

DOI: 10.35643/Info.28.2.9

Article

Scientific and technological connections between specialists in injectable oncology: a contribution to public health from the analysis of the Lattes database

Conexiones científicas y tecnológicas entre especialistas en oncología inyectable: un aporte a la salud pública a partir del análisis de la base de datos Lattes

Conexões científicas e tecnológicas entre especialistas em oncologia injetável: uma contribuição para a saúde coletiva a partir da análise da Plataforma Lattes

Henrique Koch Chaves¹ ORCID: [0000-0003-3035-6799](https://orcid.org/0000-0003-3035-6799)

Francisco Theodoro Arantes Florencio¹ ORCID: [0000-0002-2312-9603](https://orcid.org/0000-0002-2312-9603)

Carla Cristina de Freitas da Silveira¹ ORCID: [0000-0001-7376-7948](https://orcid.org/0000-0001-7376-7948)

Adelaide Maria de Souza Antunes^{2,3} ORCID: [0000-0002-2245-7517](https://orcid.org/0000-0002-2245-7517)

Jesús Pascual Mena-Chalco⁴ ORCID: [0000-0001-7509-5532](https://orcid.org/0000-0001-7509-5532)

Jorge Lima de Magalhães^{1,5,6} ORCID: [0000-0003-2219-5446](https://orcid.org/0000-0003-2219-5446)

¹ Fundação Oswaldo Cruz/FIOCRUZ – Instituto de Tecnologia em Fármacos/Farmanguinhos, Brasil, Rua Sizenando Nabuco, 100, Manguinhos, CEP: 21041180, Rio de Janeiro, Rio de Janeiro, henrique.chaves@fiocruz.br, francisco.florencio.tjii@gmail.com, carla.silveira@fiocruz.br, jorge.magalhaes@fiocruz.br

² Instituto Nacional da Propriedade Industrial, Brasil, Rua Mairink Veiga, 9, Centro, CEP: 20090-910, Rio de Janeiro, Rio de Janeiro.

³ Universidade Federal do Rio de Janeiro, Escola de Química, Brasil, Avenida Athos da Silveira Ramos, 149, CEP: 21941-909, Cidade Universitária da Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro, adelaide@eq.ufrj.br

⁴ Universidade Federal do ABC, Centro de Matemática, Computação e Cognição, Brasil, Avenida dos Estados, 5001, Bangu, CEP: 09210-580, Santo André, São Paulo, jesus.mena@ufabc.edu.br

⁵ Global Health and Tropical Medicine, GHTM, Instituto de Higiene e Medicina Tropical, IHMT, Universidade NOVA de Lisboa – UNL, Portugal, Rua da Junqueira, 100, 1349-008, Lisboa.



⁶ International Platform for Science, Technology and Innovation in Health (PCTIS), University of Aveiro, Portugal, Campus Universitário de Santiago, 3810-193, Aveiro.

Abstract

With the exponential increase in the amount of data generated daily, the interest of organizations in developing strategies for handling Big Data has intensified. An example of Brazilian Big Data is the CNPq Lattes Platform, which allows extracting information from the curricula of researchers registered therein. One of the limitations on the platform is that the information is arranged individually for each researcher and without an indexing standard, making it impossible to extract data from researchers simultaneously. To remedy this problem, the ScriptLattes software was developed, used in the present study, to extract information from the curricula of researchers in injectable oncology, to obtain a panoramic view of their technical and scientific productions. Regarding the activities carried out by these specialists in the analyzed period, a total of: 85,465 published complete articles, 24,068 technical productions and 13,496 implemented projects were observed. Networking analyzes performed applying Gephi software pointed Dr. Fernando de Queiroz Cunha, from the Faculty of Medicine of Ribeirão Preto (USP), as the most prominent researcher in the area. Despite the limitations of the Lattes Platform, the results proved to be efficient, enabling the generation of scientific and technological knowledge in injectable oncology at various institutional levels.

Keywords: INFORMATION MANAGEMENT; LATTES PLATFORM; CANCER; INJECTABLE ONCOLOGY; BIG DATA, SCRIPTLATTES.

Resumen

Con el aumento exponencial de la cantidad de datos generados diariamente, se ha intensificado el interés de las organizaciones por desarrollar estrategias para el manejo de Big Data. Un ejemplo de Big Data brasileño es la Plataforma CNPq Lattes, que permite extraer información de los currículos de los investigadores registrados en ella. Una de las limitaciones de la plataforma es que la información se ordena individualmente para cada investigador y sin un estándar de indexación, lo que hace imposible extraer datos de los investigadores simultáneamente. Para remediar este problema, se desarrolló el software ScriptLattes, utilizado en el presente estudio, para extraer información de los currículos de investigadores en oncología inyectable, con el fin de obtener una visión panorámica de sus producciones técnico-científicas. En cuanto a las actividades realizadas por estos especialistas en el período analizado, se observaron un total de: 85.465 artículos completos publicados, 24.068 producciones técnicas y 13.496 proyectos implementados. Los análisis de red realizados con la ayuda del software Gephi señalaron al Dr. Fernando de Queiroz Cunha, de la Facultad de Medicina de Ribeirão Preto (USP), como el investigador más destacado del área. A pesar de las limitaciones de la Plataforma Lattes, los resultados demostraron ser eficientes, posibilitando la generación de conocimiento científico y tecnológico en oncología inyectable en varios niveles institucionales.

Palabras clave: GESTIÓN DE LA INFORMACIÓN; PLATAFORMA LATTES; CÁNCER; ONCOLOGICOS INYECTABLES. BIG DATA; SCRIPTLATTES.

