and the need for innovation for the area; besides that, constant technological advances and the demand for continuous improvement make knowledge sharing essential for data science development.

In view of the above, it is important to research the status of scientific production related to both library and information science and data science fields, to understand the influence and the existing relationships between the two disciplines, as well as verify the production of open access content related to the research theme. The evaluation and analysis of scientific production and communication have bibliometrics as one of its main instruments, allowing the study of scientific production, to identify and measure advances and trends in science.

Among the formal instruments of scientific communication, scientific journals are one of the main vehicles for disseminating knowledge, due to the shorter time for publicization compared to other vehicles and the credibility resulting from peer review (Ferreira, 2010). Therefore, the evaluation of articles published in these journals allows the study of scientific production by areas or subjects, thus being able to identify and estimate trends and impacts on science.

The need to evaluate and analyze scientific production and communication led, in 1922, to the creation of the "statistical bibliography", developed by Wyndham Hulme, while the modern term "bibliometrics" was created only in 1934, by Paul Otlet (Medeiros; Vitoriano, 2015; Momesso; Noronha, 2017), consisting in an area of study intended on quantifying the production and usage of recorded information. This technique consists of the main tool for measuring scientific knowledge based on the analysis of publications, thus having been adopted as a methodology for the development of this research, as detailed below.

Methodology

The study is a bibliometric analysis of publications related to LIS and data science. It sought to analyze the international production on the subject, to enable a broader analysis of common study trends between the two areas. Temporal filters were not adopted during the data collection, resulting in a sample that contains publications between the period of 1996 and 2021, noting that the search

was made in the beginning of the year 2022.

Bibliometrics is a technique widely used by LIS with the purpose of investigating communication processes, especially scientific communication, quantifying these processes. The bibliometric studies were developed from existing empirical laws, such as the Lotka's law, Zipf's law, Bradford's law, etc. (Braga, 1973; Guedes; Borschiver, 2005 apud Medeiros; Vitoriano, 2015).

Since one of the research objectives is the analysis of the most productive authors in the scientific production common to the two previously mentioned areas, Lotka's law, which states that the number of authors who make n contributions in each scientific field is approximately $1/n^2$ of those who make a single contribution (Alvarado, 2002; Cândido; Garcia; Campos; Tambosi Filho, 2018; Urbizagastegui, 2008), has special relevance for the study carried out.

To achieve the proposed goals, the methodology was divided into five steps. The first consisted of the identification and selection of databases, resulting in the choice of the Web of Science and Scopus databases. These databases were chosen because they are references for scientific communication in the most diverse areas of study, in addition to being international databases that allow a closer observation of international scientific production.

In the second step, parameters and strategies for database search where defined. The descriptors "data science", "e-science", "big data" and "data mining" were used, each one always in conjunction with "information science". Since the researched databases concentrate publications from several countries, the use of English descriptors was decided due to the great use of this language for international publications. The period of coverage of the research was not limited, since it was intended to study the period of the first publications relating LIS and data science, as well as the tendency of these publications over time. The search was limited to scientific publications having LIS as the main area of study, having been found publications in the period between 1996 and 2021.

In the search, a total of 232 different articles were found distributed among the selected databases, after disregarding articles found in duplication between databases.[1]

Of the 232 mentioned articles, the Scopus database presented the highest number of results with 156 titles, while 124 titles were found in the Web of Science database, totaling 48 titles found in both bases, as shown in Figure 1.

---

[1]   The dataset, containing the list of articles that compose the researched sample, is available in: https://zenodo.org/record/6670108.
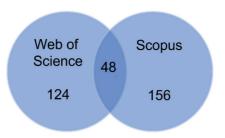
Figure 1 — Distribution of scientific production found by researched database. Source: research data, 2022

In the third step, the data processing was performed, consolidating the results of both databases, as well as exporting and tabulating these results. The exported results contained the following information: publication title, abstract, year of publication, authors, journal, and authors' institution. This information was then tabulated for analysis. The sample results were organized using electronic spreadsheets, edited using Excel and Google Sheets, both used for processing, calculation and subsequent generation of graphs related to the sample.

