

Scientific production of women in Brazil

Produção científica das mulheres no Brasil

Monique de Oliveira SANTIAGO¹  0000-0003-0163-2774

Felipe AFFONSO¹  0000-0003-2986-2379

Thiago Magela Rodrigues DIAS¹  0000-0001-5057-9936

Abstract

Researchers concentrate their efforts to understand the different female relations with science, using approaches that review their scientific and technological participation, as well as, seeking to understand their academic trajectory and performance. In this context, this study aimed to analyze the participation of women using as database the set of PhD graduates who have their curricula entered in the Lattes Platform. The data were collected and selected obtaining a set of 125,515 curricula of women who had completed their PhD. The PhD data were grouped according to the large areas of expertise (fields of science) and academic training, in which it was possible to analyze the academic evolution and the scientific and technological production of the group in a temporal manner. The different types of studies that help to understand the general aspect of women active mainly in science, besides being relevant, exhibit the characteristics of their research. This may be useful for the generation of national scientific indicators, for the management of information in the scientific area and for technological development. It is also useful to encourage and value participation of women in science.

Keywords: Bibliometric. Female gender. Large areas of expertise. Lattes platform. Woman in science.

Resumo

Pesquisadores concentram seus esforços em entender as mulheres e suas variadas relações com a ciência, trazendo abordagens que analisam a participação científica e tecnológica delas e que buscam entender suas trajetórias e desempenho acadêmico. Nesse contexto, este estudo teve como objetivo analisar a participação de mulheres na ciência utilizando como base de dados o conjunto de doutoras que possuem seus currículos cadastrados na Plataforma Lattes. Os dados foram coletados e selecionados, obtendo-se um conjunto de 125.515 currículos com o gênero feminino e a formação acadêmica doutorado concluída. Os dados das doutoras foram agrupados de acordo com as grandes áreas de atuação e formação acadêmica/titulação, através dos quais foi possível analisar a evolução acadêmica e a produção científica e tecnológica do conjunto de maneira temporal. Os vários tipos de estudos que buscam entender o aspecto geral da mulher e, principalmente, a atuação delas na ciência, além de relevantes, apresentam as características de suas pesquisas e podem ser úteis para a geração de indicadores científicos nacionais, para o gerenciamento de informações na área científica e desenvolvimento tecnológico, além de incentivar e valorizar a participação das mulheres na ciência.

Palavras-chave: Bibliometria. Gênero feminino. Grandes áreas de atuação. Plataforma Lattes. Mulheres na ciência.

Introduction

Science is based on the search for knowledge that tries to understand the phenomena of nature and their interaction with the environment, resulting in scientific production (Lopes *et al.*, 2020). The generation of knowledge

¹ Centro Federal de Educação Tecnológica de Minas Gerais, Programa de Pós-Graduação em Modelagem Matemática e Computacional. Av. Amazonas, 7675, Nova Gameleira, 30510-000, Belo Horizonte, MG, Brasil. *Correspondência para/Correspondence to:* T.M.R. DIAS. E-mail: <thiagomagela@gmail.com>.

Received on May 19, 2020, final version resubmitted on August 26, 2020 and approved on September 30, 2020.

Como citar este artigo/How to cite this article

Santiago, M. O.; Affonso, F.; Dias, T. M. R. Scientific production of women in Brazil. *Transinformação*, v. 32, e200032, 2020. <https://doi.org/10.1590/2318-0889202032e200032>



and scientific production have increased significantly over the last few years, in which the Web has become one of the tools that effectively facilitates access and dissemination of information that has allowed the development of several works that seek to analyze and understand how science evolved and how scientific collaboration occurs. Among the several studies on scientific production, one that is prominent is constituted by gender studies. Gender can be defined as an element that defines social relationship based on the perceived differences between genders, in addition to giving meaning to power relations (Scott, 1995). Therefore, it corresponds to an interdisciplinary field, that has identity as a general theme, and the representation of men and women in society. This field includes the study of the sub-field Women's Studies, which covers, among several dimensions of interest, women and their different relations with science (Leta, 2014).

The studies that focus on women and their multiple relations with science are interdisciplinary and comprehensive themes, in addition to presenting different approaches and different characteristics, such as historical, social, economic, in which they seek to trace their trajectory and career (Moschkovich; Almeida, 2015; Bellotti; Guadalupi; Conaldi, 2016; Abramo; D'Angelo; Costa, 2018; Marini; Meschitti, 2018). Despite advances in the participation of women in different contexts of science and in different segments considering academic careers, there is still a gender gap in science in the world that can and needs to be better understood. It is clear that there is clear evidence of gender inequalities in science, which probably has deep cultural and institutional roots (Valentova *et al.*, 2017). Many of these differences between the performance of men and women in science, may be related to the greater presence of men in the field (Leta, 2014). Reflecting on new ways to investigate women's productivity in science is a step towards understanding the current scenario of women's scientific performance and contribution to research.