Resumo

Com o aumento exponencial na quantidade de dados gerados diariamente, têm-se intensificado o interesse das organizações em desenvolver estratégias para o tratamento do Big Data. Um exemplo de Big Data brasileiro é a Plataforma Lattes do CNPq, que permite extrair informações dos currículos dos pesquisadores nela cadastrados. Uma das limitações na plataforma é que as informações são dispostas individualmente para cada pesquisador e sem um padrão de indexação, impossibilitando a extração dos dados dos pesquisadores simultaneamente. Para sanar tal problema, foi desenvolvido o software ScriptLattes, utilizado no presente estudo, para extrair as informações dos currículos dos pesquisadores em oncológicos injetáveis, visando obter uma visão panorâmica das produções técnicas e científicas destes. Em relação às atividades realizadas por esses especialistas no período analisado, observou-se um total de: 85.465 artigos completos publicados, 24.068 produções técnicas e 13.496 projetos implementados. As análises de rede realizadas com o auxílio do software Gephi apontaram o Dr. Fernando de Queiroz Cunha, da Faculdade de Medicina de Ribeirão Preto (USP), como o pesquisador mais proeminente na área. Apesar das limitações da Plataforma Lattes, os resultados mostraram-se eficientes, possibilitando a geração de conhecimento científico e tecnológico em oncológicos injetáveis em vários níveis institucionais.

Palavras-chave: GESTÃO DA INFORMAÇÃO; PLATAFORMA LATTES; CÁNCER; ONCOLÓGICOS INJETÁVEIS; BIG DATA; SCRIPTLATTES.

Received: 02/04/2023

Accepted: 09/08/2023

1. Introduction

According to data from the Brazilian *Ministry of Health* (MoH) (Ministério da Saúde – Brazilian term) and the *National Cancer Institute* (INCA – Brazilian term - Instituto Nacional do Cancer), the number of deaths caused by neoplasms is among the three main causes of death in Brazil (Bray et al., 2018; INCA, 2019). In the world, according to *the International Union for Cancer Control* (UICC), in the year 2020, there were about 15 million new cases of cancer, leading to 12 million deaths worldwide (Bray et al., 2018; INCA, 2019).

Currently available anticancer therapies include chemotherapeutic agents, biologics, molecular target therapy, radiotherapy, surgery, and interventional oncology (Ramos et al., 2021; Sag et al., 2016). In addition to therapy with drugs that act directly against various types of cancer, other drugs are used to minimize the toxicities caused by anticancer drugs: antiemetics, urinary protectors, corticoids and intravenous hydration, including, when indicated, red blood cell and platelet transfusions, antibiotics and growth factors (Matz & Hsieh, 2017; Oun et al., 2018).

The handling of injectable drugs is an activity of great importance and complexity developed by the pharmaceutical sector, since these drugs are not commercially available, with the possibility of customizing the preparation to meet the specific needs of the patient: individualization of the dose, adequacy of the formulation to the route of administration, adding components to the formulation and choosing the type of diluent volume appropriate to the patient's clinical condition (ASHP, 2014; Mohiuddin, 2020; Pergolizzi Jr et al., 2013). One of the great challenges of chemotherapy is the selective delivery of drugs to tumour cells with minimal interaction with healthy tissues (Webster et al., 2014)

In recent decades, due to the enormous amount of information constantly generated in the most diverse contexts of society, information has come to be classified as a strategic resource with price, cost and value (Valentim et al., 2006). In the business environment, information is considered the most important input in the decision-making process, and it is necessary to analyse information from various aspects: administrative, economic, technical-scientific, marketing, legal, environmental and political (Contador, 2010).

Information has been gaining increased relevance both in public and private organizations, which is why it is necessary to work on the Information Management process also in the sphere of public health organizations, which over the years have sought to develop, innovate and improve their services. in the face of the new society in which they find themselves. Thus, information should be considered as a fundamental part of the work practices of public health organizations and in decision-making processes (Santos & Damian, 2019).

To process the large volume of Big Data at high speed, the traditional methods previously used in the health area have shown low applicability, reinforcing the need for facilitating tools for the analysis of this data (Galvão & Valentim, 2019). An example of Big Data is the Brazilian scientific repository, the CNPq Lattes Platform, which integrates the academic curricula of professionals from all areas of knowledge, both Brazilians and foreign researchers who at some point developed projects in partnership with Brazilians researchers (A. G. C. de Brito et al., 2016). This information has been essential for technical-scientific strategies in Brazil, as they make it possible to study and reflect on scientific and technological advances in all areas of Science, Technology and Innovation.

Since the resume extraction system on the Lattes Platform occurs through the individual search by name of the person registered on the platform, a computational tool called ScriptLattes was developed, which allows the extraction and recovery of several data simultaneously (A. G. C. de Brito et al., 2016; Magalhães et al., 2020; Mena-Chalco & Junior, 2009). This program is open source based and presents, organizes and presents various information from the curricula (CVs) of registered researchers such as scholarly productions, guidelines, collaborations, geolocation, among others (Ferraz et al., 2018).

ScriptLattes has been presenting advanced results in the management of Big Data information in different areas of knowledge. Several studies using ScriptLattes are found in the Scielo, Web of Science and Google Scholar databases, such as: i) the analysis of the scientific production of researchers in neglected tropical diseases, nanotechnology, pediatrics, philosophy, physical education, genetics and reproduction freshwater fish; ii) evaluate the evolution of scientific production and collaboration networks in Brazilian regions; iii) in the analysis of the scientific production of Brazilian graduate programs, among others.(Alves et al., 2016; Fernandes & Silva, 2018; Ferraz et al., 2014, 2019; Iriart & Trad, 2020; Klepa & Pedroso, 2019; Nigro et al., 2015; Pedroso et al., 2016; Silva & Viegas, 2017; Sobral et al., 2020).

ScriptLattes is a specific open-source tool, and, in this context, the references are exclusive to the scientific and technological environment. In this context, its use, testing and academic reflections aim at constant and open construction together

with its developers. In this sense, projects were successfully executed, highlighting opportunities and challenges both in the Lattes base and in the constant improvement of ScriptLattes.

In this sense, the scientific literature shows the use of ScriptLattes to analyse the curricula of researchers in the field of oncology (Nascimento, 2020; Nascimento & Gouveia, 2021). As there is no scientometric study in the literature addressing the use of ScriptLattes in the specific field of injectable oncology, the present study was carried out to evaluate the feasibility of using the Lattes Platform and the ScriptLattes program in the management of Big Data technologies in the area in question, in which the authors of this article decided to identify and retrieve the CVs of senior specialists registered on the platform in Brazilian territory. The knowledge generated with the ScriptLattes analyses will make it possible to analyse different aspects of Brazilian science, both from a micro perspective (individual researcher and groups of researchers) and macro (collaboration network or representativeness of the area), thus allowing to identify or validate patterns of activities academic, which will result in bibliometric information about the group of researchers in question (A. G. C. D. Brito et al., 2016; Nicholson, 2006).

