CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL

UMA ANÁLISE BIBLIOMÉTRICA DA TRAJETÓRIA DAS PESQUISADORAS NA CIÊNCIA BRASILEIRA

MONIQUE DE OLIVEIRA SANTIAGO

Orientador: Prof. Dr. Thiago Magela Rodrigues Dias
Centro Federal de Educação Tecnológica de Minas Gerais

Coorientador: Prof. Dr. Alisson Marques da Silva
Centro Federal de Educação Tecnológica de Minas Gerais

BELO HORIZONTE
FEVEREIRO DE 2021

MONIQUE DE OLIVEIRA SANTIAGO

UMA ANÁLISE BIBLIOMÉTRICA DA TRAJETÓRIA DAS PESQUISADORAS NA CIÊNCIA BRASILEIRA

Dissertação apresentada ao Programa de Pós-graduação em Modelagem Matemática e Computacional do Centro Federal de Educação Tecnológica de Minas Gerais, como requisito parcial para a obtenção do título de Mestre em Modelagem Matemática e Computacional.

Área de concentração: Modelagem Matemática e Computacional

Linha de pesquisa: Sistemas Inteligentes

Orientador: Prof. Dr. Thiago Magela Rodrigues Dias
Centro Federal de Educação Tecnológica de Minas Gerais

Coorientador: Prof. Dr. Alisson Marques da Silva
Centro Federal de Educação Tecnológica de Minas Gerais

CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL
BELO HORIZONTE
FEVEREIRO DE 2021

Santiago, Monique de Oliveira
S235a Uma análise bibliométrica da trajetória das pesquisadoras na ciência brasileira / Monique de Oliveira Santiago. – 2021.
91 f.

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Modelagem Matemática e Computacional.
Orientador: Thiago Magela Rodrigues Dias.
Coorientador: Alisson Marques da Silva.
Dissertação (mestrado) – Centro Federal de Educação Tecnológica de Minas Gerais.

1. Mulheres na ciência – Estatística – Teses. 2. Curriculum vitae – Estatística – Teses. 3. Pesquisa científica – Estatística – Teses. 4. Tecnologia – Estatística – Teses. 5. Identidade de gênero no trabalho – Teses. I. Dias, Thiago Magela Rodrigues. II. Silva, Alisson Marques da. III. Centro Federal de Educação Tecnológica de Minas Gerais. IV. Título.

CDD 519.6

Elaboração da ficha catalográfica pela bibliotecária Jane Marangon Duarte, CRB 6º 1592 / Cefet/MG

SERVIÇO PÚBLICO FEDERAL
MINISTÉRIO DA EDUCAÇÃO
CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
COORDENAÇÃO DO PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL

"UMA ANÁLISE BIBLIOMÉTRICA DA TRAJETÓRIA DAS PESQUISADORAS NA CIÊNCIA BRASILEIRA"

Dissertação de Mestrado apresentada por **Monique de Oliveira Santiago**, em 23 de fevereiro de 2021, ao Programa de Pós-Graduação em Modelagem Matemática e Computacional do CEFET-MG, e aprovada pela banca examinadora constituída pelos professores:

Prof. Dr. Thiago Magela Rodrigues Dias (Orientador)
Centro Federal de Educação Tecnológica de Minas Gerais

Prof. Dr. Anisson Marques da Silva (Coorientador)
Centro Federal de Educação Tecnológica de Minas Gerais

Profª. Drª. Samile Andréa de Souza Vanz
Universidade Federal do Rio Grande do Sul

Profª. Drª. Elisângela Martins de Sá
Centro Federal de Educação Tecnológica de Minas Gerais

Visto e permitida a impressão,

Prof. Dr. Thiago de Souza Rodrigues
Coordenador do Programa de Pós-Graduação em
Modelagem Matemática e Computacional

Dedico este trabalho a meu marido Saulo e àqueles (as) que estiveram ao meu lado em toda esta trajetória.

Agradecimentos

Agradeço a Deus pelo dom da vida e por sempre abençoar meus passos.

À minha família por todo o suporte e incentivo durante essa longa caminhada, motivando-me a evoluir constantemente. Em especial, ao meu marido pela paciência e compreensão.

Ao meu orientador Professor Dr. Thiago Magela Rodrigues Dias por incentivar e acreditar no desenvolvimento deste projeto, conduzindo-me com toda atenção e dedicação.

Ao meu co-orientador Professor Dr. Alisson Marques da Silva, Professores Drª. Elisângela Martins e Dr. Gray Farias pelas contribuições e sugestões que incentivaram avançar a pesquisa.

À instituição CEFET-MG, que me proporcionou estrutura adequada para que eu pudesse exercer as minhas atividades.

À CAPES pelo financiamento desta pesquisa.

A todos os amigos, principalmente Jéssica Lopes e Fernanda Coimbra, que torceram por mim e contribuíram direta ou indiretamente para a realização deste trabalho.

“Que nada nos limite, que nada nos defina, que nada nos sujeite. Que a liberdade seja nossa própria substância, já que viver é ser livre.”
(Simone de Beauvoir)

Resumo

A temática mulheres na ciência vem sendo foco de diversos estudos que buscam entender o papel da mulher e suas variadas relações com a ciência, trazendo abordagens que analisam a participação científica e tecnológica e buscam compreender os aspectos de sua trajetória acadêmica. Neste contexto, este trabalho objetivou analisar a participação das mulheres na ciência brasileira utilizando como fonte de dados o conjunto de doutoras que possuem currículos cadastrados na Plataforma Lattes e cujo gênero registrado em seus currículos seja do sexo feminino. Após a coleta dos dados curriculares da Plataforma Lattes, foi realizada a seleção e tratamento dos dados para caracterizar o conjunto a ser analisado, obtendo um total de 125.515 currículos com gênero feminino informado e com doutorado concluído. Os dados das doutoras foram analisados, e os resultados são apresentados sob diversos aspectos como orientações, produções bibliográficas e técnicas, áreas de atuação e formação acadêmica. Em sequência, as doutoras que possuem bolsas de produtividade em pesquisa no CNPq também foram analisadas, possibilitando verificar como ocorre a colaboração científica deste conjunto, bem como, no conjunto de todas as doutoras. Logo, destaca-se que, estudar os diversos aspectos da mulher em geral e particularmente na ciência, além de ser relevante, possibilita compreender como tem evoluído suas pesquisas, podendo contribuir para uma melhor compreensão da participação das mulheres no âmbito da ciência brasileira, bem como, para formulação de políticas e estratégias que potencialize a atuação das mulheres nos ambientes acadêmicos e científicos.

Palavras-chave: Mulher na ciência. Plataforma Lattes. Produção científica e tecnológica. Gênero feminino. Bibliometria.

Abstract

The theme of women in science has been the focus of several studies that seek to understand women's role and their varied relations with science, bringing approaches that analyze scientific and technological participation and seek to understand aspects of their academic trajectory. In this context, this study aimed to analyze the participation of women in Brazilian science using as a data source the set of PhDs who have curricula registered in the Lattes Platform and whose gender registered in their curriculum is female. After the collection of curricular data from the Lattes Platform, the selection and treatment of the data was carried out to characterize the set to be analyzed, obtaining a total of 125,515 curriculum with informed female gender and completed doctorate degree. The data of the PhDs female were analyzed, and the results are presented under several aspects such as orientations, bibliographical and technical productions, areas of expertise and academic training. In sequence, the PhDs female who have research productivity grants at CNPq were also analyzed, making it possible to verify how the scientific collaboration of this group occurs, as well as, in the group of all PhDs female. Therefore, it is highlighted that studying the various aspects of women in general and particularly in science, in addition to being relevant, makes it possible to understand how their research has evolved, and may contribute to a better understanding of the participation of women in the scope of Brazilian science, as well as as for formulating policies and strategies that enhance the performance of women in academic and scientific environments.

Keywords: Woman in science. Lattes Platform. Scientific and technological production. Feminine gender. Bibliometrics.

Lista de Figuras

Figura 1 – Visão geral do LattesDataXplorer 25
Figura 2 – Processo de seleção pelo critério de gênero, preparação e visualização dos dados 27
Figura 3 – Fragmento da estrutura dos currículos XML do RCDP 28
Figura 4 – Orientações 37
Figura 5 – Distribuição de bolsas de estudo quanto as orientações concluídas . . . 38
Figura 6 – Distribuição de bolsas de estudo quanto as orientações em andamento 39
Figura 7 – Orientador principal e coorientador: mestrado e doutorado 42
Figura 8 – Sumarização das produções bibliográficas e técnicas por ano 44
Figura 9 – Quantitativo das produções das doutoras por conjuntos 46
Figura 10 – Grandes áreas de atuação 50
Figura 11 – Artigos em anais de congresso e em periódicos por grandes áreas . . . 51
Figura 12 – Artigos em anais de congresso e em periódicos per capita por grandes áreas 51
Figura 13 – Tempo de conclusão da graduação, mestrado e doutorado 54
Figura 14 – Publicações por grandes áreas e formação das doutoras 57
Figura 15 – Publicações por grandes áreas e formação das doutoras (continuação) . 58
Figura 16 – Evolução da formação acadêmica através das grandes áreas: mestrado 60
Figura 17 – Evolução da formação acadêmica através das grandes áreas: doutorado 61
Figura 18 – Rede geral das doutoras contendo todos os nós 67
Figura 19 – Componente gigante da rede contendo todos os nós 68
Figura 20 – *Clusters* identificados 69
Figura 21 – Produção bibliográfica das bolsistas e das doutoras 72
Figura 22 – Quantitativo das produções bibliográficas das bolsistas por conjuntos . . 73
Figura 23 – Redes de colaboração das bolsistas por categorias/níveis 74
Figura 24 – Redes de colaboração das bolsistas por categorias/níveis (continuação) 75

Lista de Tabelas

Tabela 1 – Leis e princípios bibliométricos, seus focos de estudo, principais aplicações e áreas de interesse 8
Tabela 2 – Percentual de orientações preenchidas por grandes áreas 40
Tabela 3 – Crescimento das orientações concluídas 41
Tabela 4 – Quantitativo para as produções bibliográficas e técnicas 44
Tabela 5 – Percentual de atualização dos currículos por produção 48
Tabela 6 – Percentual de atualização dos currículos por grandes áreas 52
Tabela 7 – Tempo de conclusão das formações acadêmicas por grandes áreas . . 55
Tabela 8 – Publicações após a conclusão do doutorado por ano 64
Tabela 9 – Publicações após a conclusão do doutorado por ano (continuação) . . . 65
Tabela 10 – Doutoras bolsistas de produtividade em pesquisa do CNPq 71
Tabela 11 – Métricas adotadas 76

Lista de Abreviaturas e Siglas

ABC	Academia Brasileira de Ciências
BITNET	*Because It's Time to NETwork* ou *Because It's There NETwork*
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico
C&T	Ciência e Tecnologia
DGPB	Diretório dos Grupos de Pesquisa no Brasil
DOS	*Disk Operating System* ou Sistema Operacional em Disco
HTML	*HyperText Markup Language*
RCD	*Repositório de Currículos de todos Doutores*
RCDP	*Repositório de Currículos das Doutoras Pesquisadoras*
URL	*Uniform Resource Locator*
WEB	*World Wide Web*
XML	*eXtensible Markup Language*

Sumário

1 – Introdução 1
1.1 Descrição do problema 3
1.2 Motivação 3
1.3 Objetivos 4
1.3.1 Objetivo geral 4
1.3.2 Objetivos específicos 4
1.4 Organização do trabalho 4

2 – Fundamentação Teórica 6
2.1 Bibliometria 6
2.2 Plataforma Lattes 9
2.3 Colaboração científica 12

3 – Trabalhos Relacionados 16
3.1 Mulheres na ciência 16
3.2 Estudos envolvendo dados da Plataforma Lattes 19
3.3 Diferenciais propostos 22

4 – Metodologia 24
4.1 Tipo do estudo 24
4.2 Dados da Plataforma Lattes 24
4.2.1 Aquisição dos dados e seleção por formação acadêmica/titulação . 24
4.2.2 Seleção pelo critério de gênero, preparação e visualização dos dados 26
4.2.3 Perfil de orientações das doutoras 29
4.2.4 Representatividade da produção bibliográfica e técnica das doutoras 30
4.2.5 Distribuição das doutoras em suas grandes áreas de atuação 32
4.2.6 Evolução da formação acadêmica das doutoras 32
4.2.7 Publicações das doutoras após a conclusão do doutorado 33
4.2.8 Colaboração científica das doutoras 34
4.2.9 Produção bibliográfica das bolsistas de produtividade em pesquisa do CNPq 34
4.3 Limitações da pesquisa 35

5 – Análise e Discussão dos Resultados 36
5.1 Perfil de orientações das doutoras 36
5.2 Representatividade da produção bibliográfica e técnica das doutoras 43

5.3 Distribuição das doutoras em suas grandes áreas de atuação 49
5.4 Evolução da formação acadêmica das doutoras 53
5.5 Publicações das doutoras após a conclusão do doutorado 63
5.6 Colaboração científica das doutoras 66
5.7 Produção bibliográfica das bolsistas de produtividade em pesquisa do CNPq 70

6 – Conclusão 78
6.1 Publicações 82
6.2 Trabalhos futuros 83

Referências 85

1 Introdução

A ciência pode ser considerada uma ferramenta essencial para a busca de conhecimento, que objetiva compreender, questionar e aprimorar os processos e métodos científicos, bem como entender os fenômenos da natureza e sua interação com o meio. A produção científica é resultado desta investigação científica e parte integrante do processo de conhecimento. O desenvolvimento de diversos estudos proporcionou um considerável crescimento das produções nos últimos anos e, atualmente, tem sido utilizada a Web para facilitar e disseminar este conhecimento, permitindo aos usuários disponibilizar e acessar os resultados das pesquisas e produções científicas, bem como, registrar informações pessoais e profissionais em sítios da Web.

Em um contexto brasileiro, a Plataforma Lattes[1] do CNPq[2] é o sítio que integra em um único sistema as bases de dados de Currículos, de Grupos de Pesquisa e de Instituições de Ensino do país. O repositório de dados curriculares é considerado um importante conjunto de dados científicos nacional, fornecendo informações de alta qualidade e possibilitando pesquisar dados dos indivíduos que estão ali cadastrados, como formação acadêmica, produção científica, dentre outros (LANE, 2010). Desse modo, por sua riqueza de informações e sua crescente confiabilidade e abrangência, as informações curriculares da Plataforma Lattes podem ser utilizadas para compreender a evolução da ciência, a produção e colaboração científica, pois, contém a trajetória acadêmica de grande parte da comunidade científica brasileira.

Normalmente os dados relacionados a produção científica estão presentes em diversos repositórios, dificultando assim a recuperação e análise dos dados. Entretanto, esse processo pode ser facilitado pela utilização dos dados curriculares da Plataforma Lattes. Como este repositório está disponível em acesso aberto, possibilita que os pesquisadores atualizem seus currículos e outros consultem os currículos através da Web (ARRUDA et al., 2009). Logo, é possível recuperar as informações incluídas nos currículos pelos próprios indivíduos e utilizar este amplo conjunto de dados para realizar análises diversas. Permitindo assim, visualizar e compreender sobre a evolução da ciência tendo como fonte de dados a produção científica dos registros cadastrados, bem como, fazer um levantamento da trajetória acadêmica e profissional dos pesquisadores. É possível também analisar dados da produção científica dos currículos cadastrados na Plataforma Lattes com o objetivo de reconhecer tendências e crescimento da produção científica, prever padrões de pesquisa, contribuir com as políticas de auxílio à pesquisa e entender como ocorre a evolução da ciência brasileira (DIAS, 2016).

[1] Plataforma Lattes: http://lattes.cnpq.br/
[2] CNPq: http://www.cnpq.br/

Em geral, os grandes repositórios de dados disponíveis atualmente possuem características próprias, padrões únicos, quantidade e diversidade de dados, ocasionando assim uma tarefa complexa para realizar um estudo e explorar esses dados. Neste cenário, tem sido utilizado a bibliometria que auxilia no processo de quantificação da comunicação escrita, utilizando métodos para análises estatísticas sobre a produção e disseminação do conhecimento aplicadas a fontes de dados científicos (ARAÚJO, 2006) e permite dispor indicadores para o planejamento nacional e evolução das pesquisas científicas.

Os dados de produção científica têm sido utilizados para gerar vários estudos com diversas abordagens, e um que tem recebido destaque entre os pesquisadores refere-se ao gênero. Gênero pode ser definido como a conexão de duas proposições: "[...] (1) o gênero é um elemento constitutivo de relações sociais baseadas nas diferenças percebidas entre os sexos e (2) o gênero é uma forma primária de dar significado às relações de poder" (SCOTT, 1995, p. 86), sendo assim, corresponde a um campo que possui como temática a identidade e a representação de homens e mulheres na sociedade. Este campo inclui o subcampo Estudo da Mulher que abrange, entre diversas dimensões de interesse, a mulher e suas variadas relações com a ciência (LETA, 2014). Os estudos que tratam de gênero são naturalmente interdisciplinares e multidisciplinares, abrangem diversas áreas do conhecimento científico (HOPPEN; VANZ, 2020, p. 2, tradução nossa)[3], e resultam em inúmeras discussões complexas sobre gênero, sexo e identidade de gênero. Contudo, tais questões estão fora do escopo desta dissertação, que considera o sexo dos indivíduos recuperados através dos arquivos do Diretório dos Grupos de Pesquisa no Brasil (DGPB)[4] para obter um cenário da participação feminina na ciência e tecnologia.

Diversos estudos contemplam a temática de investigação relacionada a questão de gênero na ciência, apresentando a quantidade de trabalhos relacionados (MEDEIROS; HOPPEN; VANZ, 2020), ou mesmo, uma reflexão dos artigos através de aspectos psicológicos, sociológicos, históricos, biológicos, incluindo uma análise da participação na produção científica entre homens e mulheres (LETA, 2014), porém, de modo geral, tentam entender o quanto os temas de gênero, mulher e ciência se correlacionam e estabelecem no dia-a-dia das universidades e instituições de pesquisa, sendo estes espaços onde se produz a ciência, produzindo ou reproduzindo as diferenças e as relações de poder (SILVA; RIBEIRO, 2014). Assim, ao analisarmos o contexto de gênero no ambiente científico focando no papel da mulher, tende a ser uma iniciativa significativa para entender o progresso feminino neste segmento, ou seja, identificar e compreender o quanto a mulher está inserida e consolidada no ambiente acadêmico, e quais os aspectos de sua trajetória na ciência.

Logo, utilizando a temática gênero para estudar a mulher na sociedade e a relação dela com a ciência, esse trabalho objetiva analisar a participação científica e tecnológica

[3] "*Gender studies topics are naturally interdisciplinary and multidisciplinary, spanning and encompassing areas in all fields*."

[4] DGPB: http://lattes.cnpq.br/web/dgp/

das doutoras pesquisadoras, investigando como suas pesquisas têm sido realizadas e como sua trajetória acadêmica tem evoluído ao longo dos anos a partir de análises bibliométricas realizadas sobre dados curriculares disponíveis na Plataforma Lattes. Esse estudo, além de apresentar uma visão das mulheres que têm realizado pesquisas, visa expor um cenário geral que pode contribuir para originar indicadores científicos nacionais e para gerenciar informações na área científica e tecnológica, bem como, para formulação de políticas e estratégias que intensifique a atuação das mulheres nos ambientes acadêmicos.

1.1 Descrição do problema

Através da produção científica, que é parte integrante do processo de produção de conhecimento, é possível desenvolver estudos com o objetivo principal de entender como a ciência tem evoluído e como a colaboração científica ocorre. Mugnaini, Digiampietri e Mena-Chalco (2014) ressaltam que, ao realizar o levantamento da produção científica de um país é possível estudar os resultados mensuráveis da ciência, tecnologia e inovação. Do mesmo modo que, ao acompanhar o fluxo de comunicação científica entre as áreas auxilia o processo de avaliação dos resultados das pesquisas. Assim, realizar um estudo para compreender os diversos aspectos científicos das mulheres é uma iniciativa significativa para conhecer a participação e o cenário atual do papel delas na ciência. Segundo Valentova et al. (2017), são claras as evidências de desigualdade entre os gêneros na Ciência Brasileira, o que provavelmente tem profundas raízes institucionais e culturais. Assim, faz necessário realizar políticas que se baseiam nas pesquisas e que possam melhorar a perspectiva e de atuação das mulheres na ciência.

Alguns mitos podem influenciar o sucesso na carreira científica das mulheres: os homens são mais produtivos que mulheres; homens têm melhor desempenho em matemática e outras ciências ditas como “duras”; homens são mais bem sucedidos nas atividades de ciência e tecnologia que mulheres; homens são melhor apresentados na Web (KRETSCHMER; KRETSCHMER, 2013). Porém, quando esses mitos se tornam persistentes, criam-se preconceitos ou atitudes em relação à avaliação do desempenho científico das mulheres, ou seja, a perpetuação destes mitos torna-se desvantajoso para elas. Por isso, torna-se tão relevante realizar estudos que reflitam e buscam novos meios de investigar sobre o assunto que permeiam as mulheres e a ciência.

1.2 Motivação

Um dos fatores que motiva o estudo para a compreensão da participação científica das mulheres é a oportunidade de entender o cenário de atuação das doutoras e apresentar seus diversos aspectos na ciência. Uma vez que, ao reconhecer que “[...] a ciência é neutra com relação às questões de gênero, revelando que os valores e as características

socialmente atribuídos às mulheres são desvalorizados na produção do conhecimento, e que desigualdades de gênero perpassam o campo científico [...]"(SILVA; RIBEIRO, 2014, p. 464), admite-se a necessidade de quebrar as barreiras das desigualdades e garante assim a excelência da pesquisa ao reter seus melhores cientistas e inovadores (VALENTOVA et al., 2017) e, portanto, possibilita "[...] criar uma ciência mais humana, livre das transformações causadas pelos séculos de exclusão de mais da metade da humanidade, as mulheres" (LETA, 2014, p. 150).