An inherent difficulty in carrying out analyses regarding a country's scientific production is generally associated with the collection of data, which are normally present in different types of repositories, making it difficult to retrieve and analyze them. However, this process can be facilitated using the Lattes Platform curriculum data repository, used in several studies to analyze scientific production (Santos; Oliveira, 2017; Gomes, 2018; Damaceno *et al.*, 2019). The Lattes Platform is considered an important set of Brazilian scientific data that provides high-quality information and allows searching for data in connection with the individuals who are registered there, such as their academic education and scientific production, among others (Lane, 2010). This data set is integrated into a single platform containing records from educational institutions, research groups in the country and curricula containing information on the academic history of a large portion of the scientific community in Brazil.

In general, the review of large repositories is not a trivial task, as it is influenced by the amount of data to be analyzed and mainly by the specific characteristics of each repository, which are unique and most of them do not have a defined structural pattern. In this framework, there is the bibliometric that assists in the quantification of written communication processes, production statistical analysis methods and dissemination of knowledge applied to scientific data sources (Araújo, 2006). So, we can say that bibliometric analyses consistently provide relevant indicators for national planning and for supporting the evolution of scientific research.

In this context, this work aims to analyze how the academic, scientific and technological participation of women has evolved, reviewing how their research has been carried out and also, how their academic trajectory has been built over the years, based on applied bibliometric analysis metrics, and curriculum data available on the Lattes Platform. This study, in addition to presenting an overview of the activities and research carried out by women, provides a general scenario that can contribute to the generation of national scientific indicators and the management of scientific and technological information.

The great motivation for this work is the possibility of understanding the scientific participation of women and to understand the current scenario in which they are inserted. Despite recognizing that science is a neutral field, the values and characteristics attributed to women are not valued and inequalities permeate the scientific

field (Silva; Ribeiro, 2014). Thus, when there is consent in breaking the barriers of inequality, we can ensure research excellence by giving equal opportunities to the best scientists and innovators (Valentova *et al.*, 2017).

Methodological Procedures

The data recorded in each of the curricula that make up the Lattes Platform are entered by the individual. The data are available on the internet, and these records contain the entire history of the researcher's professional, scientific and academic life. Taking into account the wealth of information recorded in the curricula, the selection of curricula from the Lattes Platform repository as the main source of data to measure and evaluate the scientific performance of women in Brazil is justified. So, to analyze this broad set of data, it was initially necessary to acquire the data and select the curricula by the criterion of academic background and then select by the gender criterion. Subsequently, the data were prepared to generate the specific files and, finally, to obtain visualizations of these data (Figure 1).

Thus, the first step, which corresponds to the process of data acquisition and selection of curricula by the criterion of academic formation, was carried out using the LattesDataXplorer framework developed by Dias (2016). This extractor has a set of techniques and methods responsible for the collection, selection, processing, and analysis of data from the Lattes Platform curricula. The collection of all curricula, using the above framework, was

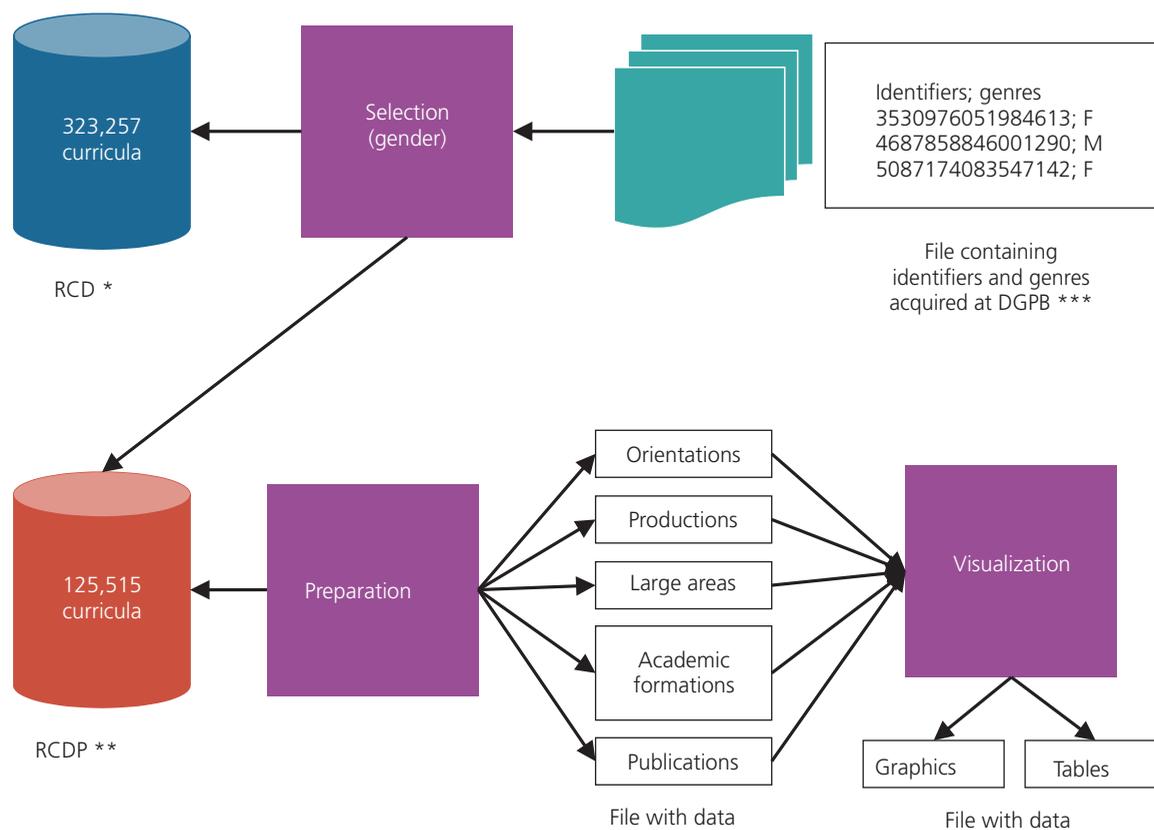


Figure 1. Gender selection process, data preparation and visualization.