2. Materials and Methods

The research was based on scientometrics and bibliometrics of shared open data in science. The free open-source program ScriptLattes in a Linux environment was used to extract the open data on the Brazilian Lattes platform. Among the filters available on the Lattes platform for data extraction, the ones listed below were selected in order to identify researchers with greater academic experience:

- a) academic level (doctorate);
- b) researcher's nationality (Brazilian and foreign);
- c) CNPq productivity fellows (all categories);
- d) presence in the directories of Brazilian Research Groups (DGP, <https://lattes.cnpq.br/web/dgp>).

In order to preliminarily identify the number of CVs available in each search strategy, in May 2022, searches were carried out in the “advanced” field of the Lattes Platform using terms in both English and Portuguese to assemble a search string: “(oncologic AND injectable), (oncology AND injectable), (cancer AND injectables), (cancer AND injectables), (antineoplastics AND injectables), (injectable AND neoplasm), (injectables AND neoplasia)”. Filters associated with the category “CNPq Productivity Fellows” and “Presence in the Directory of Research Groups” were also applied. Identification and extraction took place in March 2022. Each specialist was categorized by the Lattes ID, a 16-digit identification code for the CVs available on the platform (Mena-Chalco et al., 2014). To be able to use Scriptlattes, it was necessary to create a list of these CVs, listing a list with the ID number of each researcher on the Lattes platform (Mena-Chalco & Junior, 2009). The data were grouped, organized, processed and later made available on the web in HTML format. The variables analyzed in this work were scientific productions (articles published in journals), research projects, events, guidelines and geolocation data.

Social network analyzes (SAR) were performed with the aid of free open-source software Gephi[®], which allows visualization and exploration for all types of graphs and networks. The software in question allows easy and broad access to network data, enabling the import, visualization, filtering, browsing and grouping of data (clustering) (Bastian et al., 2009). Social Network Analyzes (SAR) focused on scientific collaboration allow the analysis of several indicators, such as co-authorship networks. These networks are formed when two or more researchers publish work together. The SAR has been used more and more in the scientific field, since they allow identifying and understanding skills, interests, demands and gaps present in the various fields of knowledge. These networks can be identified through co-written documents (Ferreira, 2009; Hayashi et al., 2012). In this type of graph, the size of the vertex represents the number of connections of each researcher, while the color represents one of the communities identified by their form of collaboration.

Two metrics were analyzed with the help of Gephi[®]: modularity and degree centrality. The degree centrality will indicate which are the authors who

collaborated the most, publishing jointly with other authors, taking into account the number of co-authors who collaborated with a given author, together with the number of publications that they made in partnership (Bordin et al., 2014). In complex networks, modularity is one of the metrics used to detect communities, groups or clusters. Modularity is a measure of the network as a whole, dividing the network into communities, according to the strength of the connections between the various vertices. Vertices more connected to each other than to the others are included in the same group. In this way, it can be inferred that a complex network with a high degree of modularity has a strong community structure, that is, the community vertices have a consistent connection with each other and a sparse connection with other communities (Vincenzo, 2008).

3. Results and Discussion

Table 1 presents the results obtained on the Lattes Platform using the search string cited in the methodology. Comparing the results obtained with the same search terms in English and Portuguese, a greater number of CVs in Portuguese was observed in all cases. A possible explanation is due to the fact that most of the CVs registered on the Lattes Platform are from Brazilian researchers, in which they index their information on the platform in the Portuguese language.

As can be seen in the table, when applying the filter “Brazilian and foreign doctors”, the number of CVs decreased substantially, reaching in some cases up to 1/4 of those found initially. It was also observed that when adding the filter “CNPq Productivity Scholars,” again there was a large decrease in relation to the previous filter. By adding the filter “Presence in the Directory of Search Groups”, the number of resumes remains practically constant in all searched terms. Another fact observed is that when using the search terms “cancer,” “oncology” and “antineoplastics” with the term “injectables”, joined by the Boolean operator “AND”, the same results were obtained in all cases.

Among the Boolean expressions used, “injectables AND neoplasia” was the one that presented the best results. In this way, it was decided to analyse, with the help of the ScriptLattes program, the curricula of senior researchers using the search filters “Brazilian and foreign doctors with any level of productivity scholarship

and presence in the GP directory”, adding a total of 535 resumes. The results with scientific and technological information on the competencies listed in the objective in full are available at the following URL address:
https://pesquisa.ufabc.edu.br/cientometria/oncologicos_injetaveis/

Table 1: Number of CVs of researchers identified in the Lattes database containing the search terms listed in the method.

Criteria selected for searching the Lattes Platform	(oncologic AND injectates)	(oncology AND injectable)	(cancer AND injectates)	(cancer AND injectables)	(antineoplastics AND injectates)	(injectable AND neoplasm)	(injectates AND neoplasia)
All (doctors and other researchers, Brazilians and foreigners)	4,549	814	4,549	44	4,549	1,809	27,174
Brazilian and Foreign Doctors	1,315	596	1,315	33	1,315	1,019	8,819
Brazilian doctors	1,305	567	1,305	31	1,305	1,000	8,721
Brazilian and foreign doctors with any level of productivity scholarship	108	102	108	7	108	88	568
Brazilian doctors with any level of productivity scholarship	106	99	106	7	106	88	558
Brazilian and foreign doctors with any level of productivity grant and presence in the DGP	103	93	103	7	103	88	535
Brazilian doctors with any level of productivity grant and presence in the DGP	101	91	101	7	101	85	526

Source: prepared by the authors based on data analysis/extraction form the Lattes Platform (2022).

Below, in Figure 1, the main page generated by ScripLattes in HTML format can be seen. The header of the page shows the links in which you can obtain relevant information about the experts analyzed, such as bibliographic production, technical production, guidelines, projects, among others.