1.3 Objetivos

1.3.1 Objetivo geral

Analisar a participação científica e tecnológica das doutoras pesquisadoras brasileiras, investigando como suas pesquisas têm sido realizadas e como sua trajetória acadêmica tem evoluído ao longo dos anos a partir de análises bibliométricas realizadas sobre dados curriculares disponíveis na Plataforma Lattes.

1.3.2 Objetivos específicos

- Fazer uma revisão de literatura sobre a participação das mulheres na ciência e estudos envolvendo dados da Plataforma Lattes;
- Obter, selecionar e preparar os dados curriculares da Plataforma Lattes;
- Traçar o perfil de orientações das doutoras;
- Analisar a representatividade da produção bibliográfica e técnica das doutoras;
- Fazer uma distribuição das doutoras em suas grandes áreas de atuação;
- Compreender a evolução da formação acadêmica das doutoras;
- Investigar as publicações das doutoras após a conclusão do doutorado;
- Analisar a colaboração científica das doutoras;
- Verificar a produção bibliográfica das bolsistas de produtividade em pesquisa do CNPq.

1.4 Organização do trabalho

O restante desta Dissertação está organizada da seguinte forma:

- No Capítulo 2, são apresentados os assuntos permeados por esta Dissertação, tais como: os conceitos e principais métricas adotadas para análise bibliométrica; o contexto histórico e informação sobre a estrutura dos currículos da Plataforma Lattes, conceitos sobre a colaboração científica e principais métricas relacionadas as redes.

- No Capítulo 3, são apresentados os trabalhos relacionados que abordam os assuntos presentes nesta Dissertação.
- No Capítulo 4, é apresentada a metodologia de pesquisa adotada durante o estudo a fim de atingir os objetivos propostos, bem como, as limitações da pesquisa.
- No Capítulo 5, é apresentada uma análise da produção científica do conjunto das doutoras pesquisadoras com currículos cadastrados na Plataforma Lattes, a partir da coleta realizada em outubro de 2019. Essa análise inclui uma visão geral do conjunto analisado, sendo a seguir realizado um estudo sobre a evolução da produção científica dessas mulheres ao longo dos anos.
- Por fim, no Capítulo 6, são apresentadas as considerações finais da Dissertação, destacando-se as suas principais contribuições e os resultados obtidos. Esse capítulo apresenta ainda a lista de publicações resultantes dos trabalhos realizados no decorrer da elaboração desta Dissertação.

2 Fundamentação Teórica

Este capítulo apresenta um estudo bibliográfico referente aos assuntos permeados por esta Dissertação. Nas próximas seções, são apresentados: conceitos de bibliometria e principais métricas adotadas para análise bibliométrica (Seção 2.1), contexto histórico e informação sobre a estrutura dos currículos da Plataforma Lattes (Seção 2.2), conceitos sobre colaboração científica e principais métricas relacionadas as redes (Seção 2.3).

2.1 Bibliometria

Os resultados das pesquisas podem ser divulgados através de canais formais como livros, capítulos de livros, artigos científicos, ou canais informais como teses, dissertações, comunicações em anais de eventos científicos. Estas atividades científicas e tecnológicas podem ser recuperadas, estudadas e avaliadas aplicando-se métodos bibliométricos que buscam construir indicadores de produção e de desempenho científico. Avaliar a ciência é uma tarefa complexa que exige domínio de conhecimentos em diferentes áreas, tais como a Ciência da Informação e a Sociologia da Ciência (SILVA; HAYASHI; HAYASHI, 2011).

A produtividade científica pode ser avaliada através de técnicas como bibliometria, cientometria, informetria e webometria. Ao mesmo tempo que essas técnicas possuem funções semelhantes, cada uma delas propõe medir a disseminação do conhecimento científico e o fluxo da informação utilizando enfoques diversos (VANTI, 2002). Contudo, a diferença essencial entre essas técnicas corresponde ao fato da bibliometria utilizar mais métodos quantitativos do que discursivos (PRITCHARD et al., 1969), e por isso, pode-se considerar a bibliometria como a área responsável por avaliar a ciência ao basear em dados da produção científica dos pesquisadores. Dias (2016) ressalta que, com seus diferentes enfoques, a análise bibliométrica pode ser utilizada para mensurar a produção científica de indivíduos em particular, grupos de pesquisa, instituições, regiões geográficas, organizações ou eventos.

O uso do termo bibliometria foi popularizado por Pritchard et al. (1969), quando sugeriu substituir em seu trabalho, o termo “bibliografia estatística” por “bibliometria”. Para ele, a bibliometria pode ser entendida como um conjunto de técnicas e métodos para tratar as informações científicas. Já Araújo (2006) ressalta o termo bibliometria como sendo uma ferramenta utilizada para obter informações quantitativas e estatísticas sobre a produção e disseminação do conhecimento aplicadas a fontes de dados científicos. Portanto, a bibliometria pode ser aplicada em diversas áreas do conhecimento na busca de informações sobre o desenvolvimento científico.

Segundo Vanti (2002, p. 155), as vantagens em se utilizar análises bibliométricas para medir o registro da informação e a difusão do conhecimento, são:

> Identificar as tendências e o crescimento do conhecimento em uma área; identificar as revistas do núcleo de uma disciplina; mensurar a cobertura das revistas secundárias; identificar os usuários de uma disciplina; prever as tendências de publicação; estudar a dispersão e a obsolescência da literatura científica; prever a produtividade de autores individuais, organizações e países; medir o grau e padrões de colaboração entre autores; analisar os processos de citação e co-citação; determinar o desempenho dos sistemas de recuperação da informação; avaliar os aspectos estatísticos da linguagem, das palavras e das frases; avaliar a circulação e uso de documentos em um centro de documentação; medir o crescimento de determinadas áreas e o surgimento de novos temas.

A bibliometria possui três leis clássicas, que são: Lei de Lotka, Lei de Bradford e Lei de Zipf. A Lei de Lotka ou Lei do Quadrado Inverso corresponde a produtividade científica de autores e pode ser determinada pela descoberta de que um número menor de pesquisadores, supostamente de maior influência em uma determinada área do conhecimento, possui grande parte da produção científica, e muitos pesquisadores, supostamente de menor prestígio, produzem pouco. Assim, em termos de produção, um grande número de pesquisadores que pouco publicam, se iguala ao pequeno número de pesquisadores mais produtivos. A Lei de Bradford ou Lei da Dispersão manifesta-se sobre o conjunto de periódicos e possui o objetivo de encontrar a extensão na qual artigos científicos de assunto específico estavam presentes em periódicos destinados a outros assuntos. É possível estimar o grau de relevância de periódicos em uma dada área do conhecimento e que os periódicos de maior qualidade ou relevância para aquela área, produzem um maior número de artigos sobre um dado assunto, formando um núcleo. A Lei de Zipf ou Lei do Menor Esforço corresponde a frequência de palavras, com o objetivo de descrever a ordem e a relação entre palavras de um determinado texto. Ao encontrar a correlação entre o número de palavras distintas e a frequência de seu uso, Zipf concluiu que existe uma regularidade na seleção e uso destas, em que um pequeno número de palavras é utilizado muitas vezes em detrimento de outras. Assim, formulou-se o princípio do menor esforço: o assunto do documento é indicado pelas palavras com maior frequência (ARAÚJO, 2006).

Diferentes abordagens baseiam-se nas três leis clássicas da bibliometria, utilizando-as individualmente ou combinadas, como meio de realizar análises complexas e assim, buscar conhecimento sobre a produção científica. Uma dessas abordagens corresponde a contabilização de publicações de um autor ou de um grupo de autores, que tem o objetivo de se obter total e média de publicações, para assim avaliar, ranquear e também classificar os autores (DIAS, 2016). Uma outra abordagem refere-se a análise de citações do conjunto de trabalhos publicados, no qual objetiva identificar padrões do conhecimento científico e descobrir:

> Autores mais citados, autores mais produtivos, elite de pesquisa, frente de pesquisa, fator de impacto dos autores, procedência geográfica e/ou institucional dos autores mais influentes em um determinado campo de pesquisa; tipo de documento mais utilizado, idade média da literatura utilizada, obsolescência da literatura, procedência geográfica e/ou institucional da bibliografia utilizada; periódicos mais citados, "core" de periódicos que compõem um campo (ARAÚJO, 2006, p. 18–19).

Como complemento às técnicas de análise bibliométrica clássicas, as autoras Guedes e Borschiver (2005) destacam em seu estudo outras teorias, no qual relacionam as principais leis e princípios bibliométricos, seus focos de estudo e suas principais aplicações na gestão da informação, em sistemas de informação e em comunicação científica e tecnológica, conforme apresentado na Tabela 1.

Tabela 1 – Leis e princípios bibliométricos, seus focos de estudo, principais aplicações e áreas de interesse

Leis e Princípios	**Focos de Estudo**	**Principais Aplicações**
Lei de Bradford	Periódicos	Estimar grau de relevância de periódicos em dada área do conhecimento.
Lei de Lotka	Autores	Estimar grau de relevância de autores em dada área do conhecimento.
Lei de Zipf	Palavras	Indexação automática de artigos científicos e tecnológicos.
Ponto de Transição (T) de Golfam	Palavras	Indexação automática de artigos científicos e tecnológicos.
Colégios Invisíveis	Citações	Identificação da elite de pesquisadores em dada área do conhecimento.
Fator de imediatismo	Citações	Estimar o grau de relevância de artigos, cientistas e periódicos científicos em determinada área do conhecimento.
Acoplamento Bibliográfico	Citações	Estimar o grau de ligação de dois ou mais artigos.
Co-citação	Citações	Estimar o grau de ligação de dois ou mais artigos.
Obsolescência da Literatura	Citações	Estimar o declínio da literatura de determinada área do conhecimento.
Vida-média	Citações	Estimar a vida-média de uma unidade de literatura de dada área do conhecimento.
Teoria Epidêmica de Golfmam	Citações	Estimar a razão de crescimento e declínio de determinada área do conhecimento.
Lei do Elitismo	Citações	Estimar a razão de crescimento e declínio de determinada área do conhecimento.
Frente de Pesquisa	Citações	Identificação de um padrão de relação múltipla entre autores que se citam.
Lei dos 80/20	Demanda de Informação	Composição, ampliação e redução de acervos.

Fonte: Adaptado de Guedes e Borschiver (2005)

O uso de indicadores bibliométricos tem contribuído para compreender os diversos aspectos, procedimentos e estruturas da atividade científica, de seu impacto, bem como,

das relações existentes entre os autores no caso da colaboração científica. Edler e Flanagan (2011) ressaltam a importância de se utilizar indicadores bibliométricos para estabelecer políticas científicas como uma maneira de internacionalizar a ciência.

Diante disso, como a bibliometria "[...] compreende a aplicação de análises estatísticas para estudar as características do uso e criação de documentos" (SPINAK, 1998, p. 142), ao considerarmos que os indicadores bibliométricos são medidas obtidas com base em dados de produções e citações e que ela se enquadra como uma abordagem quantitativa, as análises podem ser aplicadas em diversas áreas do conhecimento, possibilitando realizar estudos quantitativos e estatísticos com o objetivo de investigar e medir a produção científica.

2.2 Plataforma Lattes

O contexto histórico da Plataforma Lattes se remete a meados dos anos 80, sendo criado nesta época um sistema informatizado, denominado Banco de Currículos, para digitação dos currículos dos pesquisadores brasileiros que continham seus dados em um formulário de papel. No final dos anos 80, o CNPq já disponibilizava, através da rede remota BITNET, as buscas sobre a base de aproximadamente 30.000 currículos às universidades e instituições de pesquisa do país. No início dos anos de 90, a captação dos dados curriculares ocorreu através do formulário eletrônico BCUR, desenvolvido para o Sistema Operacional DOS, no qual os pesquisadores preenchiam o formulário e enviavam em disquete ao CNPq que carregava os dados para a base de dados. No final dos anos 90, o CNPq reuniu universitários, profissionais e técnicos para desenvolver uma única versão de currículo capaz de integrar as já existentes. Assim, em agosto de 1999 foi lançado e padronizado o Currículo Lattes, correspondendo a um formulário de currículo a ser utilizado no âmbito do Ministério da C&T e do CNPq. Desde então, o Currículo Lattes vem ampliando sua abrangência e pode ser utilizado por universidades, institutos e fundações de amparo à pesquisa como instrumento de avaliação de pesquisadores, professores e alunos. A Plataforma Lattes está em constante avaliação, reformulação e aprimoramento, buscando corrigir possíveis desvios e promovendo o aperfeiçoamento da ferramenta (PLATAFORMA LATTES, 2019).

Portanto, o CNPq através de seus consultores desenvolveu a Plataforma com o objetivo de unificar o cadastro nacional de pesquisadores e integrá-los às agências de fomento do Brasil (FERRAZ; QUONIAM; MACCARI, 2015), e agrega em um único sistema as bases de dados de Currículos, de Grupos de Pesquisa e de Instituições de Ensino do país. A Plataforma Lattes surge como um padrão para registro das informações curriculares da comunidade científica nacional por sua riqueza de informações e sua crescente confiabilidade e abrangência, tornando-se assim um elemento indispensável e

compulsório à análise de mérito e competência para as demandas de financiamentos na área de ciência e tecnologia (DIGIAMPIETRI, 2015). É uma base de dados pública, tanto no que se refere a inserção de dados quanto à recuperação das informações através da Web.

A Plataforma Lattes do CNPq pode ser acessada através do endereço eletrônico http://lattes.cnpq.br/. A página inicial contém menus, notícias, acesso direto que facilitam a navegação dos usuários, *links* para outros sítios que contém alguma compatibilidade com a plataforma ou a área acadêmica. Os cinco itens que aparecem no menu principal direcionam os usuários para níveis internos que contém informações referentes à plataforma, dados, acordos institucionais, outras bases e ajuda ao usuário. O acesso direto contém quatro grupos, “Currículo Lattes”, “Diretório de Instituições”, “Diretório de Grupos de Pesquisa” e “Painel Lattes”, com seus respectivos tópicos.

O primeiro tópico “Currículo Lattes” possui o acesso direto a inserção, atualização e pesquisa dos dados referente ao currículo da Plataforma Lattes. Ao cadastrar novo currículo é necessário preencher informações como CPF, e-mail, senha, informações pessoais, endereço, formação acadêmica, grandes áreas de atuação. Os campos de preenchimento do currículo referentes às informações pessoais e que possuem uma imagem de cadeado, corresponde a informações que não serão exibidas na consulta pública. A opção “Buscar currículo” apresenta uma interface de consulta de currículos bastante flexível e possibilita a busca simples ou avançada, bem como, uma variedade de parâmetros e filtros que podem ser selecionados conforme a necessidade da pesquisa. Ao realizar uma busca, são geradas listas de currículos que atendem aos parâmetros informados. Ao atualizar um currículo da Plataforma Lattes faz-se necessário inserir o Login (CPF ou e-mail) do usuário e a senha. Com exceção do CPF, todas as outras informações podem ser alteradas. Vale ressaltar que, um único currículo é cadastrado para cada CPF, pois o currículo é vinculado ao CPF do proprietário. O currículo da Plataforma Lattes possui doze módulos para inserção e/ou alteração de informações, que são listados abaixo com seus respectivos tópicos mais relevantes:

- Dados gerais: contém identificação pessoal como foto, nome e formas de citação bibliográfica, endereço residencial e profissional, idiomas, lista de prêmios e títulos, texto inicial do Currículo Lattes.
- Formação: compreende os dados referentes às formações e titulações, como, ensinos fundamental, médio e profissional de nível técnico, graduação, especialização, especialização (residência médica), mestrado, mestrado profissional, doutorado, aperfeiçoamento, formações complementares, pós-doutorado e/ou livre-docência.
- Atuação: abrange os dados de atuação profissional, linhas de pesquisa, membro do corpo editorial, membro do comitê de assessoramento, revisor de periódico, revisor de projeto de agência de fomento, grandes áreas de atuação.
- Projetos: contém informações referente aos projetos de pesquisa, de desenvolvimento

tecnológico, de extensão, de ensino e outros tipos de projetos.

- Produções: são distribuídas pelos tópicos produção bibliográfica, produção técnica e outra produção artística/cultural. Cada tópico agrupa diversas informações a respeito da produção bibliográfica (publicações) realizada pelo usuário.
- Patentes e Registros: abrange informações sobre patente, programa de computador registrado, cultivar protegida, cultivar registrada, desenho industrial registrado, marca registrada, topografia de circuito integrado registrada.
- Inovação: compreende informações sobre inovações como programa de computador registrado ou sem registro, marca registrada, produtos, processo ou técnicas, projetos de pesquisa, projeto de desenvolvimento tecnológico, projeto de extensão, projeto de ensino, outros projetos.
- Educação e Popularização de C&T: corresponde aos artigos completos publicados em periódicos, artigos aceito para publicação, livros e capítulos, texto em jornal ou revista, trabalhos publicados em anais de eventos, apresentação de trabalho e palestra.
- Eventos: abrange as informações correspondendo à participação ou organização em eventos, congressos, exposições, feiras e olimpíadas.
- Orientações: abrange informações sobre as orientações e supervisões tanto concluídas quanto em andamento, para dissertação de mestrado, tese de doutorado, monografia de conclusão de curso de graduação, aperfeiçoamento ou especialização, iniciação científica, supervisão de pós-doutorado, e orientação de outra natureza.
- Bancas: inclui informações tanto de participação em bancas de trabalhos de conclusão quanto da participação em bancas de comissões julgadoras.
- Citações: compreende informações quanto a citações no Web of Science (ISI)[1], no SciELO[2], no SCOPUS[3] ou outras bases bibliográficas.

A importância em manter o cadastro e atualizar constantemente o currículo da Plataforma Lattes deve ao fato que os dados curriculares são utilizados para auxiliar na concessão de bolsas e auxílios à pesquisa, e à avaliação da qualidade dos cursos (FERRAZ; QUONIAM; MACCARI, 2015). No caso do Painel Lattes[4], o mesmo usa os dados cadastrados nos currículos dos pesquisadores e fornecem dados estatísticos da produção C&T através de painéis distribuídos nos seguintes tópicos: Distribuição geográfica por área, Ranking de instituições por área, Comparativo por geografia, instituição e área, Distribuição por sexo, faixa etária e área, Distribuição por setor econômico e área, Evolução da formação de mestres e doutores. Esses painéis permitem visualizar os dados estatísticos através gráficos e tabelas de forma fácil e dinâmica. Porém todos os dados disponibilizados foram extraídos em 30 de novembro de 2016, e desde então o portal não forneceu informações atuais.

[1] WoS: https://webofknowledge.com/
[2] SciELO: https://www.scielo.org/
[3] SCOPUS: https://www.elsevier.com/solutions/scopus/
[4] Painel Lattes: http://estatico.cnpq.br/painelLattes/

Como essas informações curriculares estão disponíveis em acesso aberto, é possível recuperar os dados históricos incluídos anteriormente na plataforma pelos próprios indivíduos nela cadastrados e utilizar este amplo conjunto de dados para analisar e compreender a evolução da ciência.

2.3 Colaboração científica

A colaboração pode ser definida como uma cooperação ou um auxílio entre os autores, uma participação em obra alheia ou uma contribuição entre os indivíduos para realizar algo. Segundo Katz e Martin (1997), a colaboração científica corresponde a união do trabalho de pesquisadores, que possuem o objetivo comum de produzir e compartilhar conhecimento científico.

Para entender a complexidade da colaboração científica, faz necessário entender estruturas de redes sociais, definidas como um conjunto de indivíduos ou organizações que possuem elementos conectados por um ou mais tipos de ligações, ou seja, “redes sociais são redes formadas por indivíduos (ou algo que possa ser individualizado) com algum grau de relacionamento” (GABARDO, 2015, p. 20). A representação da rede é realizada através de um conjunto de nós ou vértices indicando indivíduos, unidos em pares por linhas ou arestas indicando conhecimento (NEWMAN, 2001). Computacionalmente, as redes são representadas através de grafos, onde um indivíduo corresponde a um nó ou vértice, e a relação entre eles corresponde a uma aresta, que pode ser direcionada ou não (DIGIAMPIETRI, 2015). Assim, a definição **G(V, E)** possui **V** como um conjunto de vértices ou nós, **E** como um conjunto de arestas e **G** o próprio grafo, no qual cada aresta representa uma conexão entre dois elementos do conjunto **V** (CORMEN et al., 2009). Grafo pode ser definido como uma abstração matemática no qual pode ser utilizado para representar uma rede complexa (GABARDO, 2015).

A colaboração científica entre pesquisadores e publicações é representada através da rede (TIAN et al., 2014), e tornou-se fundamental para a pesquisa e mais intensa nas últimas décadas com a ampliação do consenso de que a ciência é feita através de equipes colaborativas (FELIZARDO, 2018). As redes científicas possuem diversas métricas de análise, que possibilitam explorar técnicas para melhorar a compreensão destas redes. Para explorar todo o conhecimento apresentado pela rede, faz-se necessário conhecer os principais conceitos que permeiam a mesma. Assim, segundo Gabardo (2015) as principais métricas relacionadas a redes complexas são descritas a seguir:

- **Ordem**: corresponde ao número de vértices ou nós de um grafo, ou seja, representa o número de elementos em V na definição de grafo G(V,E).

- **Tamanho**: corresponde ao número de arestas ou conexões entre os vértices de um grafo, ou seja, representa o número de elementos em E na definição de grafo G(V,E).