Note: *Curricula repository of all PhD, **Curricula repository of woman PhD, ***Directory of Research Groups in Brazil.

Source: Prepared by the authors (2020).

performed in October 2019, and covered an excess of 6,300,000 records. Resumés are stored locally and have the XML format (extensible Markup Language – a markup language that contains well-defined sections and fields), making this file format more suitable for automatic data processing. In order to carry out a detailed analysis of the scientific participation of women, it was decided to limit the data through the level of academic education, reducing the number of individuals who completed the PhD (Doctor of Philosophy) level. Although PhD is not the most significant among the levels of education, as stated by Dias (2016), it is responsible for 74.51% of the articles published in journals and 64.67% published conference papers. So, after the acquisition of all the curricula registered in the Lattes Platform, the LattesDataXplorer selection module was used to select, among the curricula those of Scholars who completed the PhD, totaling a set of 323,257 curricula stored in the Curricula repository of all PhD (RCD, Curricula Repository of all PhD).

When we approach scientific production data and metrics studies, recovering the authors' in the records, indicating the gender field is a big challenge (Leta, 2014), since the Lattes Platform does not make this field available for public consultation. Therefore, in order to carry out the second stage, which corresponds to the selection by gender, it was necessary to search for this information in another database. So, the Directory of Research Groups in Brazil (DGPB) was used, which has an inventory of scientific and technological research groups active in the country, making the Census data available in an open format. The files were processed, and the identifier and gender data were obtained for all researchers, students, and technicians in the Census archives. Thus, a list of female identifiers was retrieved from the archives and it was possible to search the RCD for resumés with identifiers equal to those on the list, defining the main set of XML files with a total of 125,515 women's curricula, called the Curricula repository of female PhD (RCDP).

For the third stage, which corresponds to the preparation of the data, the structure of the RCDP files was initially identified. An XML file is organized using elements and attributes, in which it starts with the root (parent) element and runs through the child elements until reaching the attributes. The structure of the XML curricula of the Lattes Platform has the root element, called "Curriculo-vitae", and several child elements that have their elements and attributes. The XML files extracted from the Lattes Platform present a wide range of data, such as level of academic education, large areas of activity, research and extension projects, languages, bibliographic and technical productions such as conference papers and journal article, presentation work, participation in events, guidelines, among others. Thus, after accessing each XML curriculum and obtaining the specific information of that curriculum, it was stored in a collection of structured files. As each curriculum has a specific amount of information, these data were grouped as to guidelines, productions, large areas, academic education, publications. After grouping the data, the last stage was performed, which corresponds to the visualization of the data, that is, the characterization to facilitate the analysis of the data, generating graphs and tables.

As for the limitation related to data collection, the reliability of the data should be considered, but two factors made it impossible to use the entire curriculum set of the PhD for analysis. The first corresponds to the Lattes Platform, which does not provide information on gender and the second corresponds to the DGPB, as not all PhD are included in a research group. It is worth mentioning that the sample used in this study corresponds to all the PhD who are inserted or participated at some point in a research group made available in the DGPB Census Archives.

Results

The data collected from the Lattes Platform, using the LattesDataXplorer framework in October 2019, totaled more than 6,300,000 records. Of this total, records in which PhD was the level of academic education were selected, totaling 323,257 (5.13%) curricula of all PhD from different areas of scientific knowledge. These same data were

selected according to the gender criterion (using the DGPB as a validator), totaling 125,515 (1.99%) resumes of female PhD (RCDP), and these data will be used for the analyses that portray women's scientific participation.

The field of science can be defined as a clustering of different areas of knowledge that have affinity between their objects, cognitive methods, and instrumental resources (Basil, 2019). So, information related to fields of science can be recorded in the Lattes Platform curricula in different elements, allowing to choose and provide information on nine areas and their respective subareas of knowledge, which follows the classification of CNPq (National Council for Scientific and Technological Development) areas and subareas (Conselho..., 2019).

All the fields of science entered by each woman with a PhD were recovered and the one with the highest occurrence was selected, and was considered the most relevant. Thus, grouping the curricula of women with PhD by fields of science showed that there are more representatives in humanities and less representatives in engineering. They are: Humanities 24,845 (19.79%), Life Sciences 23,133 (18.43%), Biological Sciences 21,286 (16.96%), Agrarian Sciences 11,777 (9.38%), Exact and Earth Sciences 11,344 (9.04%), Applied Social Sciences 10,849 (8.64%), Linguistics, Letters and Arts 8,943 (7.13%), Engineering 5,398 (4.30%), Others 517 (0.41%). These data corroborate the study by Olinto (2011), that demonstrated that men predominate in the exact sciences careers and women predominate in the biological and health areas. An analysis that shows the underrepresentation of women at the highest levels of Brazilian science for all large fields of science is reported in the study by Valentova *et al.* (2017).