Figure 1: Main Menu generated by ScripLattes.

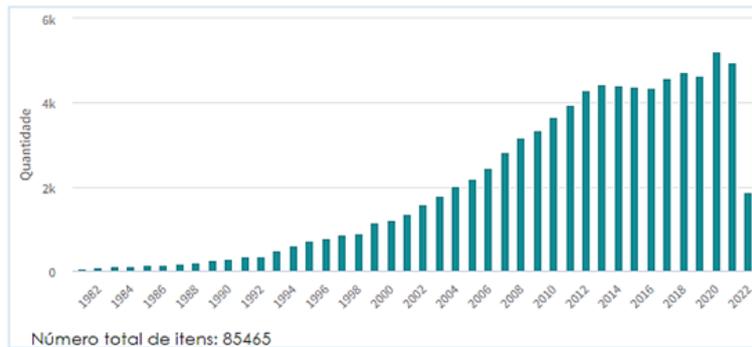


Source: Own elaboration screen capture using the ScripLattes tool, 2022.

The results obtained shows a large number of publications in the studied area. The bibliographical production of specialists concerns: papers, books, book chapters, among others. The amount identified and extracted were 264,332 products. It should be noted that, in case the analyzed researchers have carried out a publication in partnership, ScriptLattes will extract the single publications of the analyzed group, that is, it will not duplicate the information. Therefore, there may be some publications and/or technical products, which appear simultaneously listed in other senior specialists, who were listed here according to the methodology of this work. It is worth noting, once again, that ScriptLattes is constantly being updated, just like any other IT tool. However, it does not invalidate the relevance of the core competences in the area of injectable oncology identified here.

Figure 2 shows the evolution of the number of papers published by experts in the field since 1981. The published total was 85,465. It should be noted a substantial increase in 2020, where only in this year was published a total of 5,234 manuscript. In addition, it is possible to access the production by year on this page to have the list of publications sorted alphabetically by the title of the published article. For further details of the scientific production of each specialist, it is possible to access the work on the Website: https://pesquisa.ufabc.edu.br/cientometria/oncologicos_injetaveis/PB-0.html

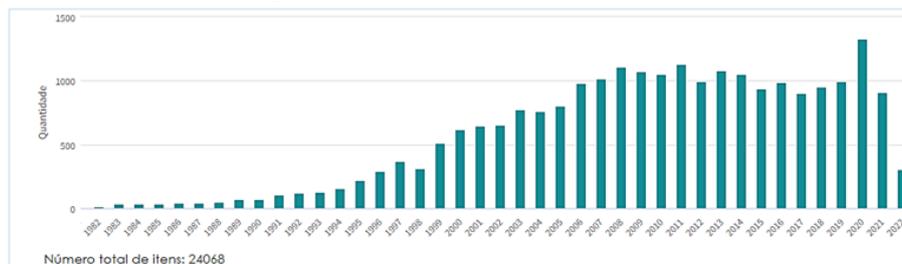
Figure 2. Evolution in the number of complete articles published in the researchers' journals analyzed by ScriptLattes between 1981 and 2022



Source: Elaborated by the authors using the ScriptLattes tool, 2022.

Regarding the technical production of the researchers analyzed with ScriptLattes, there have been increases and decreases over the years, adding up to a total of 24,068 items, among technological products (316), processes or techniques (421), technical works (11,186) and other types of technical production (12,145) (Figure 3). The first technical productions date back to 1982, with a total of 13 productions. Again, the year 2020 stands out in relation to the rest, adding up to a total of 1,331 technical productions.

Figure 3: Evolution in the technical production of researchers analyzed by ScriptLattes between 1982 and 2022.

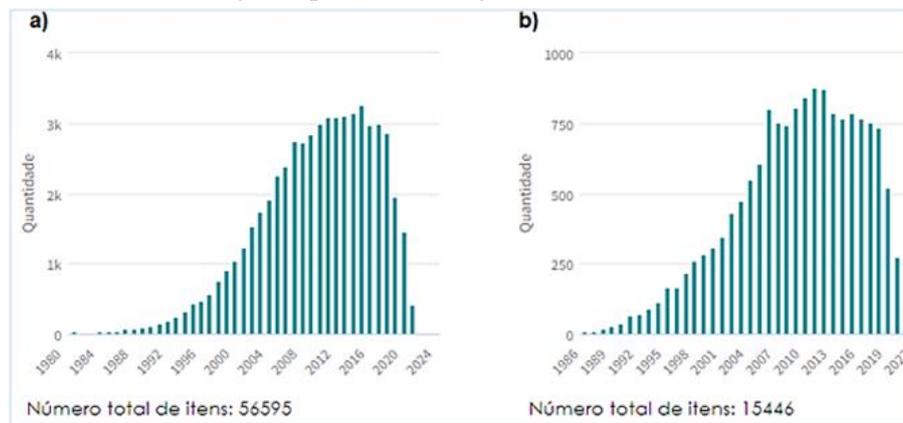


Source: extracted by the authors using the ScriptLattes tool, 2022.

In the “Orientations” (advisors) menu, it is possible to observe the orientations and supervisions in progress and concluded by these researchers, including work orientations for conclusion of graduation and specialization course (TCC), scientific initiation, masters, doctorate, postdoctoral and research orientations of another nature. When the searches were carried out on the Lattes Platform, there were a total of 4,999 supervisions/guidance in progress and 56,595 completed (Figure 4a). There is an increase in the number of completed orientations over the years, with a maximum in 2016 (3,267), followed by a decrease in the following years. Among the orientations in progress, there is a greater number of

orientations for doctoral students (2,152) while in the completed orientations, a greater number of orientations for scientific initiation students (15,446) are observed (Figure 4b). The first completed guidelines among the analyzed researchers date back to 1982, totaling 33 guidelines.

Figure 4: Evolution of number of orientations of researchers analyzed by ScriptLattes between 1982 and 2022: a) total of completed orientations; b) total number of completed scientific initiation orientations



Source: extracted by the authors using the ScriptLattes tool, 2022.