In the fourth step, the results were analyzed and depurated, identifying by titles, abstracts, and keywords those articles that suited the theme, parameters, and scope of the study, using as a requirement the need to contain in one or more of these fields the descriptors used in the previous step to identify that it was a LIS work addressing DS. At this stage, duplicate documents (present in more than one of the databases) were also identified, to disregard duplicity.

Finally, in the fifth step, the results were interpreted, identifying the number of publications by each author and institutions found, the quantity of publications per country (considering the country of origin related to the institution) and the quantity per journal. A further classification was made regarding the type of access of the journals in the sample, that is, whether the journal was open access, subscription-based or hybrid open access.

In cases of co-authorship, the individual information of each author and his/her institution was considered for quantitative calculations. For the articles produced by authors from the same institution, that institution and its country were considered only once in the quantitative calculation.

Based on the keywords of the articles, an analysis was carried out regarding the most discussed topics, to estimate the most common themes and subjects of the relationships found between DS and LIS, as shown in the results.

Scientific production connections between library and information science and data science

Of the searched databases, there was a greater concentration of publications in the Scopus database, with about 67.24% of the articles in the total universe being present in this database. However, the publications found in the Web of Science are not negligible, since 53.44% of the articles are in the database.

To outline an overview of scientific production in which the theme represented the connections between library and information science and data science, it was observed that the first publication found, according to the scope of the research, dates from 1996 and is indexed on Web of Science. It is the article "Historical perspectives on information science" by Cornish Tao, which, although published before the formal origin of the data science field, deals with the use of data mining to discover knowledge in databases; thus, considering that the automated search for information and knowledge creation is one of the fields of study for both DS and LIS, the cited article represents a possibility of intersection between the two sciences, which is why it was included in the results.

The next article dates from 2000, "Informetrics: a new area of quantitative studies", which relates data mining and information retrieval. In the period from 2000 to 2013, there was a somewhat irregular number of articles and a low average of publications (3.17 per year), except for the years of 2006 (with six articles found) and 2008 (with 11 articles identified). From 2014 onwards, a greater number of publications (average of approximately 24.12 articles/year) and an accentuated growth in publications correlating LIS topics with those of DS can be observed (Figure 2).
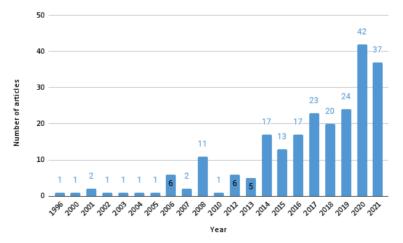

Figure 2 — Distribution of scientific production found by year of publication. Fonte: dados da pesquisa, 2022

As for authorship, a total of 569 different authors were identified, and those with the highest number of publications are shown in Table 1.

Table 1 — Authors with the highest number of publications, from 1996 to 2021

Number of articles	Author	Country
4	MUSIÑO, C. M (Dokumenta Consultoría e Integraciones)	Mexico
3	JONES, Kyle M. L. (Indiana University)	United States
	IMRAM, M. (The University of Queensland)	Australia
	PALMER, C. L. (University of Illinois)	United States
	SHAH, C. (University of Washington)	United States
	VANSCOY, A. (University at Buffalo)	United States
	ZHANG, Y. (Wuhan University)	China

Source: research data, 2022

The author with the highest number of publications identified was Celso Martínez Musiño (Dokumenta Consultoría and Integraciones) with four publications, followed by six authors with three publications and thirty-eight authors with two publications. 523 authors published a single article, thus, there is a wide diffusion of authors within the total universe of publications. Such dispersion can be explained by the fact that this is an interdisciplinary theme that allows a significant variety of contents and areas of knowledge, in addition to the fact that data science is still a new field. Lotka's law, using the counting methodology considering author and co-author(s), doesn't seem to be applicable to the sample, which may indicate that the correlation between these two areas is not at an advanced stage.

As for the authors' countries, there were a total of 42 countries, with the highest number of publications in the United States (76), followed by China (32), Brazil (25), Italy (18), and the United Kingdom (13). It is understood that even though the search for English descriptors allowed an overview of the international panorama of scientific research related to the topic, it favors countries whose native language is English. Even so, it is interesting to point out that Brazil appears with a significant number of articles, indicating not only a trend towards research about DS themes in the national LIS, but also an effort to publicize these studies

internationally, earning Brazil the spot of third greatest contributor and even placing it significantly ahead of some English-speaking countries.