A considerable number of curricula do not provide information related to the large areas of expertise 7,423 (5.91%). As it is not a mandatory information, when inserting a new academic graduation, there is no concern with filling in the data in this field or even forgetfulness can occur on the part of the woman with PhD. The RCDP data show a significant growth in relation to the data made available by the Lattes Panel, which were extracted from the curriculum base on November 30, 2016, and made available for public consultation, in which they were organized in the form of panels with descriptive data about geography, gender and age group, among others. The information referring to the panel Distribution by gender, age group, and large areas of knowledge, corresponds to 63,853 (47.50%) curricula of women with PhD divided into their fields of science (Painel Lattes, 2016). So, when comparing these data provided by the Lattes Panel with the data from the RCDP collected from the Lattes Platform, we can see that in a period of three years, most of the large knowledge areas did not have a change in positioning; only the field of science Agrarian Sciences showed a small growth. and surpassed the large areas Exact and Earth Sciences and Applied Social Sciences.

As Dias; Moita (2018) sustains, updated curricula provide a current perspective on Brazilian scientific production, as they already probably have publications entered recently. So, as 86,665 RCDP female PhD have had their curricula updated in 2019, relevant information corresponds to the percentage of updated curricula per year for each field of science (Table 1).

Table 1. Updated curricula in connection with large areas of knowledge.

Large Areas	2000 to 2004	2005 to 2009	2010 to 2014	2015	2016	2017	2018	2019
	%							
Agrarian Sciences	0.24	1.38	4.67	2.04	3.12	5.36	13.58	69.62
Applied Social Sciences	0.10	1.07	4.54	1.68	2.65	4.39	11.88	73.69
Biological Sciences	0.25	1.66	5.79	2.45	3.62	5.65	13.95	66.63
Engineering	0.13	1.72	5.46	2.50	2.93	4.89	12.25	70.12
Exact and Earth Sciences	0.27	1.71	5.65	2.28	3.24	5.65	13.05	68.14
Humanities	0.19	1.49	5.58	2.25	3.17	4.88	12.39	70.05
Life Sciences	0.14	1.20	5.16	2.14	2.89	5.30	13.68	69.49
Linguistics, Letters, Arts	0.13	1.71	6.08	2.36	2.95	4.93	12.29	69.54
Others	0.00	0.39	11.41	4.45	2.51	4.84	11.80	64.60
Uninformed	0.58	2.53	4.82	1.91	3.44	6.91	15.94	63.87

Source: Prepared by the authors (2020).

Looking at Table 1, the rate of curricula updated each year is very close among the large knowledge areas. However, if we analyze only the year 2019, the large Engineering area has the second-highest percentage of updating, although it is a large area with a low number of doctorates, 4.30%. The same applies to the field of science of Applied Social Sciences, which has the highest rate of updates, even though it is a large area with relatively few female PhD (8.63%).

The data related to scientific production have been very expressive nowadays, as in the study conducted by Elsevier (2017), which used data from the last 20 years of scientific production in different regions of the world and showed the difference in female participation. For Brazil, in the period between 2011 and 2015, just under half (49%) of the scientific studies were produced by women, in contrast to the previous period between 1996 and 2000, in which women contributed with only 38% of Brazilian scientific production.

Thus, using the most representative means of scientific production, journal articles and conference papers, the total for each field of science was summarized. For journal articles in RCDP, the totals by field of science yielded: Life Sciences 615,356; Biological Sciences 436,399; Humanities 297,098; Agrarian Sciences 252,166; Exact and Earth Sciences 176,833; Applied Social Sciences 137,093; Linguistics, Letters and Arts 94,684; Engineering 76,277; Others 5,366. The total for conference papers in RCDP was: Life Sciences 1,160,640; Biological Sciences 903,642; Humanities 753,165; Agrarian Sciences 681,955; Exact and Earth Sciences 432,215; Applied Social Sciences 282,920; Engineering 242,964; Linguistics, Letters and Arts 193,627; Others 14,135. Scientific production data from journal articles and conference papers, follow the same flow of general data, indicating that the largest contributors are in humanities, namely: Life Sciences, Biological Sciences, Humanities and Agrarian Sciences.

By dividing all articles published in the fields of science by the total number of individuals, we obtain the production *per capita*. Thus, the total in RCDP journal articles *per capita* is as follows: Life Sciences 26,6; Agrarian Sciences 21,41; Biological Sciences 20,5; Exact and Earth Sciences 15,59; Engineering 14,13; Applied Social Sciences 12,64; Humanities 11,96; Linguistics, Letters and Arts 10,59; Others 10,38. The total number of conference papers *per capita* by RCDP PhD is as follows: Agrarian Sciences 57,91; Life Sciences 50,17; Engineering 45,01; Biological Sciences 42,45; Exact and Earth Sciences 38,1; Humanities 30,31; Applied Social Sciences 26,08; Others 27,34; Linguistics, Letters and Arts 21,65.