The results obtained show a large number of events held over the years, adding a total of 7,687 between the years 1984 and 2022. There is an increase in events year after year, with emphasis on the year 2011, with a total of 175 events. As of 2019, there is a significant drop in the number of events, decreasing from 168 to 116 events. This fact may be associated with the Covid-19 pandemic, which was first reported in December 2019 in China. In this way, many events were canceled or started to be held remotely through videoconferencing platforms such as Google Meet®, Microsoft Teams®, Zoom®, Cisco Webex® e GoToMeeting® (Agbehadji et al., 2020; Esakandari et al., 2020).

Regarding the projects implemented by the analyzed researchers, a total of 13,496 were found in the analyzed period. There is an increase in the number of projects implemented from 1983 (8) to 2010 (877), followed by a consecutive drop in the following years. In the year 2021, only 221 projects had been implemented by these researchers. One of the possible explanations is due to the decrease in investments in education in Brazil between 2014 and 2018, which decreased by around 56% and in the current government suffers even more expressive cuts (Mazieiro, 2019; SENADO NOTÍCIAS, 2020). To get an idea of the size of the

cuts, in October of that year (2022), the current government blocked R\$2.4 billion from the Ministry of Education (MEC) budget, impacting the activities of the ministry, which include federal institutes of education and universities (REVISTA GALILEU, 2022; Saldaña, 2022)

The geolocation map of senior injectable oncology specialists can be identified in Figure 5. The green symbols on the map identify where each researcher works, that is, their professional address reported in their CV (Magalhães et al., 2020). By clicking on this symbol, it is possible to see relevant information about that researcher, such as his name, the university or institute where he works and the address to access the Lattes curriculum. As can be seen in the graph, there is a greater concentration of these specialists in the southeastern region of the country, with emphasis on the state of São Paulo, which is the Brazilian state with the highest absolute number of public universities, present in at least 25 cities (FIA, 2019; Jornal da USP, 2019).

Figure 5: Location map of essential skills in injectable oncology in Brazil



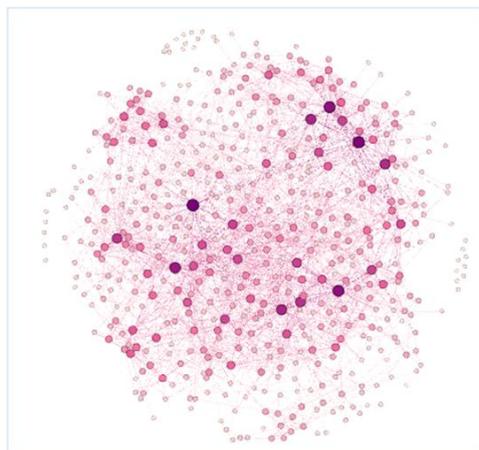
Source: extracted by the authors using the ScriptLattes tool, 2022.

In addition to the analyzes presented above, the Scriptlattes tool enables the graphical visualization of social networks through other computational tools such as Treecloud[®], Gephi[®], Cowo[®] and VOSviewer[®], already coupled to Scriptlattes (Quoniam, Ferraz, & Alvares, 2014). The academic collaboration between the

selected researchers was identified based on the publications carried out in co-authorship between them (Figures 6 and 7).

Figure 6 shows the collaboration (co-authorship in articles) network, in which the vertices or nodes (535) represent the researchers and the edges or links (2,213) represent the co-authorship between the vertices. Both the dimensions of the vertices and their color are associated with the number of connections for each researcher. The larger the size of the vertex and the darker it is, the more connected a researcher will be in the formed network. To extract the information from Figure 6, the “degree centrality” metric was used, which corresponds to the number of incident edges or the number of vertices adjacent to it (Giordano et al., 2015). The most connected vertex in the network has a degree centrality equal to 37, that is, it collaborates directly with 37 researchers identified in the context of this research. The researcher in question is Dr. Fernando de Queiroz Cunha, biologist and professor at the *Faculdade de Medicina de Ribeirão Preto (Universidade de São Paulo - USP)*, CNPq Research Productivity Scholar (Level 1A) and bibliographic production equal to 893 items. The second best-connected vertex has a degree centrality equal to 36 and corresponds to Dr. Geovanni Dantas Cassali, veterinarian and full professor at *Universidade Federal de Minas Gerais (UFMG)*, CNPq Research Productivity Scholar (Level 1A) and bibliographic production equal to 688 items.

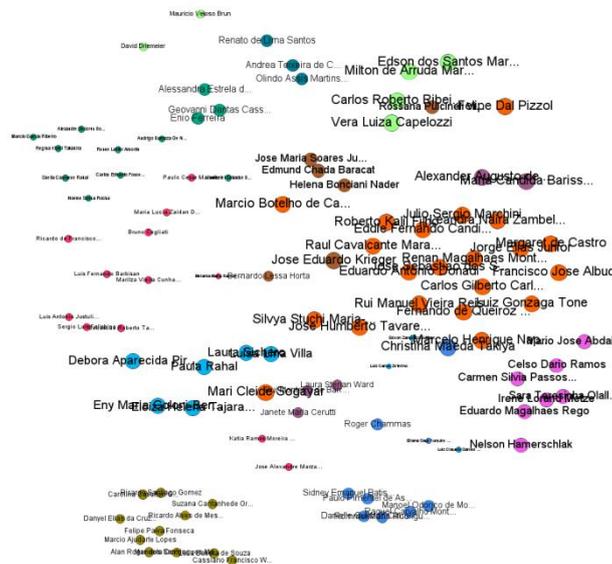
Figure 6: Co-authorship network among 298 researchers related to injectable oncology in Brazil



Source: prepared by the authors with Gephi, 2022.

For the analysis of Figure 7, the Modularity algorithm was used to determine communities or clusters based on their topology or network configuration. To facilitate the visualization of the communities formed in the collaboration network, it was decided to analyze those in which the researchers had a centrality of degree greater than or equal to 13, resulting in 101 researchers present in 11 communities.

Figure 7: Co-authorship network among the 101 most collaborative researchers with centrality of degree greater than or equal to 13



Source: elaborated by the authors with Gephi, 2022.