- **Diâmetro**: corresponde a maior distância geodésica no grafo, ou seja, refere-se ao maior caminho dentre todos os caminhos mínimos computados para todos os vértices de um grafo .

- **Grau de um vértice**: corresponde ao total de conexões que o mesmo possui. Para grafos não direcionados classifica-se o grau de um vértice simplesmente como grau (*degree*) e, para os grafos direcionados classifica-se como grau de entrada (*indegree*) e grau de saída (*outdegree*). Grau de entrada refere-se ao número de conexões que um vértice recebe e grau de saída corresponde ao número de conexões que um vértice realizou.

- **Grau médio da rede**: corresponde ao cálculo da soma do grau de todos os vértices dividido pelo número de vértices existentes. Assim, como cada aresta conecta dois vértices, a função do grau médio de um grafo não dirigido é definida por 2*|E|/|V|.

- **Densidade**: corresponde a relação entre sua ordem e seu tamanho, ou seja, encontramos a densidade do grafo G(V,E) ao medir a quantidade de arestas existentes em E pelo número máximo possível entre vértices em V. Assim, como cada aresta conecta dois vértices, a função da densidade de um grafo é definida por 2E/V.

- **Centralidade de grau**: corresponde à relação ao número de vértices do grafo conectados a um vértice em particular, ou seja, vértices centrais são importantes nos cálculos de medição. A métrica de centralidade de grau também calcula a proximidade de um dado vértice em relação aos demais.

- **Centralidade de proximidade**: corresponde a uma medida que calcula o quão próximo um vértice está dos demais vértices da rede, podendo calcular esta métrica através da soma das distâncias geodésicas em relação a todos os demais vértices da rede. Pode-se utilizar a seguinte equação para calcular a centralidade de proximidade para um vértice:

$$Cc = \frac{n-1}{\sum_{j=1}^{n} d(i,j)}$$

  Em que o valor da centralidade de proximidade é Cc, n representa o número de vértices no grafo, $\sum_{j=1}^{n}$ corresponde a soma de todas as distâncias para todos os vértices, $d(i,j)$ é o número de caminhos mais curtos ligando os vértices i e j.

- **Centralidade de intermediação**: corresponde a uma medida de centralidade baseada no número de caminhos mínimos determinado pelos vértices que passam por um ponto da rede, sendo que, quanto mais central for o vértice, mais caminhos mínimos passarão por ele. Pode-se utilizar a seguinte equação para calcular a centralidade de intermediação:

$$Cb(v) = \sum \frac{\theta ij(v)}{\theta ij}$$

  Em que o grau de centralidade é $Cb(v)$, o número de caminhos mínimos entre os vértices i e j que passam pelo vértice v é representado por $\theta ij(v)$, já θij corresponde ao número de todos os caminhos mínimos entre i e j.

- **Modularidade**: corresponde a medida de estrutura da rede, no qual os grupos de vértices ou *clusters* apresentam alto grau de similaridade ou que estão próximos uns dos outros dentro da rede. Pode-se utilizar a seguinte equação para calcular a modularidade que avalia as conexões entre os vértices:

$$Q = \frac{1}{4m} \sum_{i,j} (A_{ij} - \frac{k_i k_j}{2m})$$

  Em que a medida de modularidade é Q, $\frac{1}{4m}$ é um termo que corresponde a uma fração convencionada por razões de normalização, já a matriz de adjacências é representada por A_{ij}, e o termo $\frac{k_i k_j}{2m}$ calcula a probabilidade de uma aresta estar entre os vértices i e j, no qual m corresponde ao número total de arestas no grafo, k_i representa o grau do vértice i e k_j representa o grau do vértice j.

- **Coeficiente de aglomeração local**: para avaliar a presença de *clusters* em uma rede complexa é utilizada a métrica de coeficiente de aglomeração. O coeficiente de aglomeração local mede a propensão que um vértice específico e seus vizinhos possuem para formar um clique, sendo que clique corresponde a um subgrafo conexo onde todos os vértices estão conectados entre si. Pode-se utilizar a seguinte equação para calcular o coeficiente de aglomeração local:

$$C_i = \frac{2n_i}{k_i(k_i - 1)}$$

  Em que o coeficiente de aglomeração local para o vértice i é C_i, o número de conexões corresponde a n_i, as arestas conectadas que ligam os k_i vértices vizinhos de i até este vértice.

- **Coeficiente de aglomeração global**: ao avaliar a presença de *clusters* em uma rede complexa é utilizada a métrica de coeficiente de aglomeração, onde o coeficiente

de aglomeração global mede a propensão de uma rede em apresentar grupos ou comunidades. O cálculo é baseado no número de trios conectados presentes na rede, sendo que trio forma um triângulo e corresponde ao conjunto de três vértices conectados entre si. Pode-se utilizar a seguinte equação para calcular o coeficiente de aglomeração global:

$$C = \frac{3xNumTriangulosGrafo}{numTriosConectadosGrafo}$$

Em que o coeficiente de aglomeração global é C, $numTriosConectadosGrafo$ representa o trio de vértices constituídos por um nó central que está ligado aos outros dois e os vértices de acompanhamento são não ordenados, já $3xNumTriangulosGrafo$ corresponde ao número de trios conectados que na verdade formam triângulos, sendo que o fator 3 indica que cada triângulo corresponde a três trios.

- **Coeficiente de aglomeração médio**: corresponde à média do coeficiente de agrupamento local de todos os vértices da rede e ajuda a avaliar qual determinado modelo de rede um grafo se assemelha.

$$C = \frac{1}{N}\sum_{i} C_i$$

Em que o número de vértices na rede é representado por N e o coeficiente de aglomeração local de cada um dos vértices do grafo é representado por C_i.

As análises estatísticas representadas por gráficos, bem como as métricas descritas acima vêm sendo utilizadas para qualificar as publicações e provar a importância da rede social e colaboração entre autores. Em um domínio científico, uma rede social pode ser construída com os vértices correspondendo aos autores das publicações e as arestas correspondendo a relação de coautoria.

3 Trabalhos Relacionados

Este capítulo apresenta a síntese de um conjunto de trabalhos referente aos assuntos permeados por esta Dissertação. Nas próximas seções, são apresentados os tópicos a fim de cumprir o primeiro objetivo específico “fazer uma revisão de literatura sobre a participação das mulheres na ciência e estudos envolvendo dados da Plataforma Lattes”: trabalhos relacionados a temática “Mulheres na Ciência” (Seção 3.1); a Plataforma Lattes como principal fonte de dados (Seção 3.2); diferenciais propostos (Seção 3.3).

3.1 Mulheres na ciência

Estudos que focam na mulher e suas variadas relações com a ciência possuem diversas abordagens e apresentam caráter diversos, como o histórico, o social, o econômico, entre outros. Por ser um tema interdisciplinar e abrangente, somente um pequeno grupo de trabalhos englobando especificamente os tópicos mulher, ciência e gênero são apresentados nesta seção.

Os critérios de inclusão utilizados para seleção desses estudos foram: publicações disponíveis on-line a partir do ano 2012 e idiomas inglês e português. A pesquisa dos estudos foi realizada no ano de 2019 e os dois descritores utilizados foram combinados para otimizar a busca: “mulher *and* gênero *and* ciência” e “*woman and gender and science*”. No Portal de periódicos CAPES/MEC[1] a busca retornou 1.575 resultados para o primeiro descritor e 54.136 para o segundo, no Scielo[2] retornou 130 resultados para o primeiro descritor e 31 para o segundo, na revista Scientometrics[3] foi utilizado o segundo descritor e a busca retornou 331 artigos, no Google Acadêmico[4] os dois descritores foram agrupados e retornou 16.300 resultados. Como não foi possível ler todos os resultados, a leitura do título foi realizada nas 30 primeiras páginas e foram selecionados os que tinham relação com a temática, posteriormente realizada a leitura do resumo e excluídos os que não possuíam relação com os tópicos. Foram selecionados os estudos que apresentam uma visão e maior representatividade da mulher na ciência.

Arensbergen, Weijden e Besselaar (2012) analisaram se as diferenças de produtividade por gênero são persistentes ou se mudam ao longo do tempo. O conjunto de dados para análise abrange dentro das ciências sociais na Holambra, um período de três anos (2003 a 2005) e 1.100 pedidos de bolsas de pesquisa dos programas: bolsas de

[1] CAPES/MEC: https://www-periodicos-capes-gov-br.ezl.periodicos.capes.gov.br/index.php?
[2] Scielo: https://scielo.org/
[3] Scientometrics: https://www.springer.com/journal/11192
[4] Scholar: https://scholar.google.com.br/?hl=pt

carreira adiantadas (ECG), avançadas (ACG) e um esquema de concurso aberto (OC). As distribuições quanto ao gênero feminino e masculino referentes as publicações e citações, foram apresentadas em duas fases, primeiro para os programas ACG e OC, depois para a geração mais nova ECG. Como resultados os autores sugerem que as diferenças de desempenho por gênero estão desaparecendo. Na geração ACG e OC, os homens superaram as mulheres em termos de publicações e citações, mas não é mais o caso da geração mais nova ECG. O que permite sugerir que as diferenças de desempenho parecem desaparecer com o tempo, estando de acordo com as experiências em outras partes do sistema educativo, onde as alunas estão cada vez melhores que os alunos.

Silva e Ribeiro (2014) abordam a trajetória acadêmica e profissional de mulheres na ciência, identificando principalmente preconceitos, discriminações, conflitos, dificuldades, conquistas, relações entre trabalho e vida familiar. Foram realizadas entrevistas com seis mulheres cientistas que atuam em universidades públicas e instituição de pesquisa do Rio Grande do Sul. A análise das narrativas evidencia como as entrevistadas defrontam com muitas dificuldades para seguir a carreira científica como dupla jornada de trabalho, maternidade, produtividade em pesquisa, competição, preconceito e discriminação de gênero. As autoras ainda argumentam que a trajetória das entrevistadas na ciência foi e ainda é construída em um ambiente baseado em valores e padrões masculinos, no qual restringe, dificulta e direciona a participação das mulheres na ciência.

Leta (2014) analisa os dados da produção científica brasileira entre o período de 1997 a 2001 nos campos de Astronomia, Imunologia e Oceanografia. A autora realiza uma ampla revisão de artigos que aborda sobre a mulher na ciência brasileira, incluindo aqueles que focam os aspectos psicológicos, sociológicos, históricos e biológicos. No estudo foram apresentados os cenários: a autoria em publicações científicas de áreas específicas e em toda a produção brasileira; e a atuação em outras tarefas científico-acadêmicas. Por fim, a questão “teriam as mulheres cientistas brasileiras um desempenho inferior ao de seus colegas homens?” foi respondida com evidências sugerindo que a presença das mulheres nas áreas pode ser um fator decisivo para o nível de produção e que a maior parte das mulheres tem desempenho igual ao dos homens. As diferenças entre o desempenho de homens e mulheres estão relacionadas principalmente com a presença relativa deles no campo e com o recorte que é dado às análises.

Em Moschkovich e Almeida (2015), as autoras analisaram de acordo com o sexo, as chances de acesso ao nível mais alto da carreira e aos postos de gestão de docentes no Brasil, focando na Universidade Estadual de Campinas. Os dados sobre os professores foram provenientes da Diretoria Geral de Recursos Humanos da Unicamp, do Sistema de Informações sobre Pesquisa, Ensino e Extensão da Unicamp (Sipex) e da Plataforma Lattes do CNPq. Os resultados mostraram que embora o ingresso por concurso permita uma seleção igualitária de professores, as mulheres estão presentes em menor proporção

nas posições mais altas das universidades. Também apresentaram que ainda há uma concentração das mulheres em algumas áreas da instituição, enquanto que os homens estão espalhados por diferentes áreas. Além disso, as mulheres levam mais tempo para chegar ao topo da carreira. Por fim, os resultados apresentados no estudo apontam a necessidade em se verificar a maneira como mudanças organizacionais e a implementação de políticas devem ser formuladas para contribuir, reforçar ou abolir as desigualdades entre as carreiras de docentes.

Valentova et al. (2017) investigam a sub-representatividade das mulheres nos níveis mais altos da ciência brasileira. Os autores analisam o *gender gap* entre os anos 2013 e 2014, utilizando três fontes de dados: os bolsistas do CNPq, os membros da ABC e o montante de financiamentos concedidos para a pesquisa na chamada universal do CNPq. No estudo foram analisados os diferentes níveis de financiamento para as subáreas de Engenharia e Ciências Exatas e da Terra, de Ciências da Saúde e de Humanidades e Ciências Sociais. Por fim, os resultados evidenciam um desequilíbrio de gênero em todos os indicadores estudados na ciência brasileira. Apontam que existe uma desigualdade na concessão de auxílios, ou seja, as mulheres têm valores de fomento à pesquisa menores do que seus pares homens para as áreas de Engenharia, Ciências Exatas e da Terra e Ciências da Saúde. Contudo para Humanidades e Ciências Sociais Aplicadas, essa diferença não é observada nos detentores de bolsas de produtividade 1A e 1B.

Marini e Meschitti (2018) analisam se o gênero tem alguma discriminação no caminho para a promoção à professor titular nas universidades italianas, utilizando dados de 2013 a 2016. São investigados a desigualdade de gênero usando uma variável binária parametrizada por produtividade científica, número normalizado de vagas disponíveis, resultado de avaliação de pesquisa nacional, idade, classificação anual e posição do arquivo. Como resultados a regressão logística multinível demonstra que, entre aqueles que obtiveram a qualificação ASN (*abilitazione scientifica nazionale*) e na paridade de outras condições, os homens têm cerca de 24% mais probabilidade de serem promovidos na paridade da produção científica, o que revela uma discriminação de gênero relevante.

Em seu trabalho, Abramo, D'Angelo e Costa (2018) analisam os efeitos combinados de gênero, idade e posição acadêmica com o objetivo de entender como traços pessoais e organizacionais podem influenciar comportamentos individuais em termos de diversidade em suas atividades. O acadêmico quando aplica suas competências em áreas diferentes de sua especialização, provoca diversidade em suas atividades de pesquisa. O conjunto de dados para a análise corresponde a produção de pesquisa alcançada por todos os professores italianos nas ciências entre os anos de 2004 a 2008. Aplicou-se o modelo de regressão ao conjunto de dados e a diversidade da pesquisa foi medida através de extensão, intensidade e relacionamento cognitivo dos campos de diversidade. Eles concluíram que não há diferenças de gênero quanto à tendência a enfocar ou diversificar a pesquisa individual.

Contudo, os diferentes comportamentos de diversidade das pesquisas das mulheres em relação aos homens parecem variar com base na dimensão da diversidade analisada e na especificidade de cada disciplina, recomendando assim, cautela na identificação da diversidade da pesquisa como um co-determinante da diferença de produtividade de gênero entre homens e mulheres.

No contexto desta Dissertação que visa analisar as mulheres e sua representatividade no cenário brasileiro, Grossi et al. (2008) mapearam a participação feminina no desenvolvimento de pesquisas no Brasil, utilizando 4.970 currículos da Plataforma Lattes de mulheres que defenderam suas teses de doutorado entre os anos de 2000 a 2013. As análises foram apresentadas nos seguintes tópicos: ano de defesa de tese, realização de pós-doutorado, região geográfica em que a tese foi defendida, áreas e grandes áreas de conhecimento em que as pesquisas foram desenvolvidas, dependências administrativas onde as pesquisas foram desenvolvidas, agências financiadoras das pesquisas, tipo de carreira profissional da pesquisadora. De acordo com o grupo e período pesquisado, os resultados mostraram um crescimento constante no número de mulheres que concluem o doutorado por ano, sendo que, uma parcela de 20% investiram na conclusão ou continuação do pós-doutorado. Apresentaram também que a maior participação feminina ocorre nas áreas das Ciências Biológicas, das Ciências das Saúde e das Ciências Humanas, e que a expressiva maioria das doutoras pesquisadas atuam na docência.

3.2 Estudos envolvendo dados da Plataforma Lattes

A Plataforma Lattes é um dos repositórios de dados curriculares mais confiáveis existentes no mundo (LANE, 2010), como também, pode ser considerada uma fonte inesgotável de informações sobre a ciência brasileira. Os trabalhos apresentados nesta seção envolvem os dados da Plataforma Lattes e os critérios de inclusão utilizados para seleção desses estudos foram: publicações disponíveis on-line a partir do ano 2015 e idiomas inglês e português. A pesquisa dos estudos foi realizada no ano de 2019 e o descritor utilizado foi “Plataforma Lattes”. No Portal de periódicos CAPES/MEC a busca retornou 196 resultados, no Scielo retornou 92 resultados e no Google Acadêmico retornou 12.400 resultados. Como não foi possível ler todos os resultados, a leitura do título foi realizada nas 30 primeiras páginas e foram selecionados os que tinham relação com os dados da Plataforma Lattes, posteriormente realizada a leitura do resumo e excluídos os que não possuíam relação com o tema. Foram selecionados os estudos que utilizaram a Plataforma Lattes ou seus dados curriculares como fonte principal.

Em seu trabalho, Digiampietri (2015) realiza uma caracterização da rede social acadêmica brasileira, utilizando informações acadêmicas disponíveis e validando novas estratégias para a análise de redes sociais. Para o processo de extração, organização,

enriquecimento e análise de dados da rede social brasileira foram utilizados dados dos currículos da Plataforma Lattes, no formato XML e HTML. Para a caracterização ou detalhamento de redes sociais acadêmicas brasileiras utilizou-se diferentes medidas bibliométricas, da análise de redes sociais ou oriundas da mineração de textos. Por fim, o autor apresentou os resultados sobre novas estratégias para resolver problemas específicos relacionados à análise de redes sociais, como resolução de entidades, predição de relacionamentos, análise de tendências e identificação de áreas de atuação.

O trabalho de Dias (2016) difere dos trabalhos encontrados na literatura, pois apresenta um estudo abrangente sobre a produção científica brasileira, tendo como fonte de dados todos os currículos cadastrados na Plataforma Lattes. Para a coleta, o tratamento e a análise dos dados, o autor desenvolveu o arcabouço LattesDataXplorer, implementando técnicas para análises bibliométricas e métricas baseadas em análises de redes sociais, sendo o mesmo responsável pela extração de todo o conjunto de dados dos currículos coletados em abril de 2015. Assim, foi possível caracterizar o conjunto de dados curriculares da Plataforma Lattes, apresentando inicialmente estatísticas gerais, por áreas e por grandes áreas do conhecimento dos dados extraídos. Posteriormente, a análise da produção científica dos doutores foi apresentada como: uma visão geral que permite traçar um perfil das publicações dos doutores brasileiros, um estudo da produção científica dos doutores, uma análise baseada em redes de colaboração científica possibilitando compreender como ocorre a evolução da colaboração científica ao longo dos anos, uma análise das principais áreas de atuação. Por fim, apresentou a análise da produção científica dos docentes de pós-graduação e dos bolsistas de produtividade em pesquisa do CNPq, apresentando dados sobre a caracterização do conjunto, análise quantitativa da produção científica e análise quantitativa dos artigos publicados em periódicos. Os resultados foram apresentados em detalhes e possui caráter inédito sobre a produção cientifica brasileira, tendo em vista a abrangência das análises realizadas e pela grande quantidade de indivíduos considerados.

Brito, Quoniam e Mena-Chalco (2016) utilizam os dados da Plataforma Lattes como fonte de informação, para verificar através de uma expressão de busca e uma metodologia se é possível analisar e compreender áreas específicas do conhecimento, como a Nanotecnologia. A coleta das informações foi realizada utilizando a expressão criada pelos autores e o software ScriptLattes. O software Gephi foi utilizado para visualizar e explorar todos os tipos de redes e de técnicas de análise das mesmas. Os procedimentos descritos no estudo podem ser replicados em diversas áreas do conhecimento e permitem o uso da Plataforma Lattes como fonte de informação temática para análise e compreensão de todo seu conjunto de dados.

Em Santos e Oliveira (2017) as autoras analisam a produção de artigos científicos referente ao tema Língua Brasileira de Sinais (Libras) no período de 1987 a 2014, utilizando os dados curriculares da Plataforma Lattes. Adotou-se técnicas bibliométricas para medir a

produção em Libras, na qual foram identificados 102 autores e 241 artigos de periódicos. Os dados foram apresentados em tabelas e gráficos, obtendo uma visão geral da formação acadêmica dos pesquisadores referente ao tema, bem como a distribuição e cronologia da produção científica de artigos de periódicos sobre Libras. As autoras refletem ao final do artigo que essa temática possui apenas 16 anos de formação no Brasil apresentando um significativo crescimento na publicação de artigos e periódicos e, portanto, obtendo um reflexo positivo para o desenvolvimento da área depois da promulgação da lei de Libras.

Silva et al. (2018) desenvolveram o Sistema Lattes Web Semântico, que executa consulta semântica automática de entidades para os documentos de pesquisadores da Plataforma Lattes utilizando dados abertos conectados (*Linked Open Data*). No artigo foram apresentados conceitos, ferramentas e tecnologias relativas ao tema que possibilitaram a criação do sistema. Foi explicado o processo conceitual do arcabouço, como também um experimento que identificou a efetividade do processo de localização e extração dos termos nos documentos dos currículos. O sistema criado possibilita disponibilizar e indexar os currículos da Plataforma Lattes na Web através do HTML com RDFa (*Resource Description Framework in Attributes*, disponibilizar as informações em RDF (*Resource Description Framework*) Turtle que podem ser utilizados em diversos sistemas e domínios e, por fim, expandir a possibilidade de buscadores semânticos atuarem sobre esses currículos.