A review of these data indicates a significant change in position in some areas of activity in the production of journal articles and conference papers. For the conference papers, the first field of science that stands out is Agrarian Sciences, which appears as the fourth-largest producer of articles and ranks first in the *per capita* production. The Engineering field of science revealed a surprising and unexpected change from the penultimate to the fourth place in the total number of articles *per capita*, which shows that, although it is an area with little participation of female PhD, it is an area that produces a lot of content. As for journal articles, the production of content for Agrarian Sciences, Exact and Earth Sciences and Engineering stands out considering the production *per capita* of female PhDs, demonstrating that they are major producers of content even with fewer PhD in each of these areas. A study that addresses the publications and all areas of the Exact and Earth Sciences was published by Tuesta *et al.* (2019). A different approach on the fields of science and the modeling of social networks of collaborations between scientists can be observed in the studies of Bellotti; Guadalupi; Conaldi (2016) and Bellotti; Kronegger; Guadalupi (2016).

These data reveal the importance of scientific production of female PhDs in several fields of science. It was possible to identify the current scientific production scenario and verify that even the large fields of science that have less women participation, such as engineering, the production *per capita* presents the opposite scenario. The Lattes Platform can be considered the researcher's memory freely available on the Web. Throughout the researcher's academic and professional trajectory, it provides information on all the researcher's bibliographic, technical, and other types of publications, in addition to contemplating academic education data, areas of expertise, professional performance, and various complementary data. According to Brito; Quoniam; Mena-Chalco (2016), the Lattes

Platform is an important curricular information system, which can be used mainly for isolated assessments of groups or institutions.

All levels of academic education can be informed in the Lattes Platform curricula, and it is possible to fill in various fields such as institution, course name, initial and final year, title of work, name of academic advisor and co-advisor, different fields of science, among others. For PhD, master's degree and professional master's degrees, you can select the major areas of knowledge and choose from nine options, selecting the areas and then the sub-areas of the postgraduate course, following the CNPq classification.

To quantify master's degree and PhD, all the large areas where female PhD were active were recovered. Since it is possible to fill more than one area of knowledge, only the first large area filled out by female PhD was selected, considering it to be the most relevant among the areas inserted. Thus, by checking the academic education level of the PhD, it was possible to classify the academic level by large areas of knowledge (Table 2).

As some fields of science in the curricula of the Lattes Platform are not mandatory, the data on the initial year, the final year, or large area of knowledge that did not contain information was characterized as "Not informed" and these records were not used for the review. Thus, the master's degree 35,072 (27.94%) records and the PhD, 36,686 (29.22%) records were categorized as "Not informed". The information presented in Table 2 regarding the mean, median, and standard deviation, are relevant statistical calculations, as they allow verifying the trend and variability of the data for each of the fields of science. So, we can define these terms as follows: mean is the value that demonstrates the concentration of data distribution, that is, it can be considered a point of equilibrium of frequencies in a histogram; the median can be considered as a central trend that indicates exactly the central

Table 2. Academic education time by main fields of science.

Large areas	Master's degree				
	Total		Mean	Median	Standard deviation
	<i>n</i>	(%)			
Agrarian Sciences	9,432	(7.51)	3.17	3	0.77
Applied Social Sciences	8,039	(6.40)	3.50	3	1.25
Biological Sciences	15,43	(12.29)	3.31	3	0.91
Engineering	5,517	(4.40)	3.33	3	0.93
Exact and Earth Sciences	9,195	(7.33)	3.39	3	0.98
Humanities	19,53	(15.56)	3.61	3	1.32
Life Sciences	15,26	(12.16)	3.24	3	1.02
Linguistics, Letters and Arts	7,739	(6.17)	3.57	3	1.23
Others	299	(0.24)	3.13	3	0.73
Large areas	PhD				
	Total		Mean	Median	Standard deviation
	<i>n</i>	(%)			
Agrarian Sciences	7,918	(6.31)	4.78	5	0.91
Applied Social Sciences	7,372	(5.87)	5.10	5	1.28
Biological Sciences	15,36	(12.24)	5.13	5	0.98
Engineering	5,539	(4.41)	5.34	5	1.09
Exact and Earth Sciences	8,911	(7.10)	5.34	5	1.17
Humanities	19,86	(15.82)	5.15	5	1.21
Life Sciences	15,72	(12.52)	4.62	5	1.08
Linguistics, Letters and Arts	7,717	(6.15)	5.16	5	1.18
Others	442	(0.35)	5.04	5	0.80

Source: Prepared by the authors (2020).

value of certain data; the standard deviation demonstrates the regularity of a data set to point out the degree of oscillation of these data compared to the average of the values in the set.

Table 2 shows that the median referring to the time of academic education in all fields of science did not change for the two postgraduate levels. Checking the average data for the master's degree, the value exceeded the median in all major areas, with the lowest average training time being 3.13 years and the highest 3.61 years. As this value is influenced by values much higher than the average, it was verified that 110 records report information on the duration of academic training between 11 and 23 years. The largest standard deviation (1.32) corresponds to the large area Humanities, indicating how far the data in the set depart from the average.