Thus, it can be inferred that 101 researchers collaborate with at least 13 others (Figure 7). Despite having other communities with greater modularity, they were not represented in the graph due to the chosen parameters. The largest community formed in the network is composed of 28 people, represented in orange, with a modularity equal to 1. The second largest group is composed of 21 people, represented in green and modularity equal to 11. The smallest group is composed of by only 7 people, represented in light green color and modularity equal to 3. Researchers with larger vertex sizes in communities are research leaders in their areas and develop a high number of collaborative research.

4. Conclusions

It is no longer news that humanity is living in the information age and having as a challenge the management of the knowledge generated every day. The exponential volume of data generated both in the academy and in organizations, provide conditions never seen and dealt with in all history. Therefore, identifying, extracting and treating confidential or otherwise data has been the great challenge for companies, academies and decision makers for the formulation of strategies in their organizations.

In this sense, this work contributed to a reflection of this context for public health. Through the analysis carried out, it was possible to obtain essential and reliable information from specialists in injectable oncology in Brazil. The information made available in the form of graphs facilitates and enables those interested in the subject to have a panoramic view of how research into injectable oncology has been carried out in the country over the years, as well as making it possible to obtain indicators on some aspects of science in Brazil. Micro (individual researcher and groups of researchers) and macro (collaboration networks or representativeness of the area) analysis were carried out, showing what is been developed and published in Brazil in the areas of R, D&I.

With regard to the spatial location of specialists in injectable oncology, it is possible to state that they are mainly concentrated in the southeast region (State of São Paulo). The contrast with the other states is clearly noted. This information obtained suggests that this disparity in the location of specialists is linked to the number of larger universities in the Southeast region, as well as the economic strength of the country in the region.

Nevertheless, it is worth highlighting the contribution of the analysis of co-authorship networks. A better understanding of these academic collaboration networks can help in the process of formulating S, T&I (Science, Technology and Innovation) policies, since they favor the aid of policies for the regional deconcentration of research centers in the area and/or the concession of research grants to the most economically disadvantaged regions of Brazil.

As observed in other studies carried out with ScriptLattes, the presence of a specialist in the analyzed topic is essential, since he will have the expertise both to formulate the curriculum search strategy and to perform a critical analysis of the results provided by the crawler.

Regarding the problems faced by ScriptLattes in data extraction is due to the indexing of information by researchers in the Lattes Platform, which makes it difficult for the program to retrieve information. Although the tool allows the extraction of a large volume of data, this process ends up being imprecise, since it requires human intervention to update the intellectual production data of the specialists in the Lattes Platform. This fact often means that the information is out of date.

Another limitation from the Lattes database is that all the information obtained can be considered 100% as "core competencies" of the researchers of the present study. This fact occurs, since the "Boolean expression" used for the identification and extraction of data, they appear in the curriculum in a general way (loose and spaced throughout the curriculum), but not necessarily in the same sentence of the work and/or keyword. New versions of the ScriptLattes software should seek a solution to this limitation.

Due to this research limitation, a proposal to solve the problem would be the normalization of indexing in the Lattes Platform, that is; that there was a standardization for the indexing of words by specialists, as well as for institutions to carry out their registrations in the Lattes database.

Therefore, it is recommended that administrators of the Lattes Platform, namely the National Council for Scientific and Technological Development (CNPq), of the Ministry of Science, Technology and Innovation, seek solutions for the standardization of indexing in the Lattes database. This action will increase strategic reliability for planning, management and policy formulation activities in the area of science, technology and innovation.

Thus, it can be inferred that the results obtained with ScriptLattes can help in the information management in the field of public health, particularly injectable oncology, area focused on this study, facilitating and optimizing the use of

information in daily actions and serving as a support tool in decision-making processes.

With regard to Health Management, the information generated using ScriptLattes will also enable funding agencies in the country to monitor the productivity of research in progress, as well as predict the future results of this study. Furthermore, based on the methodology used in this work, similar studies should be carried out, contributing to the theoretical framework of scientific and technological knowledge in the country.

Acknowledgments

The authors would like to thank the Instituto de Tecnologia em Fármacos Farmanguinhos for the infrastructure and institutional support from FIOCRUZ / Ministry of Health.

References

- Agbehadji, I. E., Awuzie, B. O., Ngowi, A. B., & Millham, R. C. (2020). Review of Big Data Analytics, Artificial Intelligence and Nature-Inspired Computing Models towards Accurate Detection of COVID-19 Pandemic Cases and Contact Tracing. *International Journal of Environmental Research and Public Health*, 17(15), Artigo 15. <https://doi.org/10.3390/ijerph17155330>
- Alves, W. A. L., Santos, S. D., & Schimit, P. H. T. (2016). Hierarchical Clustering Based on Reports Generated by Scriptlattes. Em I. Nääs, O. Vendrametto, J. Mendes Reis, R. F. Gonçalves, M. T. Silva, G. Von Cieminski, & D. Kiritsis (Orgs.), *Advances in Production Management Systems. Initiatives for a Sustainable World* (Vol. 488, p. 28–35). Springer International Publishing. https://doi.org/10.1007/978-3-319-51133-7_4
- ASHP. (2014). ASHP Guidelines on Compounding Sterile Preparations. *American Journal of Health-System Pharmacy*, 71(2), Artigo 2. <https://doi.org/10.2146/sp140001>
- Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An Open Source Software for Exploring and Manipulating Networks. *Proceedings of the International AAAI Conference on Web and Social Media*, 3(1), Artigo 1.
- Bordin, A. S., Gonçalves, A. L., & Todesco, J. L. (2014). Análise da colaboração científica departamental através de redes de coautoria. *Perspectivas em Ciência da Informação*, 19, 37–52. <https://doi.org/10.1590/1981-5344/1796>
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 68(6), 394–424. <https://doi.org/10.3322/caac.21492>