Regarding the institutions linked to the authors, a total of 268 different institutions were identified, including universities, private entities, government agencies and research institutions. Table 2 orders these by highest number of publications.

Table 2 — Academic institutions with the highest number of publications

Institution	n. of published articles
Indiana University	8
University of Illinois	7
Wuhan University	7
Universidade de Brasília	6
Drexel University	5
Nanjing University	5
University of Texas	5

Source: research data, 2022

Among the institutions with publications identified in the research sample, Indiana University stands out with eight published articles. Also noteworthy are the Universities of Illinois and Wuhan (with seven articles published each) and the Universidade de Brasília with six publications. Among the seven institutions with the highest number of publications, the majority (four) are in the United States, and two Chinese institutions. Of the Brazilian institutions, the Universidade de Brasília stands out with six published articles, appearing as the fourth institution with the highest number of publications. As previously mentioned, when discussing the authors, the appearance of a Brazilian institution among those with the highest number of publications, shows a significant association of the research scenario in LIS with DS and area-related topics.

The articles found were published in a total of 111 different journals. Table 3 orders the top twelve publishers.

Table 3 — Scientific journals with the highest number of publications

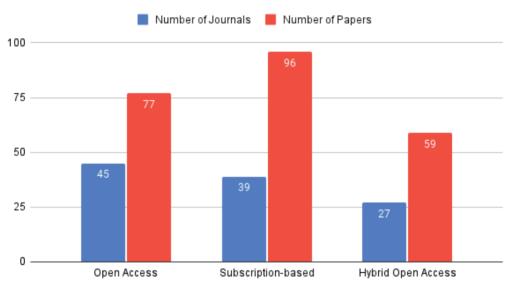| Journal	n. of published articles
Proceedings of the Association for Information Science and Technology	19
Journal of Information Science	17
Digital Libraries: The Era of Big Data and Data Science, IRCDL 2000	10
Journal of the Association for Information Science and Technology	9
Journal of Documentation	6
Advances in Classification Research Online	6
Perspectivas em Ciência da Informação	5
Online Information Review	5
Ciência da Informação	5
Library Hi Tech	5
Scientometrics	5
Journal of Library Science in China	5

Source: research data, 2022

The journal with the highest number of publications totaling 19 (8.19% of total publications) is the *Proceedings of the Association for Information Science and Technology*. The list also includes: *Journal of Information Science* (with 17 publications representing 7.33% of the total number of publications), *Digital Libraries: The Era of Big Data and Data Science*, IRCDL 2000 (with ten publications), *Journal of the Association for Information Science and Technology* (nine publications), *Journal of Documentation* and *Advances in Classification Research Online* (six publications each) and the journals *Perspectivas em Ciência da Informação*, *Online Information Review*, *Ciência da Informação*, *Library Hi Tech*, *Scientometrics* and *Journal of Library Science in China* with five publications each.
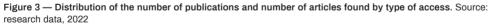
These 12 journals corresponded to 37.5% of the total number of publications, with a group of 40 journals with at least two publications (162 articles, or 69.83%) and another of 71 journals with a single article (corresponding to 30.60% of the total), thus showing a considerable concentration of articles in the first group and a wide dissemination of periodicals publishing about the subject of study.

Of 111 different journals found in the research data, only 45 (about 40.54%) are open access journals and 27 have hybrid open access (with

some open access publications and others with access limitations). Of the articles in the research sample, as shown in Figure 3, a significant number of 77 articles (about 33.19%) were published in open access journals and 59 (about 25.43%) in hybrid open access journals. Therefore, it is clear that open access journals are prolific in relation to the researched topic, even though they represent a significantly lower percentage than publications in subscription-based or hybrid open access journals. In a study such as this, in which accessibility to information is essential for the theme expansion and innovations development, this limitation is striking.

Figure 3 — Distribution of the number of publications and number of articles found by type of access. Source: research data, 2022

In the universe of open access journals, a single journal concentrated 26.25% of the total open access publications and the other 11 journals with the highest number of publications corresponded to 66.25% of the total (Table 4). It shows a large concentration of articles in a small number of journals.