The PhD, on the other hand, showed a greater variation between the fields of science for the average, with the two fields of Agrarian Sciences and Life Sciences showing a lower average than the median, respectively 4.78 and 4.62. For these two fields, the much higher values that influence the average totaled the number of 32 records that have the duration of academic training between 11 and 17 years. As for the rest of the fields of science, the total number of registrations was 226, with the duration of academic training between 11 and 25 years, which may explain the average being above the median. The largest standard deviation that indicates how far the data from the set depart from the average corresponds to the field of science Applied Social Sciences (1.28), and it does not have the highest average.

The PhD degree allows students to deepen their knowledge in a certain field; it is possible to follow the same area of knowledge or not. Thus, seeking to understand evolution across the major fields of science (the individuals were distributed in 4 time bands considering the year of completion of their training) distributed it for periods in each area of knowledge, both for the master's (Figure 2a) and for the doctorate (Figure 2b). Thus, the RCDP curricula have as their first year of training the master's corresponding to 1949 and for the doctorate 1948, divided between the periods: A: 1948 to 1965; B: 1966 to 1983; C: 1984 to 2001; D: 2002 to 2019.

Looking at Figure 2(a), Humanities is the field of science that ranked first regarding the conclusion of the master's degree in all periods. Regarding period A, in addition to this field of science, other fields that are relevant are Exact and Earth Sciences and Linguistics, Letters and Arts. In this first period, only five of the nine fields of science had female PhD with completed postgraduate studies. In period B, all major fields had female PhD with completed postgraduate studies, with Biological sciences being a large field that experienced considerable growth for this period. For periods C and D, in addition to the Humanities, the most representative in terms of education are Biological sciences and Life sciences that alternate positions between periods. As for the growth between periods, for period C the three fields of science with the highest rates correspond to Life Sciences (11.88%), Engineering (9.13%), and Agrarian Sciences (8.75%). For period D, the fields of science with the highest growth rates correspond to Others (2.84%), Agrarian Sciences (1.26%) and Biological Sciences (0.92%).

On the other hand, looking at Figure 2(b), the first period also has a graduation in only five fields of science, but the same fields of science of the master's degree do not correspond for this same period. Humanities were the fields of science with the most graduation for periods A, C and D. As for the growth between periods, for period C the three fields of science with the highest rates correspond to Engineering (33.14%), Life Sciences (22.06%) and Applied Social Sciences (13.38%). For period D, the fields of science with the highest growth rates correspond to Others (22.50%), Applied Social Sciences (5.58%) and Agrarian Sciences (5.50%).

These 71 years of the master's degree and PhD history, show the evolution of the fields of science, having as reference the year of graduation recorded in the curriculum of the Lattes Platform. The most recent period, which corresponds to 2002 to 2019, shows the fields of science Humanities, Biological Sciences and Life Sciences as the most representative; however, in the previous period it was possible to observe Exact and Earth Sciences as being the most relevant.