A tese de Gomes (2018) apresenta uma análise temporal das principais palavras-chave existentes nos artigos científicos utilizando como dados os currículos cadastrados na Plataforma Lattes. Para alcançar o objetivo de identificar e analisar temporalmente o comportamento dos principais tópicos de pesquisas de interesse dos pesquisadores brasileiros, foram utilizadas as palavras-chave de aproximadamente 14 milhões de artigos científicos referentes a todas as grandes áreas de atuação. Para coletar e tratar os dados contidos nos currículos cadastrados na Plataforma Lattes foi utilizado o arcabouço LattesDataXplorer. Na filtragem e tratamento dos dados foram realizados procedimentos que permitiram extrair e processar, as palavras-chave de cada publicação, e posteriormente analisadas com métodos bibliométricos e técnicas de análises de redes sociais. As análises foram apresentadas como caracterização dos dados e a análise temporal dos principais tópicos de pesquisa da ciência brasileira. Como resultado final, as palavras-chave dos artigos publicados em periódicos entre 1997 e 2016 foram ranqueadas utilizando a medida TF.FI e analisadas em suas respectivas grandes áreas do conhecimento. Os variados resultados apresentados podem servir de base para outros estudos que visam entender o desenvolvimento da ciência brasileira nas diversas áreas do conhecimento.

No trabalho de Damaceno et al. (2019), os autores examinaram os resultados da genealogia acadêmica brasileira dos estudantes de mestrado e doutorado. Os dados foram extraídos entre julho e agosto de 2017 através da Plataforma Lattes, totalizando 737.919 currículos, dos quais 271.370 são doutores e 466.540 mestres. Como resultados,

inicialmente corresponde sob o ponto de vista da genealogia acadêmica que a ciência no Brasil possui um certo grau de interdisciplinaridade, pois os pesquisadores que trabalham em uma das principais áreas de conhecimento orientaram alunos envolvidos em outras áreas do conhecimento. O segundo resultado corresponde as características estruturais e os padrões evolutivos em relação às áreas de conhecimento e pesquisadores, no qual eles argumentam que essas características estão relacionadas com o tempo depois que um acadêmico obteve um diploma ou hereditárias. Por último, os padrões nos níveis de treinamento que afetam as métricas topológicas mostraram curvas semelhantes para todas as métricas, assim, a adição de um mestrado ao gráfico de doutorado, não afeta o comportamento das métricas ou grau de interdisciplinaridade. Por fim, eles concluem que a genealogia acadêmica oferece uma grande oportunidade para avaliar pesquisadores e suas áreas de pesquisa sob a perspectiva do treinamento de recursos humanos.

Um estudo recente, Tuesta et al. (2019) avaliou dados da produção científica de 40.882 doutores com currículos cadastrados na Plataforma Lattes, na área de Ciências Exatas e da Terra. Os autores consideraram as publicações entre os anos 1966 e 2015 e os doutores foram agrupados em cinco conjuntos de dez anos em relação ao início do doutorado. Foram apresentados os cenários dos doutores comparando entre homens e mulheres quanto a: distribuição das publicações, evolução das publicações por ano, média de produtividade anual, tempo médio de conclusão de doutorado, média de autores por artigo. Por fim, foram apresentadas as seguintes conclusões: para as oito subáreas de Ciências Exatas e da Terra a participação dos homens é consideravelmente superior que a de mulheres (acima de 32% de diferença), com exceção de Química (diferença de 2,86%) e Oceanografia (16,16%); ocorre um crescimento na presença das mulheres no cenário de pesquisa ao longo do tempo; a porcentagem de publicações de mulheres é um pouco inferior em relação à porcentagem de sua participação; existe uma maior proporção de homens hiper produtivos, o tempo médio de duração do doutorado tem diminuído para ambos os gêneros, sendo que, na última década a diferença é quase nula; o número de colaboradores para artigos escritos por homens ou mulheres é similar para todas as décadas com raras exceções.

3.3 Diferenciais propostos

O esforço por parte dos pesquisadores em analisar a participação feminina no cenário científico, bem como os estudos que caracterizam os dados da Plataforma Lattes, compreendem inúmeros trabalhos na literatura com abordagens e focos diversos. Porém, verifica-se uma quantidade limitada de trabalhos que utilizam como fonte de dados as informações curriculares da Plataforma Lattes para analisar a produção científica das mulheres, sendo que, tanto o trabalho de Grossi et al. (2008), quanto de Tuesta et al. (2019) utilizam apenas um pequeno contexto destes dados. Consequentemente, o presente

estudo apresenta-se como a primeira análise abrangente para compreender a participação científica das mulheres ao longo dos anos de história registrada nos currículos cadastrados na Plataforma Lattes.

4 Metodologia

Este capítulo descreve a metodologia de pesquisa adotada durante o estudo para atingir os objetivos propostos. Nas próximas seções, são apresentados: o tipo de estudo (Seção 4.1); o processo de aquisição dos dados da Plataforma Lattes que compreende a seleção por formação acadêmica/titulação, seleção por gênero, preparação e visualização dos dados (Seção 4.2); limitações da pesquisa (Seção 4.3).

4.1 Tipo do estudo

Trata-se de pesquisa do tipo descritiva documental, a qual é destinado analisar a participação das doutoras pesquisadoras no aspecto científico e tecnológico, buscando verificar sua evolução acadêmica ao longo dos anos e suas pesquisas através de análises bibliométricas realizadas sobre dados curriculares da Plataforma Lattes.

Quanto aos objetivos, uma pesquisa é definida como descritiva quando relata características de populações ou fenômenos, utilizando técnicas padronizadas para a coleta de dados. Quanto aos procedimentos técnicos, uma pesquisa é definida como documental quando utiliza de materiais que não obtiveram um tratamento analítico em sua totalidade ou que podem receber outras interpretações de acordo com os objetos da pesquisa (GIL, 2008).

4.2 Dados da Plataforma Lattes

A fim de cumprir o objetivo específico “obter, selecionar e preparar os dados da Plataforma Lattes”, passou-se inicialmente pela etapa de aquisição dos dados e seleção dos currículos pelo critério de formação acadêmica/titulação e em seguida, foi realizada a seleção pelo critério de gênero. Posteriormente, foi realizada a preparação dos dados a fim de gerar os arquivos específicos e por fim obter visualizações desses dados.

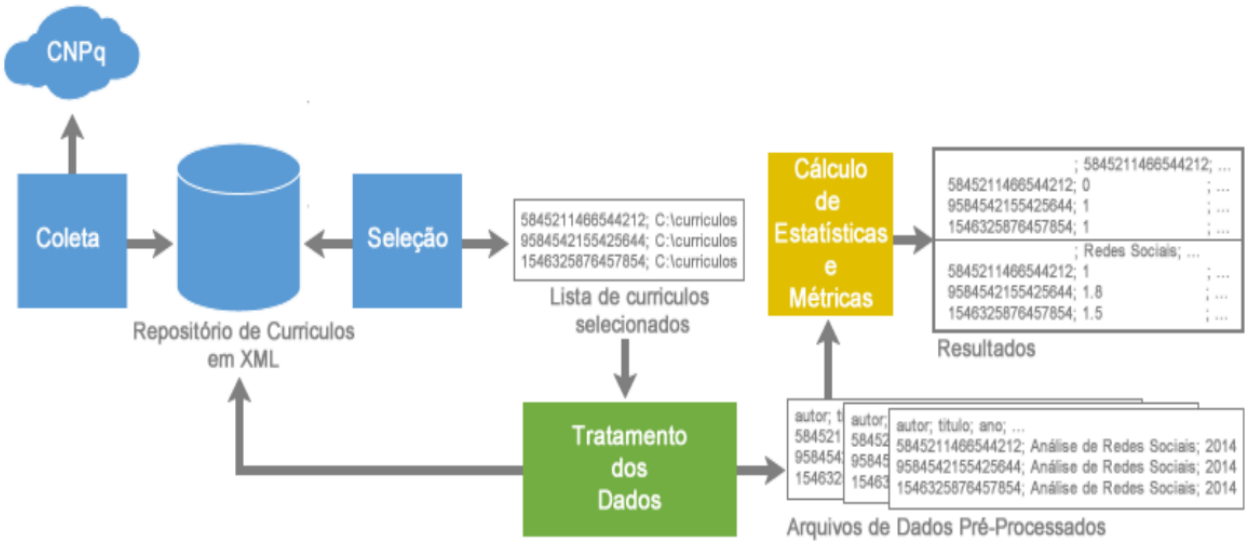
4.2.1 Aquisição dos dados e seleção por formação acadêmica/titulação

A Plataforma Lattes é um repositório de acesso aberto e todas as informações curriculares são incluídas pelo próprio indivíduo e estão disponíveis livremente na Web. Esse amplo conjunto de dados integra em um único sistema as bases de dados de Currículos, Grupos de pesquisa e de Instituições e contém todo o registro da trajetória profissional, acadêmica e produções científicas do pesquisador. Pela sua riqueza de informações e como esta fonte de dados não foi amplamente analisada, justifica-se assim a escolha da

Plataforma Lattes como fonte de dados para medir e avaliar o desempenho científico das mulheres.

O processo de extração e seleção dos dados curriculares da Plataforma Lattes foi realizado através do arcabouço LattesDataXplorer, proposto por Dias (2016). Este arcabouço contém um conjunto de técnicas e métodos responsáveis por coletar, selecionar, tratar e analisar os dados (Figura 1).

Figura 1 – Visão geral do LattesDataXplorer



Fonte: Dias (2016)

Os módulos utilizados do LattesDataXplorer para extrair e selecionar os dados curriculares da Plataforma Lattes foram coleta e seleção. Assim, o módulo coleta de dados foi realizada nas seguintes etapas:

1. Extração de URLs: responsável por obter os códigos únicos de todos os currículos cadastrados, possibilitando o acesso individual a cada currículo. Como cada currículo cadastrado na Plataforma Lattes possui um código único de oito dígitos, os mesmos são obtidos através de uma requisição na interface de consulta, gerando uma lista que contém todos os códigos de todos os currículos cadastrados.
2. Extração de IDs e Data: responsável por acessar o cabeçalho de cada currículo para extrair o seu identificador único e a data de última atualização. Além do código único, cada currículo cadastrado na Plataforma Lattes possui também um identificador único de dezesseis dígitos. Estes dados extraídos são armazenados localmente em um arquivo de identificação que contém código, identificador, data da última atualização no CNPq e data da atualização do currículo.
3. Extração dos currículos: responsável por extrair e armazenar os currículos. Caso a data de atualização na Plataforma Lattes seja divergente da data de atualização do currículo armazenado localmente, o currículo local é substituído pelo mais recente.

A coleta de todos os currículos utilizando o extrator LattesDataXplorer foi realizada em outubro de 2019, ultrapassando 6.300.000 registros. Os currículos coletados possuem o formato XML (eXtensible Markup Language - uma linguagem de marcação que contém seções e campos bem delimitados), sendo este formato de arquivo mais adequado para o processamento automático dos dados. O extrator coleta os XMLs dos currículos e os armazena em pastas, identificadas de 00 a 99. A pasta selecionada para armazenar o arquivo e o nome do mesmo são definidas de acordo com seu identificador único de 16 dígitos. Sendo que, os dois primeiros números do identificador correspondem ao nome da pasta e os 14 dígitos restantes correspondem ao nome do arquivo XML salvo. Como exemplo, o currículo com identificador 3530976051984613 seria armazenado na pasta 35 e o nome arquivo XML seria 30976051984613. Assim, todos os currículos são armazenados localmente em formato XML para fácil manipulação dos arquivos.

Para realizar uma análise detalhada da participação científica das mulheres, optou-se por limitar os dados através do nível de formação acadêmica/titulação, reduzindo o conjunto para indivíduos que possuem o nível de formação doutorado concluído. Apesar deste conjunto não ser o mais significativo entre os níveis de formação, conforme enunciado por Dias (2016), eles são responsáveis por 74,51% dos artigos publicados em periódicos e 64,67% dos artigos publicados em anais de congresso, além de possuir em geral data de atualização de seus currículos recente e notadamente são responsáveis pelo mais alto nível de formação, a saber, mestrado e doutorado.

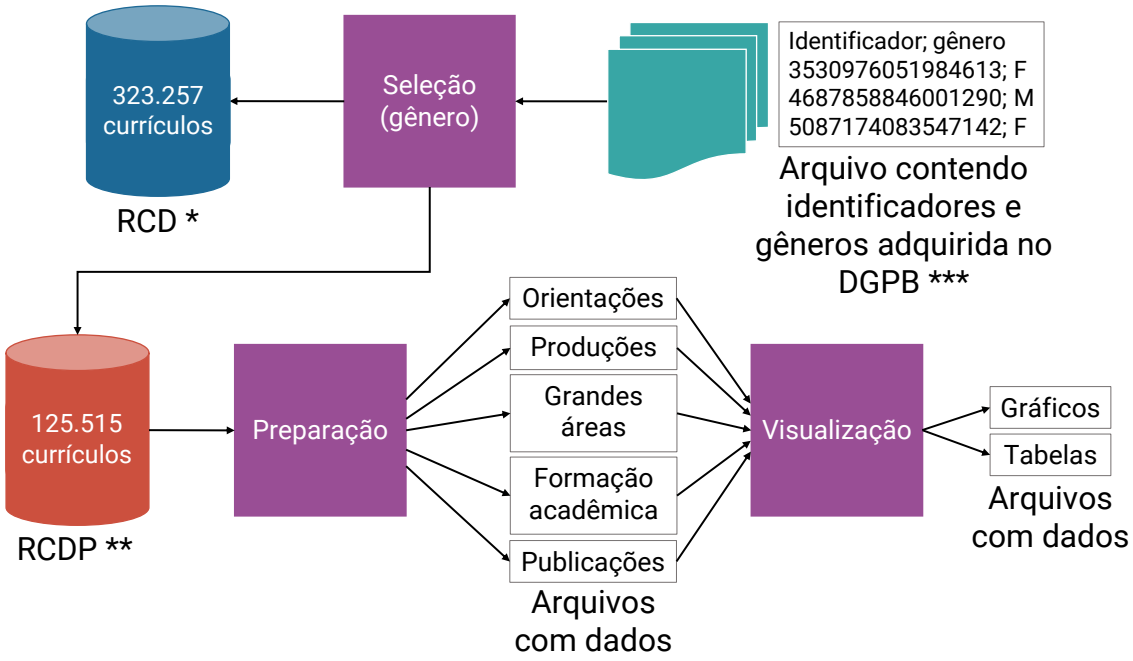
Logo, após a aquisição de todos os currículos cadastrados na Plataforma Lattes, foi utilizado o módulo de seleção do LattesDataXplorer para selecionar, dentre estes, os currículos que possuem a formação acadêmica/titulação doutorado concluído, totalizando assim um conjunto com 323.257 currículos armazenados no Repositório de Currículos de todos Doutores (RCD).

4.2.2 Seleção pelo critério de gênero, preparação e visualização dos dados

O LattesDataXplorer foi utilizado especificamente para a coleta e seleção dos dados curriculares da Plataforma Lattes, no qual obteve-se o RCD. Após esse processo, foi necessário selecionar os dados desse repositório pelo critério de gênero e assim obter um repositório de currículos das mulheres que possuem formação acadêmica/titulação doutorado concluído. Por fim, realizou-se a preparação e visualização dos dados que permitiram analisar a participação científica das doutoras pesquisadoras (Figura 2).

Identificar o sexo dos autores é um grande desafio quando abordamos a produção científica e os estudos métricos (LETA, 2014)), e como Plataforma Lattes não disponibiliza o campo sexo para a consulta pública, para realizar a etapa que corresponde a seleção pelo critério de gênero (“Seleção”), foi necessário buscar essa informação em outra fonte de dados.

Figura 2 – Processo de seleção pelo critério de gênero, preparação e visualização dos dados



Fonte: Elaboração do autor
Nota: *Repositório de Currículos de todos Doutores,
**Repositório de Currículos das Doutoras Pesquisadoras,
***Diretório dos Grupos de Pesquisa no Brasil

Desta maneira, a segunda fonte de dados utilizada para obter o gênero corresponde ao Diretório dos Grupos de Pesquisa no Brasil[1] (DGPB), que possui o inventário dos grupos de pesquisa científica e tecnológica em atividade no país, disponibilizando os dados dos Censos em formato aberto. As informações contidas no DGPB são relativas aos recursos humanos constituintes dos pesquisadores, estudantes e técnicos, bem como, linhas de pesquisa que estão em andamento, especialidades do conhecimento, setores de aplicação abrangidos, produção científica, tecnológica e artística, entre outros. Os grupos de pesquisa estão localizados em universidades, instituições de ensino superior que ofertam cursos de pós-graduação *stricto sensu*, institutos de pesquisas científicas e tecnológicas. O diretório possui dois tipos de bases: Corrente e Censitárias. As bases Correntes podem modificar diariamente pois os grupos podem ser registrados, excluídos e atualizados. As bases Censitárias são disponibilizadas em um período quase sempre bienal, sendo esta uma cópia da base Corrente, e tem como objetivo realizar um censo da capacidade de pesquisa em atividade no país (DIRETÓRIO DE GRUPOS DE PESQUISA NO BRASIL, 2020).

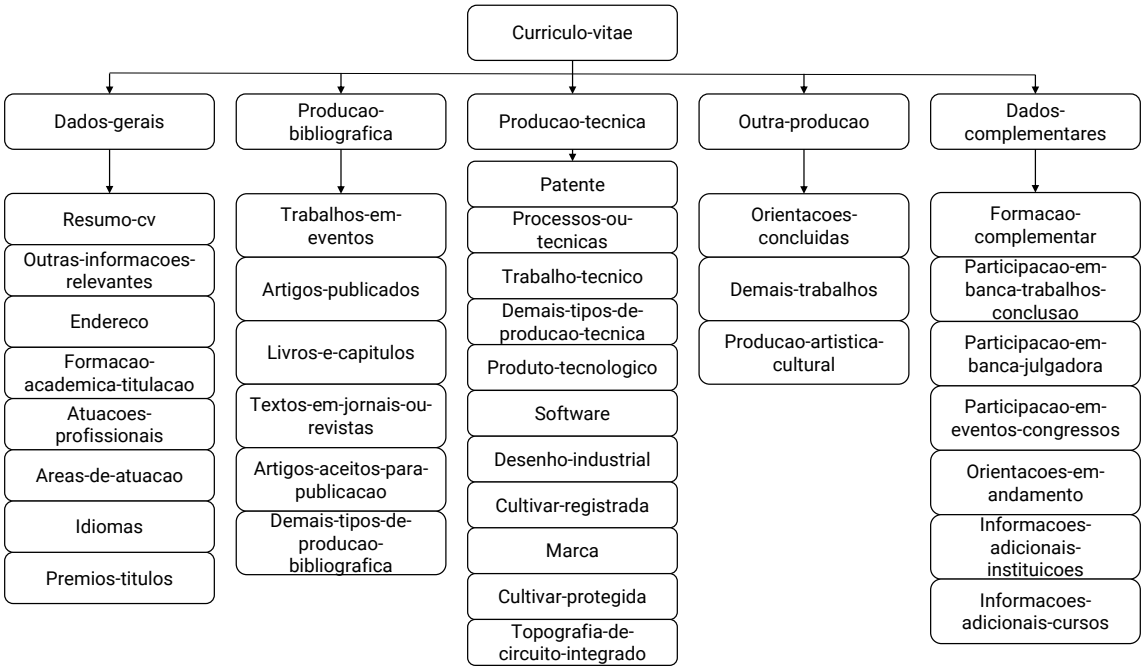
Os dados recuperados para gerar a lista de identificadores constitui-se dos arquivos Censitários de 2006, 2008, 2010, 2014 e 2016 contendo informações de pesquisadores,

[1] DGPB: http://lattes.cnpq.br/web/dgp/censos2

estudantes, técnicos e colaboradores estrangeiros que na época da disponibilização de cada Censo, os mesmos participavam de um grupo de pesquisa. Os arquivos dos Censos de 2000, 2002 e 2004 não continham o atributo sexo e por este motivo não foram utilizados para a pesquisa. Os arquivos Censitários utilizados no estudo corresponderam a 144.408 arquivos XML e depois de processados, foram identificados 888.665 registros que continham o atributo sexo preenchido. Porém dentre esses, 97 registros continham duplicidade no gênero, ou seja, em um determinado censo estava “F” e no outro estava “M”. Por este motivo, a lista de identificadores finalizou com 888.568 registros, no qual 383.595 correspondem ao gênero masculino e 504.973 ao gênero feminino. Após obter a lista de identificadores do gênero feminino, foi possível utilizá-la para buscar no RCD os registros que possuem identificadores iguais aos da lista, definindo assim o conjunto principal de arquivos XML com total de 125.515 currículos, denominado Repositório de Currículos das Doutoras Pesquisadoras (RCDP).

Para realizar a etapa de “Preparação”, inicialmente foi identificada a estrutura dos arquivos do RCDP. Um arquivo XML é organizado através dos elementos e atributos, no qual inicia pelo elemento raiz (pai) e percorre os elementos filhos até alcançar os atributos. A estrutura dos currículos XML da Plataforma Lattes apresenta vários níveis de elementos, porém para melhor visualização foi apresentado apenas os níveis iniciais (Figura 3). Assim, os arquivos XML do RCDP possuem o elemento raiz, nomeado por “Curriculo-vitae”, e seus cinco elementos filhos que possuem seus próprios elementos e atributos.

Figura 3 – Fragmento da estrutura dos currículos XML do RCDP



Fonte: Elaboração do autor

Os arquivos XML extraídos da Plataforma Lattes apresentam informações estruturadas e diversidade dos dados como, nível de formação acadêmica/titulação, grandes áreas de atuação, projetos de pesquisa e extensão, idiomas, produções bibliográficas e técnicas como artigos publicados em anais de congresso e artigos em periódicos, apresentação de trabalhos, participação em bancas, eventos, orientações, dentre outros. Assim, após identificar a estrutura dos arquivos do RCDP, a segunda etapa do módulo de "Preparação" foi realizada acessando cada currículo XML, obtendo a informação específica daquele currículo e armazenando em coleções de arquivos estruturados.