- Brito, A. G. C. D., Quoniam, L., & Mena-Chalco, J. P. (2016). Exploração da Plataforma Lattes por assunto: Proposta de metodologia. *Transinformação*, 28(1), 77–86. <https://doi.org/10.1590/2318-08892016002800006>
- Contador, J. C. (2010). Gestão de Operações. *A Engenharia de Produção a Serviço da Modernização da Empresa* (3a). <https://www.blucher.com.br>
- Esakandari, H., Nabi-Afjadi, M., Fakkari-Afjadi, J., Farahmandian, N., Miresmaeili, S.-M., & Bahreini, E. (2020). A comprehensive review of COVID-19 characteristics. *Biological Procedures Online*, 22(1), 19. <https://doi.org/10.1186/s12575-020-00128-2>
- Fernandes, F. R., & Silva, H. D. F. N. (2018). Análise da produção científica dos programas de pós-graduação e seu alinhamento com as diretrizes do sistema nacional de ciência, tecnologia e inovação: Um estudo cientométrico. *AtoZ: novas práticas em informação e conhecimento*, 7(2), 22. <https://doi.org/10.5380/atoz.v7i2.67241>
- Ferraz, R. R. N., Barnabé, A. S., Quoniam, L., Santos, A. M. dos, & Mariosa, D. F. (2018). Aspectos históricos da criação dos grupos de pesquisa em dengue no Brasil com a utilização da ferramenta computacional ScriptGP. *Ciência & Saúde Coletiva*, 23(3), 837–848. <https://doi.org/10.1590/1413-81232018233.00862016>
- Ferraz, R. R. N., Brito, A. G. C. de, Quoniam, L., & Mena-Chalco, J. P. (2019). Extraction and on line availability of indicators related to social responsibility, popularization and result prospection of research in nanotechnology in Brazil. *International Journal of Innovation*, 7(3), Artigo 3. <https://doi.org/10.5585/iji.v7i3.337>
- Ferraz, R. R. N., Quoniam, L., & Alvares, L. M. A. D. R. (2014). Avaliação de redes multidisciplinares com a ferramenta scriptlattes: Os casos da nanotecnologia, da dengue e de um programa de pós-graduação Stricto Sensu em Administração. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação*, 19(40), 67–98. <https://doi.org/10.5007/1518-2924.2014v19n40p67>
- Ferreira, G. I. S. (2009). COMISSÃO DE GRADUAÇÃO DO CURSO DE BIBLIOTECONOMIA.
- FIA. (2019, maio 30). Universidades Públicas: O que são, Importância e Lista de Instituições. FIA. <https://fia.com.br/blog/universidades-publicas/>
- Giordano, D. M., Bruning, E., & Bordin, A. S. (2015). Uso do scriptLattes e Gephi na Análise da Colaboração Científica. *Anais do Computer on the Beach*, 239–248. <https://doi.org/10.14210/cotb.v0n0.p239>
- Hayashi, C. R. M., Innocentini, M. C. P., Marcelo, J. F., & Bello, S. F. (2012). Análise de redes de colaboração científica. 35, 13.
- INCA. (2019). Estimativa 2020—Apresentação | INCA - Instituto Nacional de Câncer. <https://www.inca.gov.br/estimativa>
- Iriart, J. A. B., & Trad, L. A. B. (2020). Perfil da produção bibliográfica em Ciências Sociais e Humanas em Saúde e a percepção de pesquisadores: Avanços, limites e desafios. *Cadernos de Saúde Pública*, 36, e00085019. <https://doi.org/10.1590/0102-311X00085019>

- Jornal da USP. (2019, maio 16). Universidades públicas paulistas e o Estado de São Paulo. *Jornal da USP*. <https://jornal.usp.br/artigos/universidades-publicas-paulistas-e-o-estado-de-sao-paulo/>
- Klepa, T. C., & Pedroso, B. (2019). Analysis of the technical-scientific production of the National Council for Scientific and Technological Development (CNPq) productivity fellows in Pediatrics. *Einstein (Sao Paulo)*, 18. https://doi.org/10.31744/einstein_journal/2020AO5043
- Magalhães, J., Hir, M., Quoniam, L., Hartz, Z., & Oliveira, D. A. de. (2020). A MANAGEMENT TOOL TO AID IN THE TROPICAL OUTBREAK OF THE 21st CENTURY: SENIOR SCIENTISTS AND THEIR KNOWLEDGE OF THE TRIPLE THREAT DENGUE, ZIKA AND CHIKUNGUNYA. *Problems of Management in the 21st Century*, 15(1), 40–55. <https://doi.org/10.33225/10.33225/pmc/20.15.40>
- Matz, E. L., & Hsieh, M. H. (2017). Review of Advances in Uroprotective Agents for Cyclophosphamide- and Ifosfamide-induced Hemorrhagic Cystitis. *Urology*, 100, 16–19. <https://doi.org/10.1016/j.urology.2016.07.030>
- Mazieiro, G. (2019). Em 4 anos, Brasil reduz investimento em educação em 56%; cortes continuam—02/05/2019—UOL Educação. <https://educacao.uol.com.br/noticias/2019/05/02/em-4-anos-brasil-reduz-investimento-em-educacao-em-56.htm>
- Mena-Chalco, J. P., Digiampietri, L. A., Lopes, F. M., & Cesar, R. M. (2014). Brazilian bibliometric coauthorship networks. *Journal of the Association for Information Science and Technology*, 65(7), 1424–1445. <https://doi.org/10.1002/asi.23010>
- Mena-Chalco, J. P., & Junior, R. M. C. (2009). scriptLattes: An open-source knowledge extraction system from the Lattes platform. *Journal of the Brazilian Computer Society*, 10.
- Mohiuddin, A. K. (2020). Extemporaneous compounding: Cautions, controversies and convenience. *IP International Journal of Comprehensive and Advanced Pharmacology*, 3(4), Artigo 4. <https://doi.org/10.18231/2456-9542.2018.0028>
- Nascimento, J. O. do. (2020). Análise da produção científica brasileira em Oncologia: Um estudo comparativo entre artigos na base Lattes e os registros hospitalares de câncer. <https://ridi.ibict.br/handle/123456789/1104>
- Nascimento, J. O. do, & Gouveia, F. C. (2021). Estudo Comparativo da Produção Científica Brasileira em Câncer na Base Lattes e Casos de Câncer. *Tendências da Pesquisa Brasileira em Ciência da Informação*, 14. <https://revistas.ancib.org/index.php/tpbci/article/view/593>
- Nicholson, S. (2006). The basis for bibliomining: Frameworks for bringing together usage-based data mining and bibliometrics through data warehousing in digital library services. *Information Processing & Management*, 42(3), 785–804. <https://doi.org/10.1016/j.ipm.2005.05.008>
- Nigro, C. A., Ferraz, R. R. N., Quoniam, L., Alves, W. A. L., & Messais, L. R. R. (2015). Prestação de contas anual e quadrienal à Capes por um programa de Pós-Graduação

stricto sensu em Engenharia de Produção: Utilização da ferramenta computacional Scriptlattes-Scriptscupira. PRISMA.COM, 29, Artigo 29.