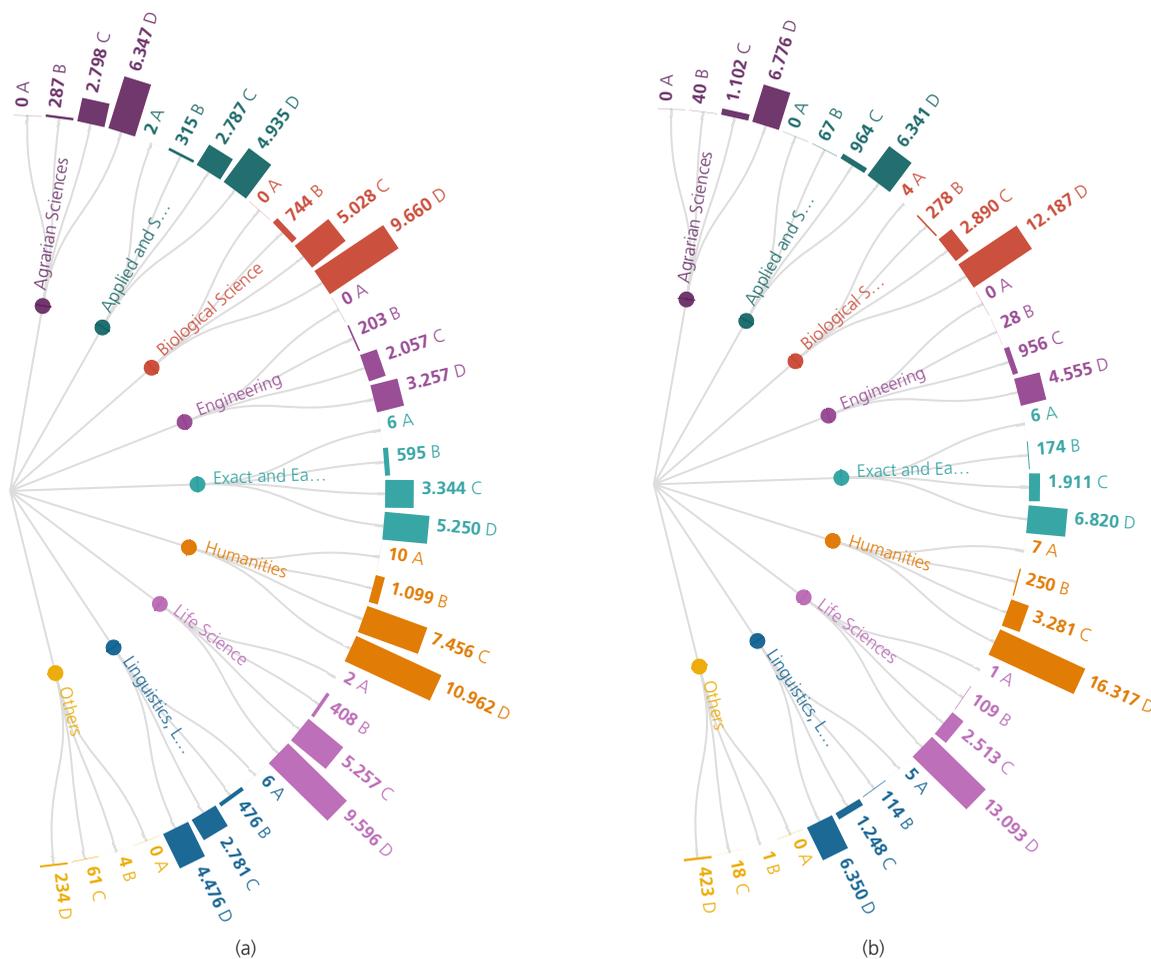


Figure 2. Education evolution: Master's degree and PhD.
 Source: Prepared by the authors (2020).

Conclusion

Women in science have been the focus of a number of studies and this approached the perspective of the fields of science and the academic graduation of female PhD. Using the Lattes Platform curricula as a source of data, it was possible to select, by the criteria of PhD graduation/degree and the female gender, the women's PhD who have already participated or are currently integrated into a research group. Thus, the preparation and analysis of the data were carried out using the RCDP, which includes 125,515 curricula of female PhDs.

The fields of science represent a relevant topic because they correspond to a clustering of women with PhD according to their affinities, methods for acquiring knowledge and resources that reflect the same socio-political context. So, by distributing PhD over the fields of science, the following topics were observed: total number of women with PhD, general total and *per capita* for scientific meetings papers and for journal articles, updating of curricula by year, the evolution of graduation with a master's degree and PhD.

The master's and PhD academic education process consists of a continuous search for knowledge to develop a set of competencies and skills that will be used in future research and teaching activities of the teacher-researcher. In this context, the evolution of the academic training of PhDs was presented, such as the time of academic training,

the evolution of academic training per year for a master's degree and PhD. It was possible to observe how some fields of science are more representative than others. The large fields of science of humanities, biological sciences and life sciences, stand out due to their representativeness.

Such large scientific areas stand out for having a considerable representation of women, working mainly in academic careers. On the other hand, the great area of Engineering stands out, being the least representative, which results in a small number of women being trained and, consequently, acting in the field of science that compose it.

The studies that have as a theme the general aspects of women and, particularly in science, are relevant and can contribute to the generation of national scientific indicators and the management of information in the scientific and technological areas. As future work, it is expected that an analysis of the bibliographic production of the researchers who are CNPq research productivity fellows and the scientific collaboration of the researching PhD be carried out.

Acknowledgements

The authors are grateful to *Centro Federal de Educação Tecnológica de Minas Gerais* and *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* for their assistance in the research.

Contributors

M. O. SANTIAGO and T. M. R. DIAS participated in all stages of writing the article. F. AFFONSO contributed to the collection, treatment and analysis of data.

References

- Abramo, G.; D'Angelo, C. A.; Costa, F. D. The effects of gender, age and academic rank on research diversification. *Scientometrics*, v. 114, n. 2, p. 373-387, 2018. Doi: <https://doi.org/10.1007/s11192-017-2529-1>.
- Araújo, C. A. Bibliometria: evolução histórica e questões atuais. *Em Questão*, v. 12, n. 1, p. 11-32, 2006. Disponível em: <https://www.seer.ufrgs.br/EmQuestao/article/view/16>. Acesso em: 18 ago. 2020.
- Bellotti, E.; Guadalupi, L.; Conaldi, G. Comparing fields of sciences: multilevel networks of research collaborations in Italian academia. In: Lazega, E.; Snijders, T.A.-B. (ed.). *Multilevel network analysis for the social sciences*. [S.l.]: Springer, 2016. p. 213-244. Doi: https://doi.org/10.1007/978-3-319-24520-1_9.
- Bellotti, E.; Kronegger, L.; Guadalupi, L. The evolution of research collaboration within and across disciplines in Italian academia. *Scientometrics*, v. 109, n. 2, p. 783-811, 2016. Doi: <https://doi.org/10.1007/s11192-016-2068-1>.
- Brasil. Ministério da Ciência, Tecnologia, Inovações e Comunicação. *Recursos Humanos: indicadores sobre o ensino de pós-graduação*. Brasília: Ministério da Ciência, Tecnologia, Inovações e Comunicação, 2019. Disponível em: http://antigo.mctic.gov.br/mctic/opencms/indicadores/detalhe/Recursos_Humanos/RH_3.5.4.html. Acesso em: 29 abr. 2020.
- Brito, A. G. C.; Quoniam, L.; Mena-Chalco, J. P. Exploração da Plataforma Lattes por assunto: proposta de metodologia. *Transinformação*, v. 28, n. 1, p. 77-86, 2016. Doi: <http://dx.doi.org/10.1590/2318-08892016002800006>.
- Conselho Nacional de Desenvolvimento Científico e Tecnológico. *Tabela de áreas do conhecimento*. Brasília: CNPq, 2019. Disponível em: <http://lattes.cnpq.br/documents/11871/24930/TabeladeAreasdoConhecimento.pdf/d192ff6b-3e0a-4074-a74d-c280521bd5f7>. Acesso em: 8 dez. 2019.
- Damaceno, R. J. *et al.* The brazilian academic genealogy: evidence of advisor: advisee relationships through quantitative analysis. *Scientometrics*, v. 119, n. 1, p. 1-31, 2019. Doi: <https://doi.org/10.1007/s11192-019-03023-0>.
- Dias, T. M. R. *Um estudo da produção científica brasileira a partir de dados da Plataforma Lattes*. 2016. 181 f. Tese (Doutorado em Modelagem Matemática e Computacional) – Centro Federal de Educação Tecnológica de Minas Gerais, Belo Horizonte, 2016.
- Dias, T. M. R.; Moita, G. F. Um retrato da produção científica brasileira baseado em dados da plataforma lattes. *Brazilian Journal of Information Science*, v. 12, n. 4, p. 62-74, 2018. Doi: <https://doi.org/10.36311/1981-1640.2018.v12n4.08.p62>.
- Elsevier. *Gender in the global research landscape*. [S.l.]: Elsevier, 2017. Available from: https://www.elsevier.com/__data/assets/pdf_file/0008/265661/ElsevierGenderReport_final_for-web.pdf. Access on: Mar. 1, 2020.
- Gomes, J. O. *Uma análise temporal dos principais tópicos de pesquisa da ciência brasileira a partir das palavras-chave*