Como cada currículo possui uma quantidade específica de informações, esses dados foram agrupados quanto a orientações, produções, grandes áreas, formação acadêmica, publicações. Após agrupar os dados, foi realizado uma caracterização para facilitar as análises dos dados, denominado módulo de "Visualização". Nesta etapa, para cada coleção de arquivos estruturados foram gerados gráficos e tabelas para facilitar as análises.

4.2.3 Perfil de orientações das doutoras

A preparação do arquivo de dados a fim de cumprir o objetivo "traçar o perfil de orientações das doutoras" foi realizada recuperando de todos os currículos do RCDP as informações relativas a orientações concluídas e orientações em andamento para as categorias Pós-doutorado, Doutorado, Mestrado, Especialização, Graduação, Iniciação científica e Outras.

Assim, como critério de inclusão para as orientações concluídas, o elemento do arquivo XML utilizado corresponde a "outra-producao"[2]. Dentro deste, continha o elemento "orientacoes-concluidas", no qual buscou-se e realizou-se a contagem de todos os registros para os seguintes elementos: "orientacoes-concluidas-para-pos-doutorado", "orientacoes-concluidas-para-doutorado", "orientacoes-concluidas-para-mestrado" e "outras-orientacoes-concluidas". Contudo, como o elemento "outras-orientacoes-concluidas" engloba as orientações relativas à especialização, graduação, iniciação científica e outros, dentro dele buscou-se o elemento "dados-basicos-de-outras-orientacoes-concluidas", no qual realizou-se a contagem de todos os registros para os seguintes atributos: "monografia-de-conclusao-de-curso-aperfeicoamento-e-especializacao", "trabalho-de-conclusao-de-curso-graduacao", "iniciacao-cientifica" e "orientacao-de-outra-natureza".

Igualmente, o elemento do arquivo XML utilizado como critério de inclusão para as orientações em andamento corresponde a "dados-complementares". Dentro deste, continha o elemento "orientacoes-em-andamento", no qual buscou-se e realizou-se a contagem de todos os registros para os seguintes elementos: "orientacao-em-andamento-de-

[2] Todos os nomes dos elementos do arquivo XML descritos na Figura 3 e a partir da subseção 4.2.3 não apresentam a ortografia correta. Essas informações foram suprimidas intencionalmente com o objetivo de apresentar os nomes reais dos elementos encontrados no XML, no qual correspondem aos mesmos utilizados no desenvolvimento dos códigos para obtenção dos resultados.

pos-doutorado", "orientacao-em-andamento-de-doutorado", "orientacao-em-andamento-de-aperfeicoamento-especializacao", "orientacao-em-andamento-de-graduacao", "orientacao-em-andamento-de-iniciacao-cientifica" e 'outras-orientacoes-em-andamento".

Para este tópico não foram aplicados critérios de exclusão, em razão da contagem realizada por cada categoria.

Além da contagem, para cada currículo do RCDP buscou-se os seguintes atributos para as orientações concluídas: dentro do elemento "dados-basicos-de-outras-orientacoes-concluidas", atributos "natureza" e "ano"; no elemento "detalhamento-de-outras-orientacoes-concluidas", atributos "flag-bolsa", "numero-id-orientado", "nome-do-orientado" e "tipo-de-orientacao"; o elemento "areas-do-conhecimento", buscou-se o elemento "area-do- conhecimento" e por fim o atributo "nome-grande-area-do-conhecimento". Já para as orientações em andamento buscou-se os seguintes atributos: dentro do elemento "dados-basicos-da-orientacao-em-andamento-de-pos-doutorado", atributos "natureza" e "ano"; no elemento "detalhamento-da-orientacao-em-andamento-de-pos-doutorado": "flag-bolsa", "numero-id-orientando", "nome-do-orientando" e "tipo-de-orientacao"; o elemento "areas-do- conhecimento", buscou-se o elemento "area-do-conhecimento-1" e por fim o atributo "nome-grande-area-do-conhecimento". Para os dois tipos de orientações somente a primeira grande área de conhecimento foi recuperada, considerando ser a primeira preenchida pela doutora a mais relevante dentre as outras. Para este tópico os critérios de exclusão foram: os registros que não continham dados de grande áreas e ano preenchidos foram categorizados como "Não informados" e os mesmos não foram utilizados para as análises.

Enfim, a conclusão da etapa de pré-processamento para as orientações das doutoras pesquisadoras obteve-se arquivos estruturados contendo o total e os atributos para cada registro de cada uma das categorias Pós-doutorado, Doutorado, Mestrado, Especialização, Graduação, Iniciação científica e Outras, tanto para orientações concluídas, quanto para as orientações em andamento.

4.2.4 Representatividade da produção bibliográfica e técnica das doutoras

A preparação do arquivo de dados a fim de cumprir o objetivo "analisar a representatividade da produção bibliográfica e técnica das doutoras" foi realizada recuperando de todos os currículos do RCDP as informações referentes as produções bibliográficas e técnicas.

Assim, como critério de inclusão para as produções bibliográficas e técnicas, dentro do elemento "producao-bibliografica" do arquivo XML, buscou-se os elementos "trabalhos-em-eventos", "artigos-publicados", "livros-e-capitulos", "textos-em-jornais-ou-revistas" e "demais-tipos-de-producao-bibliografica". Para o elemento "trabalhos-em-eventos", buscou-se e realizou-se a contagem de todos os registros contendo o elemento "trabalho-em-

eventos". Para o elemento "artigos-publicados", buscou-se e realizou-se a contagem de todos os registros contendo o elemento "artigo-publicado". Para o elemento "livros-e-capitulos", buscou-se os dois elementos "livros-publicados-ou-organizados" e "capitulos-de-livros-publicados", sendo que no primeiro buscou-se e realizou-se a contagem de todos os registros contendo o elemento "livro-publicado-ou-organizado" e para o segundo buscou-se e realizou-se a contagem de todos os registros contendo o elemento "capitulo-de-livro-publicado". Para o elemento "textos-em-jornais-ou-revistas", buscou-se e realizou-se a contagem de todos os registros contendo o elemento "texto-em-jornal-ou-revista". Para o elemento "demais-tipos-de-producao-bibliografica", buscou-se e realizou-se a contagem de todos os registros contendo os elementos "outra-producao-bibliografica", "prefacio-posfacio", "traducao" e "partitura-musical".

Igualmente, dentro do elemento "producao-tecnica" do arquivo XML, buscou-se os elementos "trabalho-tecnico" e "demais-tipos-de-producao-tecnica". Para o elemento "trabalho-tecnico", buscou-se e realizou-se a contagem de todos os registros contendo os elementos: "parecer", "elaboracao-de-projeto", "outra", "relatorio-tecnico", "servicos-na-area-da-saude", "extensao-tecnologica", "nao-informado". Para o elemento "demais-tipos-de-producao-tecnica", buscou-se e realizou-se a contagem de todos os registros contendo os elementos: "curso-de-curta-duracao-ministrado", "desenvolvimento-de-material-didatico-ou-instrucional", "outra-producao-tecnica", "editoração", "relatorio-de-pesquisa", "manutencao-de-obra-artistica" e "maquete".

Foi realizado a contagem geral e por ano para cada uma das produções bibliográficas e técnicas. Sendo que, na contagem por ano, o procedimento foi idêntico ao descrito acima, adicionando na busca mais um elemento correspondente a "dados-basicos" e recuperando o atributo "ano" em cada uma das produções. Também foram recuperados a data e hora de atualização para todos os currículos, buscando o elemento "curriculo-vitae", os atributos "hora-atualizacao" e "data-atualizacao".

Para este tópico não foram aplicados critérios de exclusão, em razão da contagem realizada por cada produção bibliográfica e técnica.

Enfim, a conclusão da etapa de pré-processamento para a representatividade da produção bibliográfica e técnica das doutoras pesquisadoras obteve-se dois arquivos estruturados. No primeiro obteve-se o total para cada produção bibliográfica e técnica para cada registro do RCDP, sendo eles: artigos publicados em periódicos, artigos publicados em anais de congresso, capítulos de livros, demais produções técnicas, livros, outras produções bibliográficas, textos em jornais e revistas, trabalhos técnicos. No segundo obteve-se o total por ano para as mesmas produções bibliográficas e técnicas descritas acima.

4.2.5 Distribuição das doutoras em suas grandes áreas de atuação

A preparação do arquivo de dados a fim de cumprir o objetivo “fazer uma distribuição das doutoras em suas grandes áreas de atuação”, foi realizada recuperando de todos os currículos do RCDP as informações referentes as áreas de atuação profissional.

Assim, como critério de inclusão para as grandes áreas de atuação, o elemento do arquivo XML utilizado corresponde a “dados-gerais”. Dentro deste, continha o elemento “areas-de-atuacao”, no qual buscou-se o elemento: “area-de-atuacao” e por fim, o atributo “nome-grande-area-do-conhecimento”. Para realizar essa análise foi selecionada a maior ocorrência dentre as grandes áreas recuperadas de cada currículo, considerando esta a mais importante dentre elas.

Igualmente, buscou-se também as publicações bibliográficas dos artigos em periódicos e em anais de congresso por grandes áreas. Assim, como critério de inclusão para as produções bibliográficas, dentro do elemento “producao-bibliografica” do arquivo XML, buscou-se os elementos “trabalhos-em-eventos” e “artigos-publicados”. Para o elemento “trabalhos-em-eventos”, buscou-se e realizou-se a contagem de todos os registros contendo o elemento “trabalho-em-eventos”. Para o elemento “artigos-publicados”, buscou-se e realizou-se a contagem de todos os registros contendo o elemento “artigo-publicado”.

Como critérios de exclusão, os registros que não continham dados de grandes áreas preenchidos foram categorizados como “Não informados” e os mesmos não foram utilizados para as análises.

Enfim, a conclusão da etapa de pré-processamento para a distribuição das doutoras pesquisadoras em suas grandes áreas de atuação obteve-se um arquivo estruturado contendo como colunas: grande área, total de artigos em periódicos, total de artigos em anais de congresso para todos os registros do RCDP.

4.2.6 Evolução da formação acadêmica das doutoras

A preparação do arquivo de dados a fim de cumprir o objetivo “compreender a evolução da formação acadêmica das doutoras” foi realizada recuperando de todos os currículos do RCDP as informações referentes a formação acadêmica para os níveis: graduação, mestrado e doutorado.

Assim, como critério de inclusão, dentro do elemento “dados-gerais” do arquivo XML, buscou-se no elemento “formacao-academica-titulacao” os seguintes elementos “graduacao”, “mestrado” e “doutorado”. Para o elemento “graduacao” buscou-se os seguintes atributos: “nome-curso”, “ano-de-inicio”, “ano-de-conclusao” e “nome-instituicao”. Para os elementos “mestrado” e “doutorado”, buscou-se os atributos: “nome-curso”, “ano-de-inicio”, “ano-de-conclusao” e “nome-instituicao”, como também o elemento “areas-do-conhecimento”, que contém o elemento “area-do-conhecimento-1” com os seguintes atributos: “nome-da-

especialidade", "nome-da-sub-area-do-conhecimento", "nome-da-area-do-conhecimento" e "nome-grande-area-do-conhecimento". Somente a primeira grande área de conhecimento foi recuperada, considerando ser a primeira preenchida pela doutora a mais relevante dentre as outras.

Como critérios de exclusão, os registros que não continham dados preenchidos quanto ao ano inicial, ano final ou grande área do conhecimento foram categorizados como "Não informados" e os mesmos não foram utilizados para as análises.

Enfim, a conclusão da etapa de pré-processamento para a evolução da formação acadêmica das doutoras pesquisadoras obteve-se um arquivo estruturado contendo como colunas os registros para cada nível de formação graduação, mestrado e doutorado: nome do curso, ano de início, ano de conclusão, nome da instituição. Como também as colunas referentes a área de conhecimento para mestrado e doutorado: nome da especialidade, nome da subárea do conhecimento, nome da área do conhecimento e nome grande área do conhecimento.

4.2.7 Publicações das doutoras após a conclusão do doutorado

A etapa de preparação do arquivo de dados a fim de cumprir o objetivo "investigar as publicações das doutoras após a conclusão do doutorado" foi realizada recuperando de todos os currículos do RCDP as informações referentes as produções bibliográficas: artigos completos publicados em periódicos, trabalhos publicados em anais de eventos, livros e capítulos de livros.

O primeiro critério de inclusão utilizado corresponde ao ano de conclusão do doutorado, sendo esse campo a base para a contagem dos registros de publicações. Assim, dentro do elemento "dados-gerais" do arquivo XML, buscou-se no elemento "formacao-academica-titulacao". Dentro deste, buscou-se o elemento "doutorado" que continha o atributo "ano-de-conclusao".

O atributo "ano-de-conclusao" sendo um dado numérico maior que zero correspondia a um critério de inclusão para a contagem da produção bibliográfica por ano. Assim, como outros critérios de inclusão, dentro do elemento "producao-bibliografica" do arquivo XML, buscou-se os elementos "trabalhos-em-eventos", "artigos-publicados" e "livros-e-capitulos". Para o elemento "trabalhos-em-eventos", buscou-se o elemento "trabalho-em-eventos" e dentro dele o elemento "dados-basicos-do-trabalho", que continha o atributo "ano-do-trabalho". Para o elemento "artigos-publicados", buscou-se o elemento "artigo-publicado" e dentro dele o elemento "dados-basicos-do-artigo", que continha o atributo "ano-do-artigo". Para o elemento "livros-e-capitulos", buscou-se os elementos "livros-publicados-ou-organizados" e "capitulos-de-livros-publicados". Dentro de "livros-publicados-ou-organizados", buscou-se o elemento "livro-publicado-ou-organizado" e dentro dele o elemento "dados-basicos-do-livro",

que continha o atributo “ano”. Dentro de “capitulos-de-livros-publicados”, buscou-se o elemento “capitulo-de-livro-publicado” e dentro dele o elemento “dados-basicos-do-capitulo”, que continha o atributo “ano”.

Como critérios de exclusão, os registros que não continham ano de conclusão do doutorado ou essa informação não correspondiam a um dado numérico maior que zero não foram utilizados para as análises.

Enfim, a conclusão da etapa de pré-processamento para as publicações das doutoras pesquisadoras após a conclusão do doutorado obteve-se um arquivo estruturado para cada produção bibliográfica, no qual contém o somatório de publicações por ano, a partir do primeiro ano de conclusão do doutorado para todos os registros do RCDP para artigos completos publicados em periódicos, trabalhos publicados em anais de eventos, livros e capítulos de livros.

4.2.8 Colaboração científica das doutoras

A etapa de preparação do arquivo de dados a fim de cumprir o objetivo “analisar a colaboração científica das doutoras” foi realizada recuperando de todos os currículos do RCDP as informações referentes aos autores das produções bibliográficas e técnicas.

O critério de inclusão utilizado corresponde ao autor da produção. Assim, dentro do elemento “dados-gerais” do arquivo XML, buscou-se todos os atributos particulares de cada doutora nos elementos “formacao-academica-titulacao”, “areas-de-atuacao” e “endereco”. Já nas publicações científicas e técnicas, buscou-se nos elementos “producao-bibliografica” e “producao-tecnica” o atributo relacionado aos autores. Foi recuperado uma lista de todos os autores para cada produção de cada doutora contendo nome e/ou identificador e posteriormente gerou-se uma lista de conexão entre os autores.

Enfim, a conclusão da etapa de pré-processamento para a colaboração científica das doutoras pesquisadoras, obteve-se dois arquivos estruturados. O primeiro contém informações de cada doutora, como identificador, nome, estado, país, titulação, grande área, área, instituição e cidade. Já o segundo arquivo, contém somente os identificadores, sendo que, cada linha do arquivo refere-se a uma conexão entre a doutora e seus colaboradores para cada produção.

4.2.9 Produção bibliográfica das bolsistas de produtividade em pesquisa do CNPq

A preparação do arquivo de dados a fim de cumprir o objetivo “verificar a produção bibliográfica das bolsistas de produtividade em pesquisa do CNPq” foi realizada recuperando de todos os currículos do RCDP as informações referentes as produções bibliográficas e técnicas para as doutoras que são bolsistas de produtividade em pesquisa do CNPq.

A informação relativa aos bolsistas de produtividade em pesquisa do CNPq não aparece nos arquivos XML recuperados. Sendo assim, o arcabouço LattesDataXplorer recuperou essa informação através do cabeçalho de cada currículo cadastrado na Plataforma Lattes e gerou arquivos contendo os identificadores de todos os bolsistas de produtividade em pesquisa do CNPq para os níveis: 1A, 1B, 1C, 1D e 2. De posse dessas listas, foi realizada a filtragem dos arquivos para recuperar as bolsistas que são doutoras do RCDP, sendo este o critério de inclusão. Então, recuperando no RCDP os registros que possuem identificadores iguais aos dos arquivos, gerou-se novos arquivos que possuem apenas os identificadores das bolsistas de produtividade em pesquisa do CNPq e que são doutoras pesquisadoras para os níveis: 1A, 1B, 1C, 1D, 2 e geral. Através dos registros destas listas, foi realizado o mesmo procedimento descrito nas subseções 4.2.4 e 4.2.8.

Enfim, a conclusão da etapa de pré-processamento para as produções bibliográficas das doutoras pesquisadoras que são bolsistas de produtividade em pesquisa do CNPq, obteve-se vários arquivos estruturados para cada nível: 1A, 1B, 1C, 1D, 2 e geral. O arquivo utilizado para analisar a produção contém o somatório referente a cada produção bibliográfica. Para a colaboração foram gerados dois arquivos, o primeiro contendo informações de cada doutora, como identificador, nome, estado, país, titulação, grande área, área, instituição e cidade. E o segundo arquivo, contém somente os identificadores, sendo que cada linha do arquivo corresponde a conexão das doutoras para todos os registros do RCDP.

4.3 Limitações da pesquisa

No que se refere a coleta dos dados, deve-se considerar a confiabilidade dos dados, porém dois fatores impossibilitaram utilizar todo o conjunto de currículo das doutoras para análise. O primeiro corresponde a Plataforma Lattes que não disponibiliza as informações quanto ao gênero e o segundo corresponde ao DGPB, pois não são todas as doutoras que estão inseridas em um grupo de pesquisa. Vale ressaltar que a amostra utilizada neste estudo corresponde a todas as doutoras que estão inseridas ou participaram em algum momento de um grupo de pesquisa disponibilizado nos arquivos Censitários do DGPB.

5 Análise e Discussão dos Resultados

Os dados coletados da Plataforma Lattes, utilizando o arcabouço LattesDataXplorer em outubro de 2019, totalizaram mais de 6.300.000 registros. Desse total, foram selecionados os registros com o nível de formação acadêmica/titulação doutorado concluído, totalizando em 323.257 (5,13%) currículos de todos os doutores das diversas áreas do conhecimento científico. Esses mesmos dados foram selecionados pelo critério de gênero (utilizando o DGPB como validação), totalizando 125.515 (1,99%) currículos das doutoras pesquisadoras, sendo que, esses dados serão utilizados para as análises desse capítulo que retratam a participação científica feminina. Nas próximas seções, são apresentados: o perfil de orientações das doutoras (Seção 5.1); estudo sobre a representatividade da produção bibliográfica e técnica das doutoras (Seção 5.2); distribuição das doutoras em suas grandes áreas de atuação (Seção 5.3); evolução da formação acadêmica das doutoras (Seção 5.4); as publicações das doutoras após a conclusão do doutorado (Seção 5.5); a colaboração científica das doutoras (Seção 5.6); a produção bibliográfica das bolsistas de produtividade em pesquisa do CNPq (Seção 5.7).

5.1 Perfil de orientações das doutoras

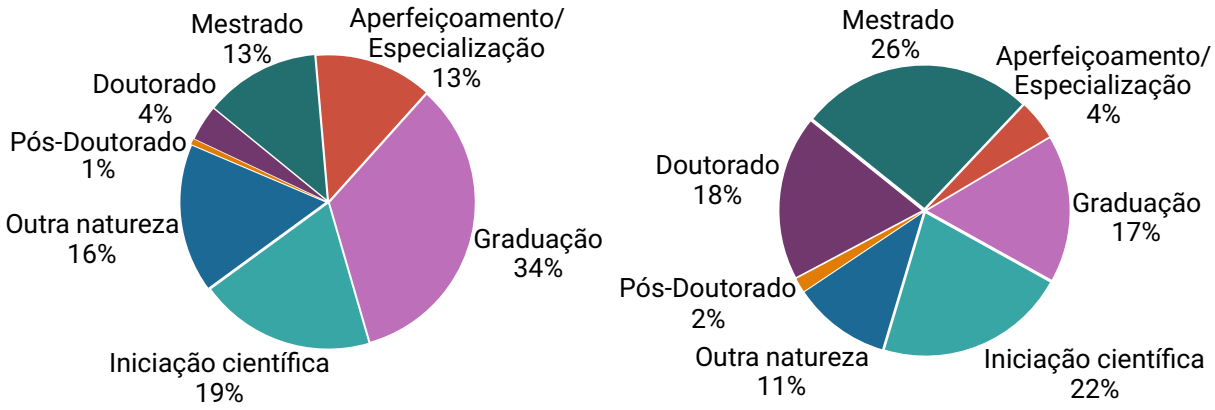
Uma análise relevante corresponde às orientações, pois os doutores, no decorrer de sua trajetória acadêmica, são responsáveis pela orientação dos alunos nos seus diferentes níveis de formação:

> A orientação ocorre em diversos campos científicos, como nos cursos técnicos de nível médio, cursos de graduação, iniciação científica, especialização, residência multiprofissional e uniprofissional em saúde, mestrado e doutorado. O pós-doutorado não corresponde a um grau acadêmico, e não apresenta a figura do orientador, tendo apenas um supervisor igualmente com título de doutor e com notável experiência acadêmico-científica, que acompanhará o andamento da pesquisa do respectivo pós-doutorando. (LOPES et al., 2020, p. 3858)

Na Plataforma Lattes as orientações podem ser cadastradas através do menu “Orientações”, sendo possível escolher entre as opções: “Orientações e supervisões concluídas” e “Orientações e supervisões em andamento”. As orientações concluídas correspondem a todas as orientações realizadas pelas doutoras desde o início de sua carreira e que já estão finalizadas, enquanto que as orientações em andamento são aquelas que estão em desenvolvimento e ainda não foram finalizadas. Ao inserir uma orientação, tanto para concluídas, quanto para em andamento, é possível escolher dentre uma das opções: Dissertação de mestrado, Tese de doutorado, Monografia de conclusão de curso de

aperfeiçoamento/especialização, Trabalho de conclusão de curso de graduação, Iniciação Científica, Supervisão de pós-doutorado, Orientação de outra natureza. Logo, quantificando os dados do RCDP temos um total de 3.455.229 orientações concluídas (Figura 4a) e 334.570 orientações em andamento (Figura 4b) distribuídas pelas sete opções de formação acadêmica/titulação.