Table 4 — Open access scientific journals with the highest number of publications

Journal	n. of published articles
Advances in Classification Research Online	6
Perspectivas em Ciência da Informação	5
Journal of Library Science in China	5
Ciência da Informação	5
Library Philosophy and Practice	4

Source: research data, 2022.

The open access journals found by the research corresponded to 40.54% of the total number of journals, with 24.32% of the total corresponding to mixed access journals and 35.13% with subscription access. Thus, a considerable percentage of open access journals compose the research universe.

Is significant that, among the open access journals with the highest number of publications, two titles are Brazilian, the journals *Perspectivas em Ciência da Informação* and *Ciência da Informação*, both with five publications. Also, most publications by Brazilian authors were made in open access journals.

The results presented seem to indicate the existence of strong possibilities for research correlating LIS and DS, in different areas of study, as discussed below.

Trends in related studies between library and information science and data science

The subjects and areas covered by the articles present a great variety, with a significant number presenting intersections of LIS with administration, education, medicine, etc. To this end, applications or theories of data science were also addressed.

From the sample, it is possible to identify study trends correlating LIS and DS, especially by analyzing roles that information professionals can assume in the context of activities related to data or curricular changes in LIS resulting from the advent of DS, this being the most debated theme in the sample and found in 34 of the articles. Furthermore, it was found a significant number of articles dedicated to the analysis of the predominance of LIS themes over DS ones, totaling 22 articles.

It is also possible to perceive possibilities of applications in scientific communication, especially for scientific production metrics (28 articles related to metrics were found).

Two topics of common interest between the two areas presented a significant number of studies in the researched sample, these being data mining (with a total of 26 articles dealing directly with this topic) and big data (with 24 articles). The results show a tendency to use data mining and big data associated to metrics studies (as aids in the information exploration process), knowledge organization and knowledge representation, in which LIS techniques appear as conductors, to allow better use of information in the fields mentioned, as several articles were also found aiming to understand the domains of these two subjects in the field of LIS.

Studies focused on information literacy competence in the context of data and knowledge management are significant in the sample, with 16 and 15 articles, respectively. There was also a tendency of studies focused on classification subareas (12 articles identified) and 11 articles referring to digital libraries, iSchools and library-oriented applications.

Other topics were also approached, such as knowledge representation, user studies, e-Science, information and recommendation systems, open data, information representation and data science applications in the fields of design and information retrieval.

Figure 4 shows the most discussed topics relating to the total research sample.

It is understood, considering the limited number of articles relating the two fields of study and the different approaches found even in similar themes, that these findings better represent possible tendencies of correlated studies than effective connections between the areas. At the same time, the diversity of the collected sample may indicate the existence of ample possibilities for collaborations between library and information science and data science.
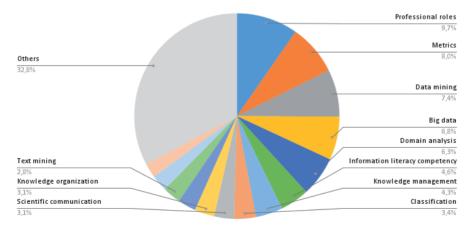

Figure 4 — Percentage of the main themes found in relation to the total sample. Source: research data, 2022

While the diversity of themes seems to point to an inconstant and non-systematic approach between LIS and DS, it also seems to indicate that the possibilities of intersections between them are extensive.

Final considerations

The analysis made in the present bibliometric study allows to observe relevant aspects about the theme of data science produced by researchers in information science, whose scientific production has increased in the last six years, which may indicate a growing interest of LIS in themes related to DS practice.

The intersection between fields of LIS and DS seems to be a growing research trend, which can be perceived by the increase in interconnected scientific production in recent years and the number of countries of origin of the articles found. Although many articles were produced by institutions in the United States, it is understood that this result is directly related to the methodology used of searching through English-based descriptors and databases, since a significant amount of the scientific production isn't published in that language; however, it is also understood that this methodology allowed the analysis of a global scene.

The significant number of articles addressing changes related to DS in the curricula of LIS and Librarianship programs or the emergence of new acting areas for information professionals because of DS denote that the changes introduced by the current technological evolution and the emergence of DS has been causing interest to LIS.