de publicações científicas. 2018. 127 f. Tese (Doutorado em Modelagem Matemática e Computacional) – Centro Federal de Educação Tecnológica de Minas Gerais, Belo Horizonte, 2018.

Lane, J. Let's make science metrics more scientific. *Nature*, v. 464, n. 7288, p. 488, 2010. Doi: <https://doi.org/10.1038/464488a>.

Leta, J. Mulheres na ciência brasileira: desempenho inferior? *Revista Feminismos*, v. 2, n. 3, p. 139-152, 2014. Disponível em: <https://cienciasmedicasbiologicas.ufba.br/index.php/feminismos/article/viewFile/30039/17771>. Acesso em: 17 ago. 2020.

Lopes, E. F. B. et al. A relação entre orientador e orientando no processo de produção científica. *Brazilian Journal of Development*, v. 6, n. 1, p. 3854-3868, 2020. Doi: <https://doi.org/10.34117/bjdv6n1-273>.

Marini, G.; Meschitti, V. The trench warfare of gender discrimination: evidence from academic promotions to full professor in Italy. *Scientometrics*, v. 115, n. 2, p. 989-1006, 2018. Doi: <https://doi.org/10.1007/s11192-018-2696-8>.

Moschkovich, M.; Almeida, A. M. F. Desigualdades de gênero na carreira acadêmica no Brasil. *Dados*, v. 58, n. 3, p. 749-789, 2015. Doi: <http://dx.doi.org/10.1590/00115258201558>.

Olinto, G. A inclusão das mulheres nas carreiras de ciência e tecnologia no Brasil. *Inclusão Social*, v. 5, n. 1, p. 68-77,

2011. Disponível em: <http://revista.ibict.br/inclusao/article/view/1667>. Acesso em: 10 ago. 2020.

Painel Lattes. *Distribuição por sexo, faixa etária e grande área de atuação*. Brasília: CNPq, 2016. Disponível em: <http://estatico.cnpq.br/painelLattes/sexofaixaetaria/>. Acesso em: 20 jan. 2020.

Santos, S. A. D.; Oliveira, M. A produção científica sobre língua brasileira de sinais (LIBRAS) presente nos currículos Lattes do CNPq. *Perspectivas em Ciência da Informação*, v. 22, n. 4, p. 23-46, 2017. Disponível em: <https://periodicos-des.cecom.ufmg.br/index.php/pci/article/view/22535>. Acesso em: 17 ago. 2020.

Scott, J. W. Gênero: uma categoria útil de análise histórica. *Educação e Realidade*, v. 20, n. 2, p. 71-99, 1995. Disponível em: <https://www.seer.ufrgs.br/educacaoerealidade/article/viewFile/71721/40667>. Acesso em: 18 ago. 2020.

Silva, F. F.; Ribeiro, P. R. C. Trajetórias de mulheres na ciência: "ser cientista" e "ser mulher". *Ciência e Educação*, v. 20, n. 2, p. 449-466, 2014. Doi: <http://dx.doi.org/10.1590/1516-73132014000200012>.

Tuesta, E. F. et al. Análise da participação das mulheres na ciência: um estudo de caso da área de ciências exatas e da terra no Brasil. *Em Questão*, v. 25, n. 1, p. 37-62, 2019. Disponível em: <https://www.seer.ufrgs.br/EmQuestao/article/view/80193>. Acesso em: 18 ago. 2020.

Valentova, J. V. et al. Underrepresentation of women in the senior levels of Brazilian science. *PeerJ*, v. 5, p. 1-20, 2017. Doi: <https://doi.org/10.7717/peerj.4000>.