Figura 4 – Orientações



(a) Concluídas (b) Em andamento

Fonte: Elaboração do autor

Um aspecto relevante na Figura 4 refere-se as formações acadêmicas/titulações: mestrado, doutorado e pós-doutorado. Nestas três formações, a soma das porcentagens para o gráfico de orientações concluídas corresponde a 18%, enquanto que para o gráfico das orientações em andamento essa soma corresponde a 46%. Uma das hipóteses para essa diferença de percentual diz respeito ao fato de que, como os doutores são responsáveis pela formação dos alunos nos principais programas de pós-graduação *stricto sensu* no Brasil, eles tendem a orientar mais alunos de graduação e menos alunos de pós-graduação no início de sua carreira e com o passar dos anos após a conclusão do doutorado o número de orientações da graduação diminui e as orientações da pós-graduação aumentam. No entanto, como são consideradas todas as orientações em todo o seu histórico, as orientações concluídas de nível mais baixo, como a graduação, possuem maior representatividade no conjunto. Vale ressaltar sobre a duração das orientações de cada uma das formações acadêmicas/titulações, que podem contextualizar sobre os percentuais obtidos, por exemplo, o mestrado dura de 18 a 24 meses, e o doutorado de 24 a 48 meses.

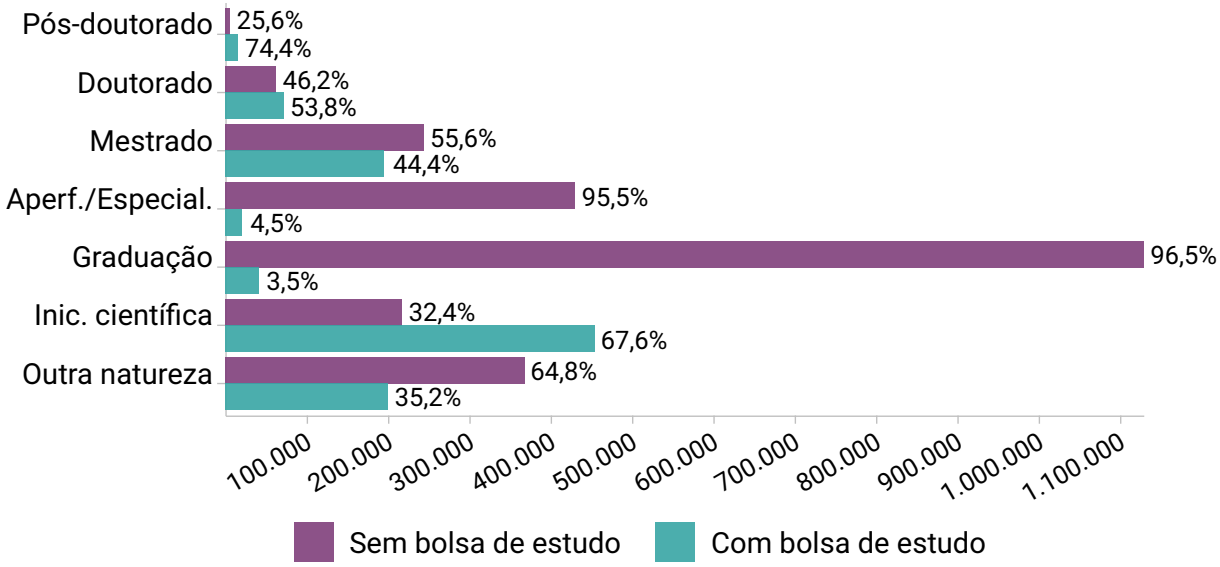
Para as orientações concluídas, no Pós-doutorado do RCDP, apenas 7.167 doutoras pesquisadoras realizaram a supervisão de 21.235 pós-doutores e as duas primeiras conclusões de Pós-doutorado ocorreram em 1980, sendo supervisionadas por uma única doutora. Quanto a pós-graduação Doutorado, 22.342 doutoras orientaram 134.359 doutores e as duas primeiras conclusões ocorreram em 1969, sendo orientadas por uma única doutora. Já para a formação Mestrado, 45.593 doutoras orientaram 439.864, e a primeira orientação

de mestrado foi realizada em 1966. Para a Especialização, 45.609 doutoras orientaram 450.036 especialistas e as duas primeiras orientações de especialização foram realizadas em 1966. Para a Graduação, 73.609 doutoras orientaram 1.169.858 graduados e a primeira orientação realizada foi no ano 1968. Para o nível Iniciação científica, 60.981 doutoras orientaram 672.106 alunos e as três primeiras orientações foram realizadas no ano 1962. Por fim, para a opção Outra natureza, 45.283 doutoras realizaram a orientação de 567.771 orientados e as duas primeiras orientações realizadas em 1962.

Já para as orientações em andamento, no Pós-doutorado do RCDP, apenas 3.788 doutoras pesquisadoras realizam a supervisão de 5.435 pós-doutorandos. Quanto ao nível de formação ao Doutorado, 22.390 doutoras orientam 61.961 doutorandos. Já para a formação Mestrado, 33.311 doutoras orientam 87.935 mestrandos. Para a Especialização 6.077 doutoras orientam 14.922 orientados. Para o nível de formação Graduação, 19.699 doutoras orientam 55.424 graduandos. Na Iniciação científica, 30.085 doutoras orientam 72.114 alunos. Por fim, para a opção Outra natureza, 11.193 doutoras realizam a orientação de 36.779 orientandos.

Bolsa de estudo corresponde a um benefício fornecido ao estudante durante o período de sua formação, concedido por uma agência financiadora. Na Plataforma Lattes, ao preencher as informações quanto a orientação, é informado pela doutora se o orientado possuiu auxílio de bolsa de estudo ou não, e qual a agência que financiou. Assim, sumarizando esta informação, foi possível agrupar se o orientado foi contemplado com bolsas de estudo ou não, para cada formação acadêmica/titulação referente as orientações concluídas (Figura 5).

Figura 5 – Distribuição de bolsas de estudo quanto as orientações concluídas

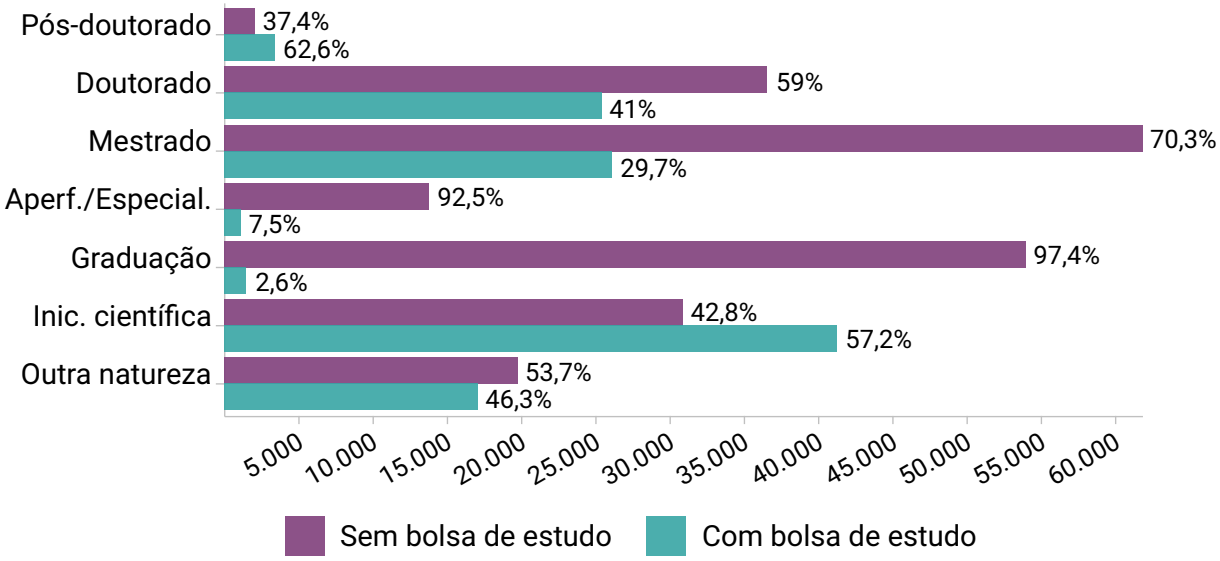


Fonte: Elaboração do autor

É relevante observar na Figura 5, como o percentual de bolsas foi significativo para as formações iniciação científica, doutorado e pós-doutorado. Dentre as sete opções, a Graduação possui o menor percentual de bolsas de estudo, 3,5%, e Pós-doutorado o maior, 74,36%.

Quanto as orientações em andamento, ao preencher as informações na Plataforma Lattes, as doutoras informam se o orientando possui algum auxílio de bolsa de estudo ou não, e qual a agência financiadora. Assim, sumarizando esta informação, foi possível agrupar se o orientando foi contemplado com bolsas de estudo ou não, para cada formação acadêmica/titulação referente as orientações em andamento (Figura 6).

Figura 6 – Distribuição de bolsas de estudo quanto as orientações em andamento



Fonte: Elaboração do autor

É relevante observar na Figura 6, como o percentual de bolsas está sendo significativo para as formações iniciação científica e pós-doutorado. Principalmente para mestrado e doutorado, é notável o percentual de queda no total de bolsistas se comparado ao percentual respectivo da Figura 5. Dentre as sete opções, a Graduação possui o menor percentual de bolsas de estudo, 2,6% e Pós-doutorado o maior, 62,6%. Outro aspecto relevante nas orientações corresponde ao atributo grande área preenchido pela doutora ao inserir uma orientação. A grande área pode ser definida como uma “aglomeração de diversas áreas do conhecimento, em virtude da afinidade de seus objetos, métodos cognitivos e recursos instrumentais refletindo contextos sociopolíticos específicos” (MINISTÉRIO DA CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES, 2019). O percentual de grandes áreas inseridos nas orientações é relativamente baixo. Para as pós-graduações doutorado e mestrado, este item está inserido em 58,1% e 55,9%, respectivamente, já para as demais formações acadêmicas/titulações, o preenchimento deste atributo é inferior a 47%. Já para

as orientações em andamento, todas as categorias de formação possuem preenchimento abaixo de 37%. Reunindo as informações quanto as orientações preenchidas, foi distribuído percentualmente para cada formação acadêmica/titulação as informações de orientações concluídas e em andamento de acordo com as grandes áreas (Tabela 2).

Tabela 2 – Percentual de orientações preenchidas por grandes áreas

Grandes áreas		**Pós-d.**	**Dout.**	**Mest.**	**Ap./Esp.**	**Grad.**	**Inic.C.**	**Outra n.**
Ciências	Concl.	9,24%	9,47%	7,66%	2,55%	6,05%	10,18%	13,27%
Agrárias	Andam.	7,02%	6,37%	5,00%	4,11%	4,49%	8,91%	10,08%
Ciências	Concl.	22,36%	16,12%	11,00%	3,86%	6,38%	15,01%	10,96%
Biológicas	Andam.	18,85%	10,18%	6,68%	3,99%	5,65%	12,17%	9,13%
Ciências	Concl.	12,61%	19,12%	18,32%	30,11%	20,93%	19,94%	20,80%
da Saúde	Andam.	14,96%	16,43%	17,09%	31,96%	20,96%	20,67%	18,41%
Ciênc. Exatas	Concl.	10,15%	7,94%	7,17%	2,68%	4,91%	10,17%	6,62%
e da Terra	Andam.	7,38%	6,77%	5,62%	2,58%	4,55%	8,83%	7,63%
Ciências	Concl.	20,70%	23,42%	26,33%	34,54%	24,98%	19,34%	22,40%
Humanas	Andam.	23,16%	29,63%	32,93%	29,74%	29,58%	20,88%	27,70%
Ciênc. Sociais	Concl.	6,11%	7,32%	11,90%	15,08%	24,28%	9,46%	11,09%
Aplicadas	Andam.	9,68%	10,56%	13,45%	14,80%	19,72%	11,34%	10,06%
Engenharias	Concl.	6,61%	5,68%	6,66%	1,98%	4,50%	6,87%	4,09%
	Andam.	5,53%	5,64%	5,42%	1,69%	4,96%	6,28%	3,52%
Ling. Letras	Concl.	11,97%	10,62%	10,34%	8,64%	7,60%	8,63%	10,26%
e Artes	Andam.	12,91%	13,74%	12,56%	10,34%	9,67%	10,32%	12,12%
Outra	Concl.	0,25%	0,31%	0,62%	0,55%	0,37%	0,38%	0,52%
	Andam.	0,51%	0,68%	1,26%	0,79%	0,41%	0,62%	1,35%
Total	Concl.	9.499	78.058	245.987	194.785	482.497	316.508	198.918
	Andam.	1.952	22.575	30.961	5.162	17.056	22.556	10.848

Fonte: Elaboração do autor

Assim, como pode ser observado na Tabela 2, para as orientações concluídas, a categoria Pós-doutorado com 9.499 orientações preenchidas possui as duas grandes áreas mais representativas igual a Ciências Biológicas (2.124) e Ciências Humanas (1.966). Para as categorias Doutorado, Mestrado, Especialização e Outra natureza, as duas grandes áreas mais representativas correspondem a Ciências Humanas e Ciências da Saúde. Já na Graduação, as duas grandes áreas mais representativas são Ciências Humanas e Ciências Sociais Aplicadas. Por fim, as duas grandes áreas mais representativas para a Iniciação científica são Ciências da Saúde e Ciências Humanas.

Ainda observando a Tabela 2, porém focando nas orientações em andamento, o Pós-doutorado com 1.952 orientações em andamento preenchidas, possui as duas grandes áreas mais representativas iguais a Ciências Humanas e Ciências Biológicas. Para as demais formações acadêmicas/titulações, as duas grandes áreas mais representativas correspondem a Ciências Humanas e Ciências da Saúde, alterando esta posição na formação Especialização.

Ao preencher as orientações na Plataforma Lattes, a doutora informa o ano de término nas orientações concluídas para cada tipo de formação acadêmica/titulação, e

da mesma forma, o ano de início para as orientações em andamento para cada tipo de formação acadêmica. Deste modo, foram recuperados todos os anos preenchidos nas orientações concluídas e realizado uma sumarização por ano, bem como, o percentual de crescimento para cada formação acadêmica/titulação referente aos dez últimos anos (Tabela 3).

Tabela 3 – Crescimento das orientações concluídas

Ano		Pós-d.	Dout.	Mest.	Ap./Esp.	Grad.	Inic.C.	Outra n.
2009	Quant.	974	6.341	20.950	28.535	67.743	34.940	31.364
	Cresc.	21,75%	10,07%	9,70%	6,28%	-1,04%	6,09%	7,81%
2010	Quant.	1.227	6.597	22.153	30.522	71.437	40.017	37.159
	Cresc.	25,98%	4,04%	5,74%	6,96%	5,45%	14,53%	18,48%
2011	Quant.	1.343	7.288	24.389	33.757	72.892	43.707	40.600
	Cresc.	9,45%	10,47%	10,09%	10,60%	2,04%	9,22%	9,26%
2012	Quant.	1.601	8.448	26.843	31.759	66.332	47.144	41.534
	Cresc.	19,21%	15,92%	10,06%	-5,92%	-9,00%	7,86%	2,30%
2013	Quant.	1.869	9.360	29.020	26.177	75.179	48.962	42.949
	Cresc.	16,74%	10,80%	8,11%	-17,58%	13,34%	3,86%	3,41%
2014	Quant.	2.134	10.608	30.648	28.042	78.192	48.433	45.434
	Cresc.	14,18%	13,33%	5,61%	7,12%	4,01%	-1,08%	5,79%
2015	Quant.	2.337	10.827	32.863	28.280	72.561	47.919	43.256
	Cresc.	9,51%	2,06%	7,23%	0,85%	-7,20%	-1,06%	-4,79%
2016	Quant.	2.291	10.044	34.430	24.201	77.931	47.618	43.688
	Cresc.	-1,97%	-7,23%	4,77%	-14,42%	7,40%	-0,63%	1,00%
2017	Quant.	2.059	9.772	31.609	18.478	80.772	45.300	42.107
	Cresc.	-10,13%	-2,71%	-8,19%	-23,65%	3,65%	-4,87%	-3,62%
2018	Quant.	1.738	9.416	25.741	16.600	70.194	30.954	33.754
	Cresc.	-15,59%	-3,64%	-18,56%	-10,16%	-13,10%	-31,67%	-19,84%

Fonte: Elaboração do autor

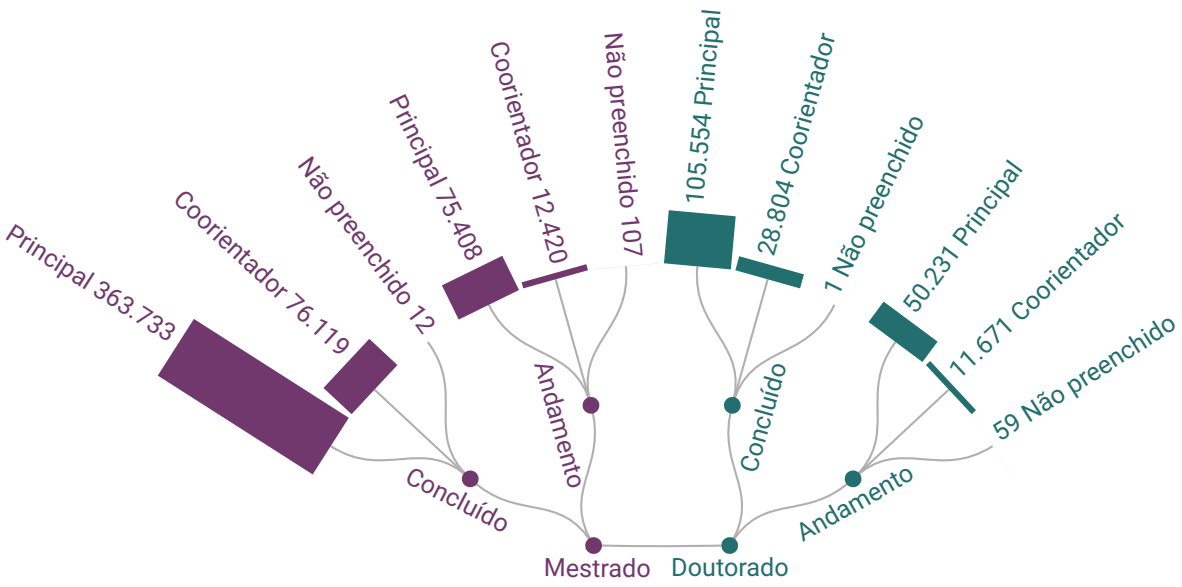
Assim, a Tabela 3 apresenta o decorrer de dez anos para todas as opções de formação referente a taxa de crescimento anual, iniciando em 2008 e finalizando em 2018. Não foi incluído o ano de 2019, pois os dados para o estudo foram coletados neste ano e como a quantidade para cada formação acadêmica/titulação estava abaixo de 50% com relação a 2018 os mesmos poderiam não ter sido incluído por todas as doutoras, e neste caso não são relevantes para o estudo. Nos anos 2010 e 2011, todas as formações tiveram crescimento positivo em relação ao ano anterior, enquanto que nos outros anos, ao menos um ano de alguma formação apresentou uma redução. O ano de 2018 é o único ano em que a quantidade de orientações teve redução no crescimento para todas as formações. As categorias Pós-doutorado e Doutorado tiveram crescimento quanto as orientações de 2009 até 2015, e Mestrado um ano a mais, até 2016.

A construção do conhecimento não ocorre de maneira isolada, é um processo interativo entre orientador/orientando, que resulta em desenvolvimento pessoal e outros inúmeros benefícios para ambos como enunciado por Ferreira, Furtado e Silveira (2009, p. 172):

> Para o orientando, resulta em crescimentos pessoal, profissional e acadêmico, encorajamento, direção, desenvolvimento de senso crítico, independência e auto-confiança. Já para o orientador ocasiona em aumento da satisfação pessoal, estímulo, oportunidade de manter-se atualizado em termos de técnicas e conhecimento, aumento da habilidade para atrair novos colaboradores pra projetos atuais e futuros, além de proporcionar oportunidade para "criar um legado" em gerações futuras na linha de pesquisa.