These changes and new research trends are becoming more prominent and should allow the emergence of new research possibilities, as well as new correlations with data science and possibly changes in the interdisciplinary relationships of library and information science with technological areas, especially with computer science, based on the growing approach of DS objects of study, such as big data, data mining, cloud storage, social networks, etc. by LIS.

The study found a wide range of themes in the interface between data science and library and information science, thus being able to identify possible areas of study correlating both fields of knowledge, such as professional roles, scientific communication metrics, data mining, big data, content analysis, information literacy competence, and knowledge management.

It is understood that the sample size found by this study is still concise, despite the growing trend identified. Thus, there is potential for the development of studies that more fully explore the perspectives of relationship between the areas, especially considering the similarities between the objects of study and

the interdisciplinary nature of both. The research enabled to identify some of these possibilities, and so aims to contribute to the discussion in library and information science about its new roles in the denominated age of information and digital data.

The increasingly central role of information and knowledge in contemporary society leads to the adoption of new roles by library and information science, allowing greater dialogue with other disciplines, due to the interdisciplinary essence of the field. This scenario also reinforces the importance of open science in this environment of innovation as a facilitator to the sharing of scientific research, data and even computer code, promoting greater cooperation between researchers and the public for informational and scientific progress.

References

ALVARADO, R. U. A Lei de Lotka na bibliometria brasileira. *Ciência da Informação*, Brasília, v. 31, n. 2, p. 14-20, ago. 2002. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19652002000200002&lng=en&nrm=iso. Acesso em: 27 maio 2022.

BRAGA, G. M. Relações bibliométricas entre frente de pesquisa (*research front*) e revisões da literatura: estudo aplicado à ciência da informação. *Ciência da Informação*, Brasília, v. 2, n. 1, p. 9-26, 1973. Disponível em: http://revista.ibict.br/ciinf/article/view/20. Acesso em: 27 maio 2022.

CÂNDIDO, R. B.; GARCIA, F. G.; CAMPOS, A. L. S; TAMBOSI FILHO, E. Lei de Lotka: um olhar sobre a produtividade dos autores na literatura brasileira de finanças. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação*, v. 23, n. 53, p. 1-15, set./dez. 2018. Disponível em: https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2018v23n53p1/37239. Acesso em: 27 maio 2022.

CAO, L. Data science: a comprehensive overview. *ACM Computing Surveys*, v. 50, n. 3, jun. 2018. Disponível em: https://dl.acm.org/doi/10.1145/3076253. Acesso em: 27 maio 2022.

_____. Data science and analytics: a new era. *International Journal of Data Science and Analytics*, v. 1, n. 1, p. 1-2, abr. 2016. Disponível em: https://link.springer.com/article/10.1007/s41060-016-0006-1. Acesso em: 27 maio 2022.

CLEVELAND, W. S. Data science: an action plan for expanding the technical areas of the field of statistics. *International Statistical Review*, v. 69, n. 1, p. 21-26, abr. 2001. Disponível em: https://doi.org/10.1111/j.1751-5823.2001.tb00477.x. Acesso em: 27 maio 2022.

CONEGLIAN, C. S.; DIEGER, R.; SANTARÉM SEGUNDO, J. E.; CAPTREZ, M. O papel estratégico da web semântica no contexto do *big data*. In: WORKSHOP DE INFORMAÇÃO, DADOS E TECNOLOGIA, 1., Santa Catarina, 2017. *Anais ...* Santa Catarina: Universidade Federal de Santa Catarina, 2017. Disponível em: https://repositorio.ufsc.br/handle/123456789/180289. Acesso em: 27 maio 2022.

DHAR, V. Data science and prediction. *Communications of the ACM*, v. 56, n. 12, p. 64-73, 2013. Disponível em: https://dl.acm.org/doi/10.1145/2500499. Acesso em: 27 maio 2022.

FERREIRA, A. G. C. Bibliometria na avaliação de periódicos científicos. *DataGramaZero: Revista de Ciência da Informação*, v. 11, n. 3, 2010. Disponível em: https://brapci.inf.br/index.php/res/v/7110. Acesso em: 27 maio 2022.

FRICKÉ, M. Big data and its epistemology. *Journal*

of the Association for Information Science and Technology, v. 66, n. 4, p. 651-661, 2015. Disponível em: https://asistdl.onlinelibrary.wiley.com/doi/full/10.1002/asi.23212. Acesso em: 27 maio 2022.