- Oun, R., Moussa, Y. E., & Wheate, N. J. (2018). The side effects of platinum-based chemotherapy drugs: A review for chemists. *Dalton Transactions* (Cambridge, England: 2003), 47(19), 6645–6653. <https://doi.org/10.1039/c8dt00838h>
- Pedroso, B., Picinin, C. T., & Cantorani, J. R. H. (2016). Production of productivity scholarship researchers in Physical Education in the 2010-2012 triennium: “Top 30” journals with the highest number of publications. *Revista Brasileira de Cineantropometria & Desempenho Humano*, 18, 731–736. <https://doi.org/10.5007/1980-0037.2016v18n6p731>
- Pergolizzi Jr, J. V., Labhsetwar, S., & LeQuang, J. A. (2013). Compounding Pharmacies: Who is in Charge? *Pain Practice*, 13(3), Artigo 3. <https://doi.org/10.1111/papr.12033>
- Ramos, M. J., Rito, P. da N., & Vieira, V. V. (2021). Monitoramento ambiental na manipulação de medicamentos oncológicos injetáveis à luz das normativas vigentes. *Vigilância Sanitária em Debate*, 9(2), Artigo 2.
- REVISTA GALILEU. (2022). Entenda como o corte de R\$ 2,4 bi do MEC afeta as universidades federais—Revista Galileu | Sociedade. <https://revistagalileu.globo.com/Sociedade/noticia/2022/10/entenda-como-o-corte-de-r-24-bi-do-mec-afeta-universidades-federais.html>
- Sag, A. A., Selcukbiricik, F., & Mandel, N. M. (2016). Evidence-based medical oncology and interventional radiology paradigms for liver-dominant colorectal cancer metastases. *World Journal of Gastroenterology*, 22(11), Artigo 11. <https://doi.org/10.3748/wjg.v22.i11.3127>
- Saldaña, P. (2022, outubro 5). Governo Bolsonaro corta R\$ 2,4 bi do MEC, e universidades veem situação insustentável. *Folha de S.Paulo*. <https://www1.folha.uol.com.br/educacao/2022/10/governo-bolsonaro-corta-r-24-bi-do-mec-e-universidades-veem-situacao-insustentavel.shtml>
- Santos, B. R. P. dos, & Damian, I. P. M. (2019). Análise da Gestão da Informação na Atenção Básica em Saúde: Um estudo em Unidades de Saúde da Família. *Informação & Sociedade: Estudos*, 29(3). <https://periodicos.ufpb.br/index.php/ies/article/view/45547>
- SENADO NOTÍCIAS. (2020). Senadores criticam corte de R\$ 4,2 bi do orçamento da Educação para 2021. Senado Federal. <https://www12.senado.leg.br/noticias/materias/2020/08/12/senadores-criticam-corte-de-r-4-2-bi-do-orcamento-da-educacao-para-2021>
- Silva, C. E. M. V. D., & Viegas, E. M. M. (2017). An analysis of the Brazilian researchers on freshwater fish genetics and reproduction as coauthors in academic papers. *Animal Reproduction*, 14(Suppl. 1), 1245–1252. <https://doi.org/10.21451/1984-3143-AR0055>
- Sobral, N. V., Duarte, Z., Santos, R. N. M. dos, & Mello, R. C. (2020). Redes de colaboração científica na produção de conhecimento em doenças tropicais negligenciadas no Brasil: Estudo a partir da plataforma LATTES do CNPq. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação*, 25, 01–22. <https://doi.org/10.5007/1518-2924.2020.e72981>

Valentim, M. L. P., Alcará, A. R., Fedel, B., Cervantes, B. M. N., Moraes, C. R. B. de, Barbalho, C. R. S., Silva, E. O., Carvalho, E. L. de, Serzedello, N. T. B., Zwaretch, N. S., Woida, L. M., Molina, L. G., Sala, F., Gelinski, J. V. V., North, K., Tomaél, M. I., Chiara, I. G. D., Lenzi, L. A. F., Catarino, M. E., ... Santos, J. C. dos. (2006). Informação, conhecimento e inteligência organizacional. Em *Oficina Universitária. Oficina Universitária*.
https://ebooks.marilia.unesp.br/index.php/lab_editorial/catalog/book/321

Vincenzo, N. (2008). Modularity for community detection: History, perspectives and open issues. 67.

Webster, W. D., Parks, G. T., Titov, D., & Beasley, P. (2014). The production of radionuclides for nuclear medicine from a compact, low-energy accelerator system. *Nuclear Medicine and Biology*, 41 Suppl, e7-15. <https://doi.org/10.1016/j.nucmedbio.2013.11.007>

Editor's note

The editor responsible for the publication of this article is Mario Barité.

Author Contribution Note

La siguiente tabla describe los roles y contribuciones de cada autor/autor del manuscrito. Para ello, se adoptó la Contributor Roles Taxonomy* (CRediT), en la que se marcó con una X el rol/contribución del respectivo autor(es):

Roles y contribuciones	Henrique Koch Chaves	Francisco Theodoro Arantes Florencio	Carla Cristina de Freitas da Silveira	Adelaide Maria de Souza Antunes	Jesús Pascual Mena-Chalco	Jorge Lima de Magalhães
Concepción del manuscrito			X	X		X
Escritura manuscrita	X					X
Metodología	X				X	X
Curación de datos	X	X		X	X	
Discusión de resultados	X		X	X		X
Análisis de los datos	X		X		X	X

* <https://credit.niso.org/>