Neste contexto, a principal função do orientador compreende direcionar os orientandos ao longo de sua trajetória acadêmica, mantendo "relações singulares, intersubjetivas, complexas e ricas em detalhes com os orientandos, e, desta convivência, resultam dissertações e teses que contribuem para a sistematização e consolidação do conhecimento científico em determinada área" (LEITE.FILHO; MARTINS, 2006, p. 100). Para as formações mestrado e doutorado, é possível incluir orientador principal e coorientador. Assim, foi realizada a sumarização para estas duas formações, mestrado e doutorado, quanto as orientações concluídas e em andamento (Figura 7).

Figura 7 – Orientador principal e coorientador: mestrado e doutorado



Fonte: Elaboração do autor

É relevante observar na Figura 7 a representação importante das doutoras como orientadoras principais nos níveis de mestrado e doutorado, sendo que nas orientações concluídas corresponde a 78,56% para doutorado e 82,69% para o mestrado, já nas orientações em andamento corresponde a 81,07% para o dourado e 85,75% para o mestrado. É importante ressaltar que no decorrer do mestrado e doutorado é possível alterar o orientador principal e incluir coorientador

Os docentes-pesquisadores se envolvem em diversas atividades de pesquisa e ensino. Bourdieu (2004) introduziu o conceito de capital científico, onde as atividades

com maior prestígio e reconhecimento na ciência são de caráter científico “puro”, ou seja, aquelas relacionadas diretamente com o desenvolvimento científico, como uma produção bibliográfica. Já as atividades focadas somente no ensino não rendem capital “puro”, pois os docentes realizam atividades invisíveis e que não geram produção bibliográfica, e consequentemente, não alcançam reconhecimento pelo trabalho realizado.

Independentemente do valor simbólico, para que ocorra a produção bibliográfica, se faz necessária que haja uma parceria e uma relação unilateral entre orientador e orientado. É importante ressaltar que o processo de orientação envolve diversas dimensões além da produtiva, como sociais e afetivas. É imprescindível que ambas as partes se mantenham em constante diálogo e fácil comunicação, para que possam aprender, evoluir e construir em conjunto. Contudo, ressaltando o processo produtivo, é fundamental que o orientando seja receptivo ao receber orientações e aconselhamentos para pesquisas e leituras, e simultaneamente, o orientador seja capacitado para orientar e esteja em constante aprimoramento de suas habilidades como docente (LOPES et al., 2020), podendo os dois em conjunto desenvolver o projeto e tenham bons frutos do mesmo. Ao refletirmos sobre o processo de produção bibliográfica, a orientação pode ser considerada um importante aspecto para o meio acadêmico.

5.2 Representatividade da produção bibliográfica e técnica das doutoras

A produção bibliográfica é o resultado da pesquisa científica, no qual, possui como objetivos estudar os fenômenos da natureza e suas relações com o meio, bem como responder às necessidades da sociedade (LOPES et al., 2020). As produções bibliográficas e técnicas são partes integrantes desse processo de produção de conhecimento científico, no qual possui como elementos os artigos em periódicos, livros, capítulos de livros, anais de congressos, resumos, teses, dissertações e monografias, entre outros meios de divulgação e comunicação da ciência (DOMINGUES, 2014).

Os diversos tipos de produções podem ser cadastrados na Plataforma Lattes através do menu “Produções”, sendo possível escolher dentre as opções Produção bibliográfica, Produção técnica e Outra produção artística/cultural. As opções que compõe a Produção Bibliográfica são: Artigos completos publicados em periódicos; Artigos aceitos para publicação; Livros e capítulos; Texto em jornal ou revista (magazine); Trabalhos publicados em anais de eventos; Apresentação de trabalho e palestra; Partitura musical; Tradução; Prefácio, posfácio; Outra produção bibliográfica. Já em Produção Técnica, as opções correspondem a: Assessoria e consultoria; Extensão tecnológica; Programa de computador sem registro; Produtos; Processos ou técnicas; Trabalhos técnicos; Cartas, mapas ou similares; Curso de curta duração ministrado; Desenvolvimento de material didático ou instrucional; Editoração;

Manutenção de obra artística; Maquete; Entrevista, mesas redondas, programas e comentários na mídia; Relatório de pesquisa; Redes sociais, websites e blogs; Outra produção técnica. Por fim, as opções de cadastro para Outra produção artística/cultural são: Artes cênicas; Música; Artes visuais; Outra produção artística/cultural.

Como o conjunto de doutores são responsáveis pela maioria das produções científicas cadastradas na Plataforma Lattes (DIAS, 2016), realizou-se um levantamento das produções bibliográficas e técnicas das doutoras pesquisadoras. Assim, dentre todos os tipos de produções disponibilizados, foram selecionadas oito que possuem maior relevância para sumarizar e apresentar na Tabela 4 as produções do conjunto RCDP.

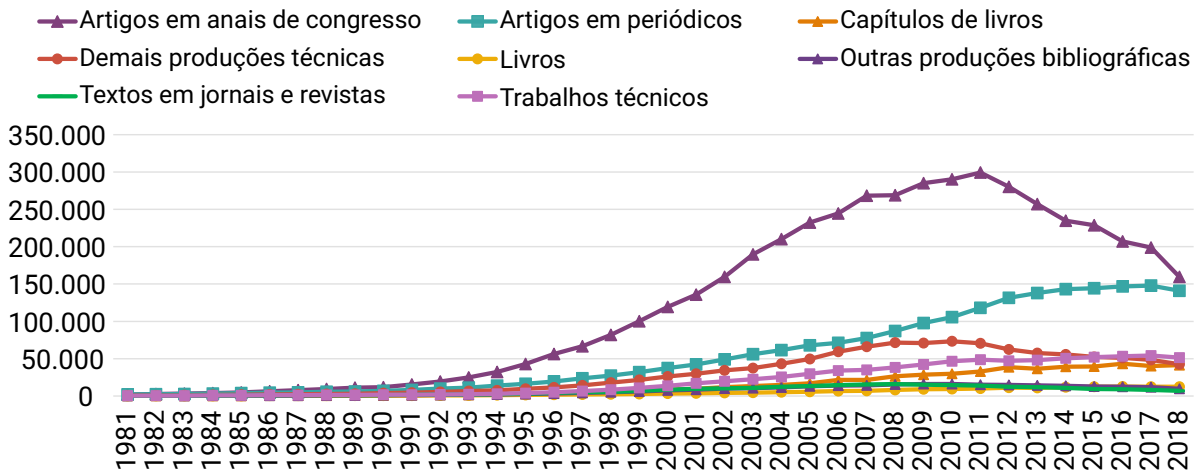
Tabela 4 – Quantitativo para as produções bibliográficas e técnicas

Tipo de produção	**Total**
Artigos em anais de congresso	4.811.242
Artigos em periódicos	2.155.370
Capítulos de livros	571.110
Demais produções técnicas	1.262.293
Livros	184.343
Outras produções bibliográficas	282.503
Textos em jornais e revistas	269.718
Trabalhos técnicos	807.089

Fonte: Elaboração do autor

As publicações podem ocorrer de diversas formas, porém para o conjunto RCDP a publicação e divulgação de conteúdo se concentra nos artigos em periódicos e nos artigos em anais de congresso. Livros e capítulos de livros possuem a quantidade bem inferior se comparado aos artigos. Utilizando estes dados, realizou-se uma sumarização das produções das doutoras por ano, conforme apresentado na Figura 8.

Figura 8 – Sumarização das produções bibliográficas e técnicas por ano



Fonte: Elaboração do autor

Para esta análise foram contabilizados todos os trabalhos registrados na Plataforma Lattes por currículo do RCDP, ou seja, o mesmo trabalho registrado em currículos distintos foi contabilizado em sua totalidade. A primeira publicação registrada no RCDP para artigos em periódicos ocorreu em 1940 e anais de congresso em 1948. Neste cenário de publicações por ano é possível perceber que, nos primeiros quarenta e um anos (1940 a 1981) a produção bibliográfica permaneceu tímida e constante para todos os tipos de produções. Possivelmente, uma hipótese para explicar esse fato pode ser relacionado ao lançamento e padronização do currículo da Plataforma Lattes que ocorreu em agosto de 1999, no qual os dados referentes as publicações anteriores a este período podem não ter sido divulgados pelas doutoras. Baseando no total de 7.412 doutoras formadas entre os anos de 1952 a 1999, disponível na aba "Evolução Formação" dos dados estatísticos do Painel Lattes[1], permite introduzir outra hipótese para explicar o fato descrito acima, que está relacionada ao número de doutoras, que eram poucas na época.

É relevante observar na Figura 8 que, a partir do início da década de 1980, ocorreu um aumento em todos os tipos de produções bibliográficas e tecnológicas. As produções que se evidenciaram das demais correspondem aos artigos em periódicos e em anais de congresso. Com relação aos artigos em periódicos, o mesmo apresentou um crescimento significativo até 2013, apresentou pouco crescimento após esse ano e queda em 2017. Já as produções dos artigos em anais e congressos tiveram um aumento considerável com ápice no ano de 2011 e queda significativa após esse ano. Esse declive acentuado referente foi tão expressivo que desde o ápice até o ano de 2018, apresentou uma queda de 46,68%, chegando ao final do último ano com valor total de artigos próximo ao total dos artigos em periódicos. Essa diminuição considerável no número de artigos em anais de congresso apresenta o mesmo comportamento no estudo de Dias (2016), para o total de publicações por ano referente ao conjunto de todos os doutores com currículos cadastrados na Plataforma Lattes.

São diferentes hipóteses que podem estar relacionadas ao declive acentuado dos artigos em anais de congresso a partir de 2011. Uma delas refere-se à classificação da produção bibliográfica utilizados pela CAPES, se considerarmos que o sistema de avaliação da produção influencia as ações dos indivíduos. Fato esse impulsionado pela não consideração dos artigos em anais de congresso nas avaliações dos programas de pós graduação. Assim, os artigos em periódicos tornam-se mais interessantes para a publicação, pois influenciam nos conceitos dos programas no qual as doutoras participam, e os mesmos direcionam seus esforços para esse tipo de publicação.

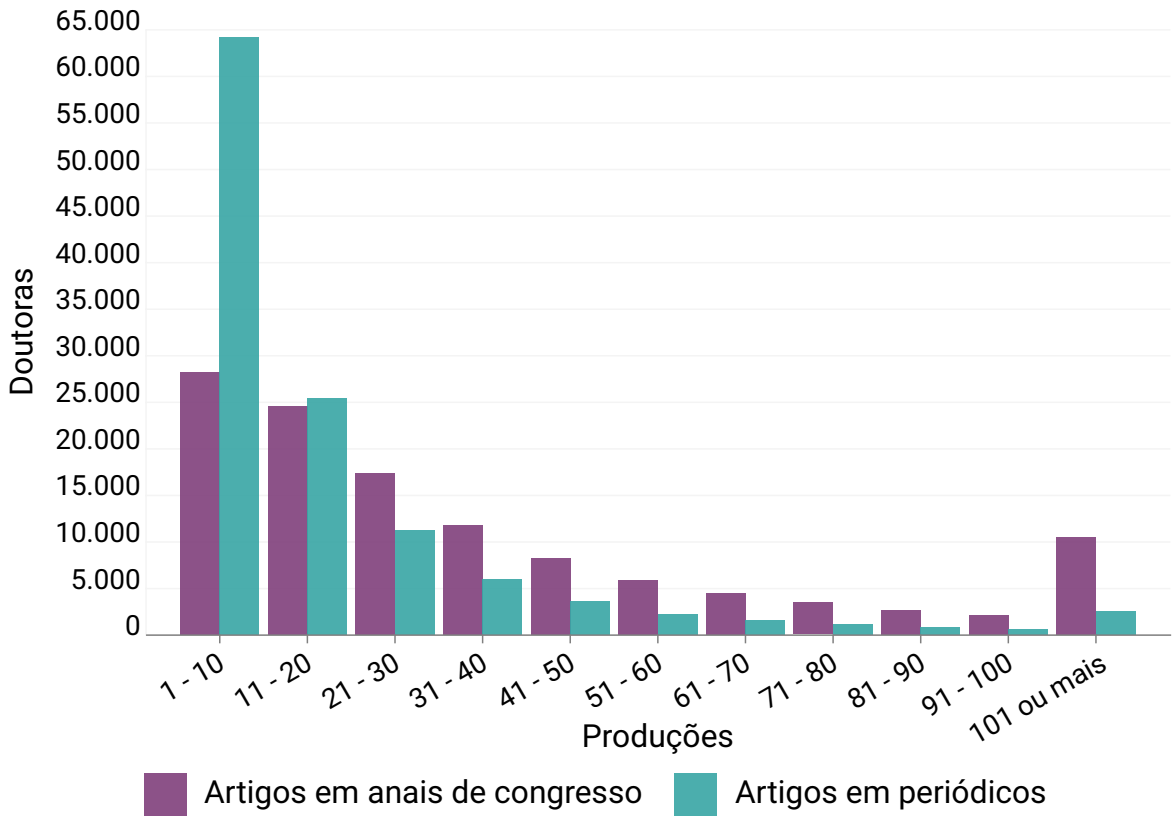
Atualmente, a "classificação dos periódicos é organizada através do módulo Qualis da Plataforma Sucupira, que coleta dados para avaliação dos programas de pós-graduação

[1] Painel Lattes - download dos dados dos painéis: http://estatico.cnpq.br/painelLattes/. Dados extraídos da base de currículos em 30 de novembro de 2016 e disponibilizados para consulta pública.

recomendados pela Capes" (FREIRE; FREIRE, 2017). O Qualis-Periódicos é "um sistema usado para classificar a produção científica dos programas de pós-graduação no que se refere aos artigos publicados em periódicos científicos" (SUCUPIRA, 2020), no qual afere a qualidade da produção a partir da análise da qualidade dos veículos de divulgação. A avaliação é periódica e possui as classificações de 2010-2012 e 2013-2016, no qual possui indicativos de qualidade A1, mais elevado; A2; B1; B2; B3; B4; B5; C - peso zero. A mensuração da produção utilizando o sistema Qualis proporciona o amadurecimento da pesquisa através da propagação do conhecimento, uma vez que a CAPES realiza a avaliação e incentiva a produção entre os docentes dos cursos de pós-graduação (VOGEL, 2015).

Continuando com a análise das produções bibliográficas das doutoras pesquisadoras e objetivando demonstrar a proporção delas com os números de produções, foi recuperado o total de artigos de cada doutora e distribuído em conjuntos. Assim, obteve-se dez conjuntos contendo um intervalo de 10 anos em cada e para o último conjunto agrupou-se as doutoras que possuem mais de 100 artigos, apresentando então no eixo vertical a quantidade de doutoras e no eixo horizontal o número de produções, para artigos em periódicos e em anais de congresso (Figura 9).

Figura 9 – Quantitativo das produções das doutoras por conjuntos



Fonte: Elaboração do autor

É importante ressaltar que, o número de doutoras do RCDP que não possuem produções corresponde a 6.509 (5,19%) para artigos em periódicos e 6.720 (5,35%) para artigos em anais de congresso. Já referente ao último conjunto, que corresponde as produções iguais ou maiores que 101 artigos, é importante ressaltar que o total de doutoras para os artigos em periódicos corresponde a 2.498 (1,99%) e para os artigos em anais de congresso 10.474 (8,34%) doutoras fazem parte deste conjunto.

Ao observarmos a Figura 9, o primeiro conjunto possui muita representação para os artigos em periódicos ao contemplar um total de 51,11%, ou seja, mais da metade das doutoras possui total de produções entre 1 até 10 artigos, que corresponde a um número baixo de produção para uma grande quantidade de doutoras se comparado aos outros conjuntos. O segundo conjunto possui menos do que a metade do conjunto anterior (20,24%), e esta queda ocorre sucessivamente para os seguintes. O conjunto que corresponde a 101 ou mais, compreende um total de 688.543 artigos em periódicos para 1,99% doutoras, ou seja, possui uma quantidade significativa de artigos para um pequeno número de doutoras. Já para os artigos em anais de congresso, os três primeiros conjuntos contemplam um total de 55,79%, ou seja, mais da metade das doutoras produziram de 1 até 30 artigos. O conjunto que corresponde a 101 ou mais, compreende um total de 1.772.670 artigos em anais de congresso para 8,34% doutoras, ou seja, possui uma quantidade significativa de produções para um pequeno número de doutoras.

Essa análise também foi realizada para livros e capítulos de livros. Um número considerável de doutoras do RCDP não possui produções, correspondendo a 77.540 (61,78%) para livros e 45.805 (36,49%) para capítulos de livros. Para as produções em livros, o conjunto mais representativo corresponde ao primeiro, de 1 a 10 livros publicados, totalizando 44.597 (35,53%) doutoras. São 12 (0,01%) doutoras agrupadas no conjunto 101 ou mais, sendo a maior produção de 134 livros. Já para os capítulos de livros, o primeiro conjunto também é o mais significativo, compreendendo 64.613 (51,48%) doutoras. São 89 (0,07%) doutoras agrupadas no conjunto 101 ou mais, sendo que a maior produção corresponde a 333.

É importante ressaltar que 13 doutoras hiper produtivas possuem cada uma mais de 500 artigos em periódicos, sendo que, a maior produção corresponde a 737 artigos. Já para artigos em anais de congresso 17 doutoras hiper produtivas possuem cada uma mais de 750 artigos produzidos, sendo que, a maior produção corresponde a 1.827. Duas doutoras são hiper produtivas tanto para artigos em anais de congresso, quanto para artigos em periódicos e possuem a mesma grande área Ciências da Saúde, são elas: Vera Luiza Capelozzi e Angela Maggio da Fonseca. As outras doutoras hiper produtivas possuem as seguintes grandes áreas: Ciências da Saúde 14, Ciências Exatas e da Terra 6, Ciências Agrárias 2, Ciências Biológicas 2, Engenharias 2, Ciências Humanas 1, Não informado 1.

Referente aos artigos em periódicos, as 13 doutoras hiper produtivas realizaram

a orientação de 3.246 pessoas (0,09% do total de orientações concluídas), e orientam 245 pessoas (0,07% do total de orientações em andamento). A média dessas orientações concluídas corresponde a 190,94 e a mediana 194, valores superiores aos de todas as doutoras (27,53 média e 12 mediana). Já a média e a mediana das orientações em andamento corresponde a 14,4 e 11 respectivamente, também superiores aos valores de todas as doutoras (2,64 média e 1 mediana). Para os artigos em anais de congresso que possui 17 doutoras hiper produtivas, elas realizaram a orientação de 1.772 pessoas (0,05% do total de orientações concluídas), e orientam 144 pessoas (0,04% do total de orientações em andamento). A média dessas orientações concluídas corresponde a 136,3 e a mediana 138, valores superiores aos de todas as doutoras (27,53 média e 12 mediana). Já a média e a mediana das orientações em andamento corresponde a 11,07 e 9 respectivamente, também superiores aos valores de todas as doutoras (2,64 média e 1 mediana). Um estudo destacando as hiper produtivas pode ser observado por Tuesta et al. (2019), no qual 5% das mulheres publicaram mais de 130 artigos. Vale ressaltar que, a Plataforma Lattes não disponibiliza informação de idade ou ano de nascimento na consulta pública do currículo do pesquisador e por isso, não é possível identificar se a hiper produção é um indicador dependente ou não da idade do pesquisador.

Os dados dos currículos da Plataforma Lattes são utilizados como uma das formas de avaliação para concursos oficiais, para editais de financiamento de projetos de pesquisa, por instituições de ensino, pelo CNPq, para concessão de bolsas, prêmios e títulos. Todas estas formas de avaliações motivam os pesquisadores brasileiros, principalmente para aqueles que estão ligados a instituições de ensino, manterem seus currículos atualizados (PERUCCHI, 2015). Assim, relacionando a Figura 9 com as datas de atualizações dos currículos, foi possível mapear e distribuir o percentual de atualização por ano para cada conjunto, conforme apresentado na Tabela 5.

Tabela 5 – Percentual de atualização dos currículos por produção

Produção	**2000 a 2004**	**2005 a 2009**	**2010 a 2014**	**2015**	**2016**	**2017**	**2018**	**2019**
0	0,54%	3,56%	14,63%	5,52%	6,78%	8,96%	18,79%	41,24%
1 - 10	0,22%	1,64%	5,82%	2,53%	3,71%	6,44%	16,20%	63,44%
11 - 20	0,17%	1,25%	4,11%	1,59%	2,40%	4,33%	11,29%	74,86%
21 - 30	0,12%	1,11%	4,00%	1,42%	1,85%	3,36%	8,30%	79,84%
31 - 40	0,20%	0,94%	3,54%	1,21%	1,90%	2,76%	7,65%	81,79%
41 - 50	0,20%	1,12%	2,78%	1,43%	1,74%	2,44%	6,03%	84,27%
51 - 60	0,05%	1,29%	3,60%	1,15%	1,85%	2,40%	6,14%	83,51%
61 - 70	0,06%	1,15%	3,95%	1,28%	1,47%	2,36%	6,63%	83,10%
71 - 80	0,18%	0,80%	2,68%	1,70%	0,89%	2,68%	6,25%	84,82%
81 - 90	0%	0,86%	2,44%	0,98%	0,98%	1,83%	5,13%	87,78%
91 - 100	0,52%	0,70%	2,61%	1,57%	1,92%	1,92%	4,88%	85,89%
101 ou mais	0,12%	0,76%	2,36%	0,92%	1,40%	1,76%	5,04%	87,63%

Fonte: Elaboração do autor

Assim, observando a Tabela 5, para as doutoras que não possuem artigos verifica-se o percentual de atualização baixo, se comparado aos outros. Uma hipótese para este fato pode estar relacionada a não atuação das doutoras no ensino e pesquisa. Como um dos requisitos ao participar de processos seletivos ou mesmo submeter projetos nos editais é ter informação atualizada no currículo, caso estas doutoras não estiverem atuando no ensino e pesquisa, não têm a necessidade de realizar periodicamente esta atualização. Já os três últimos conjuntos, no qual as doutoras pesquisadoras produziram mais de 81 artigos, possuem a data de atualização com os maiores percentuais para o ano de 2019. Vale ressaltar que, a atualização geral dos currículos apresenta um percentual baixo, porém a Plataforma Lattes não disponibiliza informação de currículo ativo ou não na consulta pública do currículo do pesquisador, visto que essa é uma informação relevante para calcular o percentual de atualização, não foi possível descartar as pesquisadoras inativas.