HU, S. Technology impacts on curriculum of Library and Information Science (LIS): a United States (US) perspective. *Libres*: Library and Information Science Research Electronic Journal, v. 23, n. 2, p. 1-9, set. 2013. Disponível em: https://www.libres-ejournal.info/1033/. Acesso em: 27 maio 2022.

LIN WANG. Twinning data science with information science in schools of library and information science. *Journal of Documentation*, v. 74, n. 6, p. 1.243-1.257, 2018. Disponível em: https://www.emerald.com/insight/content/doi/10.1108/JD-02-2018-0036/full/html. Acesso em: 27 maio 2022.

LOWRIE, I. Algorithmic rationality: epistemology and efficiency in the data sciences. *Big Data & Society*, v. 4, p. 1-13, jan./jun. 2017. Disponível em: https://journals.sagepub.com/doi/full/10.1177/2053951717700925. Acesso em: 27 maio 2022.

MEDEIROS, J. M. G.; VITORIANO, M. A. V. A evolução da bibliometria e sua interdisciplinaridade na produção científica brasileira. *RDBCI*: Revista Digital de Biblioteconomia e Ciência da Informação, v. 13, n. 3, p. 491-503, set./dez. 2015. Disponível em: https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/8635791. Acesso em: 27 maio 2022.

MOMESSO, A. C.; NORONHA, D. P. Bibliométrie ou bibliometrics: o que há por trás de um termo? *Perspectivas em ciência da informação*, v. 22, n. 2, abr./jun. 2017. Disponível em: https://www.scielo.br/j/pci/a/X4xTbMZNdVXV3MNsYFRnSbQ. Acesso em: 27 maio 2022.

PINHEIRO, L. V. R. Ciência da informação: desdobramentos disciplinares, interdisciplinaridade e transdisciplinaridade. In: GONZÁLEZ DE GÓMEZ, M. N.; ORRICO, E. G. D. (org.). *Políticas de memória e informação*: reflexos na organização do conhecimento. Natal: Editora UFRN, 2006. p. 111-141. Disponível em: http://ridi.ibict.br/handle/123456789/18. Acesso em: 27 maio 2022.

PROVOST, F.; FAWCETT, T. Data science and its relationship to big data and data-driven decision making. *Big Data*, v. 1, n. 1, p. 51-59, 2013. Disponível

em: https://www.liebertpub.com/doi/10.1089/big.2013.1508. Acesso em: 27 maio 2022.

SARACEVIC, T. Information science. *Journal of the American Society for Information Science*, v. 50, n. 12, p. 1.051-1.063, 1999. Disponível em: https://asistdl.onlinelibrary.wiley.com/doi/epdf/10.1002/%28SICI%291097-4571%281999%2950%3A12%3C1051%3A%3AAID-ASI2%3E3.0.CO%3B2-Z. Acesso em: 27 maio 2022.

SMART, P. et al. Open science and open innovation in a socio-political context: knowledge production for societal impact in an age of post-truth populism. *R&D Management*, v. 49, n. 3, p. 279-297, 2019. Disponível em: https://onlinelibrary.wiley.com/doi/full/10.1111/radm.12377. Acesso em: 27 maio 2022.

SOUZA, R. R.; ALMEIDA, M. B.; BARACHO, R. M. A. Ciência da informação em transformação: *big data*, nuvens, redes sociais e web semântica. *Ciência da Informação*, v. 40, n. 2, p. 159-173, 2013. Disponível em: https://revista.ibict.br/ciinf/article/view/1379. Acesso em: 27 maio 2022.

URBIZAGASTEGUI, R. A produtividade dos autores sobre a Lei de Lotka. *Ciência da informação*, v. 37, n. 2, p. 87-102, maio/ago. 2008. Disponível em: https://revista.ibict.br/ciinf/article/view/1214. Acesso em: 27 maio 2022.

VIRKUS, S.; GAROUFALLOU, E. Data science and its relationship to library and information science: a content analysis. *Data Technologies and Applications*, v. 54, n. 5, p. 643-663, 2020. Disponível em: https://doi.org/10.1108/DTA-07-2020-0167. Acesso em: 27 maio 2022.