Conforme enunciado por Dias e Moita (2018), currículos com data atualizada proporcionam uma análise atual dos dados e da produção científica brasileira, já que currículos que possuem informações mais recentes provavelmente já possuem trabalhos recém-publicados registrados. Assim, sendo a produção bibliográfica uma importante atividade acadêmica, as doutoras do RCDP são ativas e atualizam seus currículos constantemente, principalmente àquelas que publicam uma quantidade elevada de artigos.

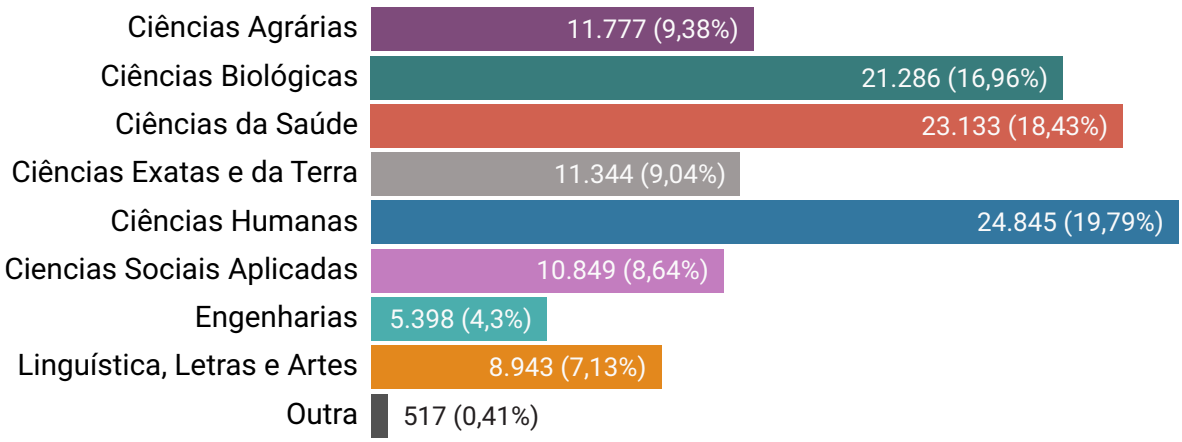
5.3 Distribuição das doutoras em suas grandes áreas de atuação

As informações que descrevem as atuações dos indivíduos podem ser cadastradas nos currículos da Plataforma Lattes através do menu “Atuação -> Áreas de atuação”, permitindo escolher e informar nove áreas e suas respectivas subáreas de conhecimento. A Plataforma Lattes utiliza a classificação de áreas e subáreas do CNPq[2].

Utilizando as grandes áreas de conhecimento preenchidas por cada doutora em seu respectivo currículo, foi selecionada a que apresentou maior ocorrência, considerando esta a mais relevante dentre as outras. Assim, agrupando os currículos das doutoras pesquisadoras por grandes áreas de atuação (Figura 10), tem-se como grandes áreas de atuação mais representativas relacionadas a humanas e menos representativas correspondendo as engenharias. Esses dados corroboram com o estudo de Olinto (2011), comprovando que os homens predominam nas carreiras exatas e as mulheres têm predominância nas áreas biológicas e saúde. Uma análise que apresenta a sub-representatividade das mulheres nos níveis mais altos da ciência brasileira para todas as grandes áreas do conhecimento é observado no estudo de Valentova et al. (2017).

[2] Tabela de Áreas do Conhecimento CNPq: https://cutt.ly/RglVN56

Figura 10 – Grandes áreas de atuação



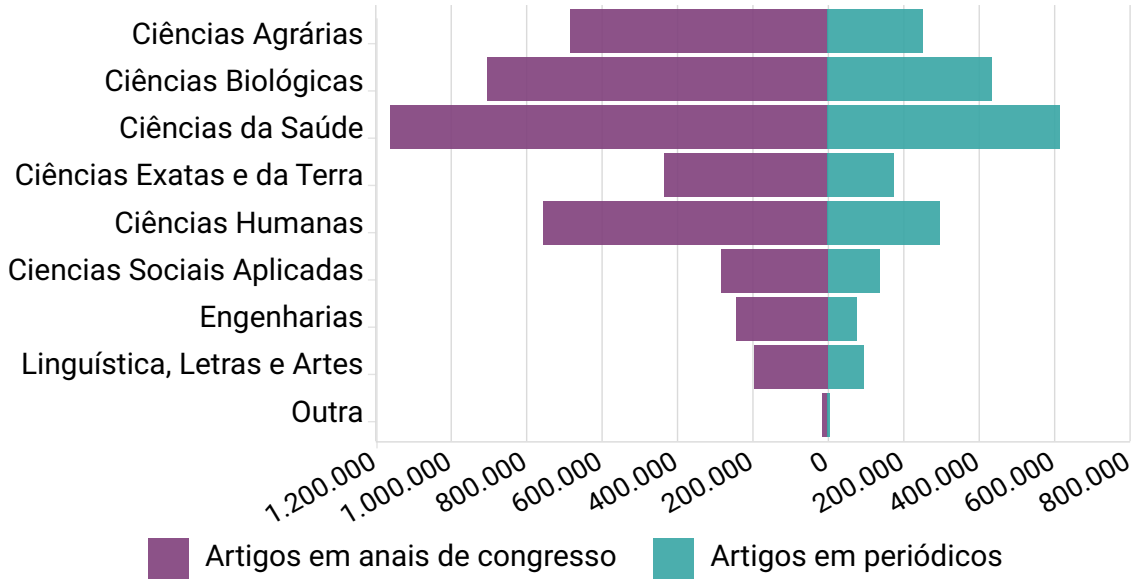
Fonte: Elaboração do autor

Um número considerável de currículos não possui informação relacionada a grande área de atuação, 7.423 (5,91%). Isso pode ser explicado pelo fato que, por não ser uma informação obrigatória, ao inserir uma nova formação acadêmica não existe a preocupação em preencher os dados desse campo, ou mesmo pode ocorrer esquecimento por parte das doutoras. Outro aspecto a ser levado em consideração refere-se a dificuldade em associar o curso com a área de atuação, principalmente para os cursos interdisciplinares e pós-graduações recentes, que resulta em, as doutoras não informam dados para este campo.

Os dados relacionados a produção científica têm sido muito expressivos atualmente, como no estudo conduzido por Elsevier (2017), que utilizou dados dos últimos 20 anos relacionados a produção cientifica em várias regiões do mundo e no qual, demonstra a diferença da participação feminina em diferentes perspectivas. Em um dos tópicos, no qual apresenta os dados do Brasil, refere-se ao período entre 2011 e 2015, onde pouco menos da metade (49%) dos estudos científicos foram produzidos por mulheres, contrastando com o período anterior entre 1996 a 2000, no qual as mulheres contribuíram com menor quantidade, apenas 38% da produção científica brasileira. Neste mesmo estudo, são apresentados diversos outros aspectos da diferença de gênero, demonstrando em números como as mulheres se apresentavam no período anterior e como foi evidente seu avanço no período mais atual, para o Brasil e outros regiões.

Seguindo nesta temática de produção científica, atrelada ao conceito da Tabela 4, no qual apresenta o quantitativo das produções científicas e tecnológicas do RCDP e permite visualizar que a divulgação de conteúdo concentra-se nos artigos em anais de congresso e nos artigos em periódicos, foi possível sumarizar os dados das grandes áreas, conforme apresentado na Figura 11.

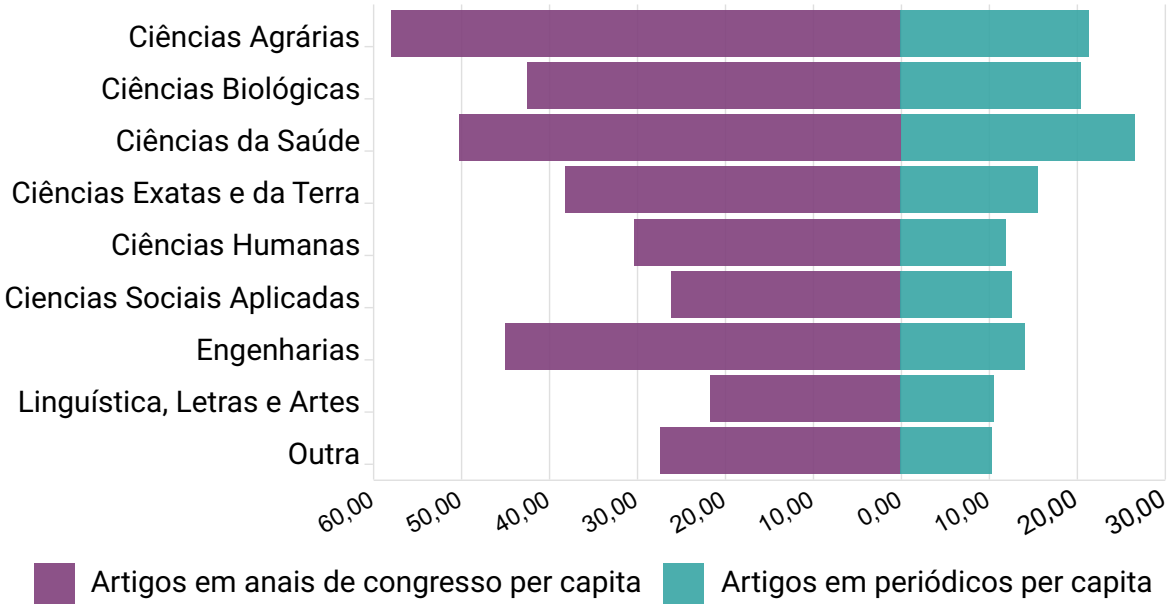
Figura 11 – Artigos em anais de congresso e em periódicos por grandes áreas



Fonte: Elaboração do autor

Observando a Figura 11, as grandes áreas de atuação para os artigos em anais de congresso e artigos em periódicos seguem o mesmo fluxo de grandes áreas mais representativas relacionadas às humanas da Figura 10. Porém, ao dividir o total de artigos pelo total de indivíduos de acordo com a grande área de atuação correspondente, temos um número de produções per capita, conforme apresentado na Figura 12.

Figura 12 – Artigos em anais de congresso e em periódicos per capita por grandes áreas



Fonte: Elaboração do autor

A análise da Figura 12 torna-se mais interessante, pois é possível verificar a produção sem influência do número total de doutoras de algumas áreas. Ao confrontarmos as duas figuras anteriores, percebe-se uma alteração considerável de posição entre algumas áreas de atuação na produção de artigos em anais de congresso e artigos em periódicos. Para os artigos em anais de congresso, a primeira área que se destaca em relação as outras é Ciências Agrárias, no qual na Figura 11 a mesma apresenta-se como a quarta maior produtora de artigos e na Figura 12 aparece em primeiro lugar. Já a área de Engenharias, teve uma alteração surpreendente e inesperada do sétimo (Figura 11) para o terceiro lugar (Figura 12), o que nos permite apurar que mesmo sendo uma área que possui pouca participação das doutoras, é uma área que produz muito conteúdo.

Já para os artigos em periódicos as produções per capita estão mais igualitárias dentre as grandes áreas. Confrontando as duas figuras anteriores, percebe-se o quanto as grandes áreas Ciências Agrárias, Ciências Exatas e da Terra e Engenharias se destacam em relação as outras grandes áreas pela quantidade de produção per capita em relação ao total de doutoras correspondente a cada uma delas, demonstrando que são grandes produtoras de conteúdo mesmo com quantidade menor de doutoras. Um estudo que apresenta as publicações e abrange todas as áreas da grande área Ciências Exatas e da Terra pode ser observado no estudo de Tuesta et al. (2019). Já uma abordagem distinta sobre as áreas e a modelagem de redes sociais de colaborações entre cientistas, pode ser observada nos estudos de Bellotti, Guadalupi e Conaldi (2016) e Bellotti, Kronegger e Guadalupi (2016).

Os currículos que possuem a data de atualização mais recente, proporcionam uma perspectiva atual da produção científica brasileira, pois provavelmente já possuem publicações recentes registradas (DIAS; MOITA, 2018). Assim, como 86.665 doutoras do RCDP possuem seus currículos atualizados em 2019, uma informação relevante corresponde ao percentual de atualização dos currículos por ano para cada grande área (Tabela 6).

Tabela 6 – Percentual de atualização dos currículos por grandes áreas

Grandes áreas	**2000 a 2004**	**2005 a 2009**	**2010 a 2014**	**2015**	**2016**	**2017**	**2018**	**2019**
Ciências Agrárias	0.24%	1.38%	4.67%	2.04%	3.12%	5.36%	13.58%	69.62%
Ciências Biológicas	0.25%	1.66%	5.79%	2.45%	3.62%	5.65%	13.95%	66.63%
Ciências da Saúde	0.14%	1.20%	5.16%	2.14%	2.89%	5.30%	13.68%	69.49%
Ciências Exatas e da Terra	0.27%	1.71%	5.65%	2.28%	3.24%	5.65%	13.05%	68.14%
Ciências Humanas	0.19%	1.49%	5.58%	2.25%	3.17%	4.88%	12.39%	70.05%
Ciências Sociais Aplicadas	0.10%	1.07%	4.54%	1.68%	2.65%	4.39%	11.88%	73.69%
Engenharias	0.13%	1.72%	5.46%	2.50%	2.93%	4.89%	12.25%	70.12%
Linguística, Letras e Artes	0.13%	1.71%	6.08%	2.36%	2.95%	4.93%	12.29%	69.54%
Não informado	0.58%	2.53%	4.82%	1.91%	3.44%	6.91%	15.94%	63.87%
Outra	0%	0.39%	11.41%	4.45%	2.51%	4.84%	11.80%	64.60%

Fonte: Elaboração do autor

Observando a Tabela 6, o percentual de atualização dos currículos em cada ano é bem próximo entre as grandes áreas. Mas se analisarmos somente o ano de 2019, as grandes áreas, Ciências Humanas, Ciências Sociais Aplicadas e Engenharias possuem mais de 70% de seus currículos atualizados. É relevante confrontar esta informação com os dados das Figura 10, pois para as duas grandes áreas Ciências Sociais Aplicadas e Engenharias, mesmo tendo menor número de doutoras possuem maiores percentuais de atualização de seus currículos.

Esses dados nos revelam a importância da produção científica das doutoras pesquisadoras nas diversas grandes áreas de atuação. Foi possível identificar o cenário de produções científica atual e verificar que mesmo as áreas que possuem menos participação feminina, como as Engenharias, a produção per capita apresenta o cenário oposto.

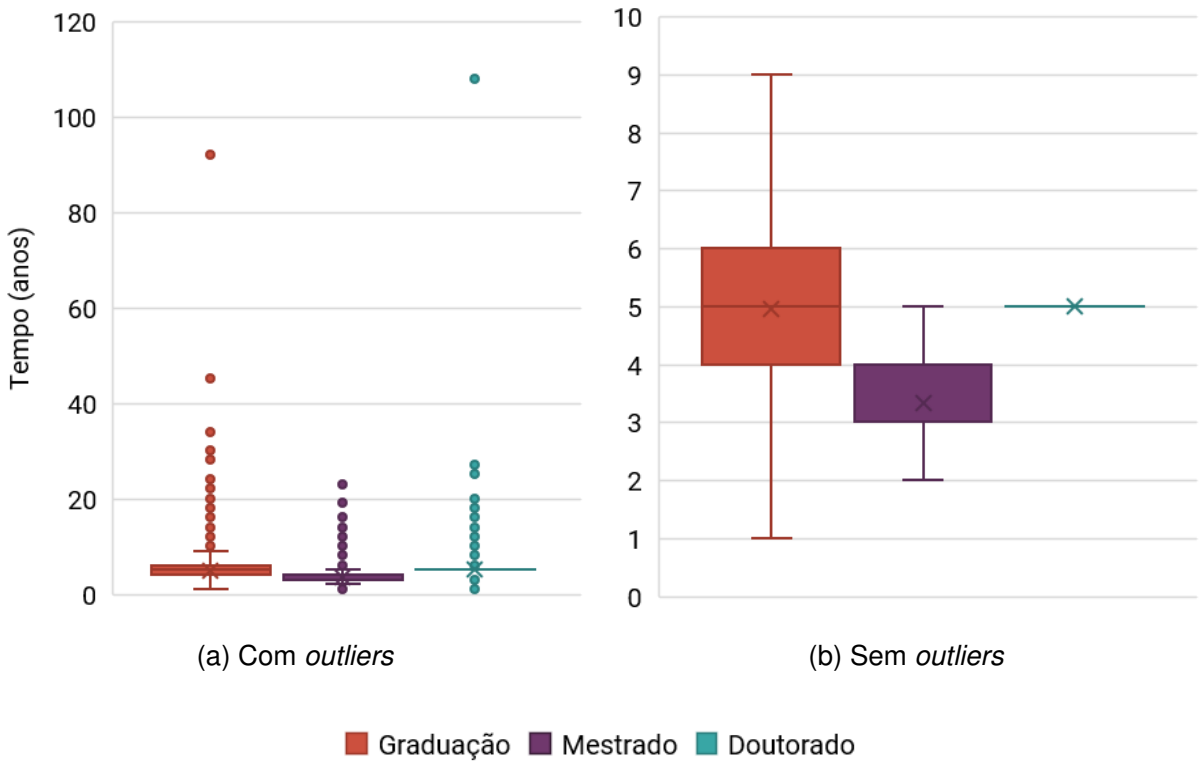
5.4 Evolução da formação acadêmica das doutoras

Os registros de formações acadêmicas podem ser informados nos currículos da Plataforma Lattes através de três tópicos: Formação acadêmica/titulação; Pós-doutorado e/ou livre-docência e Formação complementar. No tópico referente a Formação acadêmica/titulação é possível cadastrar dados de formação escolhendo dentre as opções: Doutorado; Mestrado; Mestrado profissional; Especialização; Especialização - Residência médica; Graduação; Ensino profissional de nível técnico; Ensino médio (2º grau); Ensino Fundamental (1º grau) e Aperfeiçoamento. Para todos os níveis de formações acadêmicas é possível preencher vários campos como instituição, nome do curso, ano inicial e final, título do trabalho, nome do orientador e coorientador, áreas, entre outros. Para as formações acadêmicas doutorado, mestrado e mestrado profissional, pode-se selecionar as grandes áreas de conhecimento e escolher dentre nove opções, selecionando as áreas e posteriormente as subáreas que se relacionam com a pós-graduação cursada, segundo a classificação de áreas e subáreas do CNPq.

Considerando o tempo de conclusão de cada titulação um tópico pertinente para análise da evolução acadêmica das doutoras pesquisadoras, foi realizado um levantamento geral referente as formações graduação, mestrado e doutorado. Foram recuperados do RCDP os dados referentes ao ano inicial e ano final de cada doutora para cada uma dessas titulações listadas acima e posteriormente realizado o cálculo do tempo de conclusão para cada formação acadêmica, usando a fórmula (AnoFinal-AnoInicial)+1. Deste modo, foi possível realizar a comparação entre as três formações acadêmicas conforme apresentado na Figura 13, sendo que, no eixo vertical apresenta o tempo de conclusão (anos) e no eixo horizontal as formações acadêmicas. Estes dados foram apresentados através do *boxplot* ou gráfico de caixa, utilizado para avaliar a distribuição empírica dos dados. Ao gerar os gráficos, tornou-se difícil a análise devido aos *outliers* (Figura 13a), e por isso na

(Figura 13b), foram ocultados os *outliers* para melhorar a visualização.

Figura 13 – Tempo de conclusão da graduação, mestrado e doutorado



(a) Com *outliers* (b) Sem *outliers*

Fonte: Elaboração do autor

Alguns campos do currículo da Plataforma Lattes não possuem preenchimento obrigatório e assim, os dados referentes ao ano inicial ou ano final que não continham informações foram caracterizados como “Não informado” e esses registros não foram utilizados para estas análises. Ressaltando que, o total de registros categorizados como Não informado corresponde a: graduação 4.283 (3,41%), mestrado 8.296 (6,6%) e doutorado 410 (0,33%).

A primeira parte da análise do *boxplot* inicia pelos traços que se estendem de ambos os lados da caixa, representando a amplitude dos dados para o fundo de 25% e o topo de 25% dos valores do conjunto, excluindo *outliers*. Assim, percebe-se na Figura 13 a diferença entre o tempo de conclusão para as três formações. Na graduação, como era esperado, o tempo de formação tem uma variação maior, ou seja, tem amplitude maior que compreende entre 1 e 9 anos. O mestrado apresenta o valor mínimo 2 e máximo 5 anos, e no doutorado a amplitude é menor, ou seja, possui pouca variação dos dados, com valores mínimo e máximo igual a 5 anos. Um segundo tópico relevante equivale aos valores que estão dentro da caixa do *boxplot* e representam 50% do total do conjunto de dados, ou seja, correspondem aos valores centrais do conjunto e distância entre o primeiro e o terceiro quartis. Assim, observando a Figura 13, para a graduação a frequência de formação

compreende entre 4 e 6 anos, para o mestrado entre 3 e 4 e para doutorado 5 anos. Outro tópico da análise corresponde a mediana, que é uma medida comum do centro dos dados e é representada pela linha dentro da caixa, já a média é representada pelo X. Assim, observando a Figura 13b, a graduação é simétrica pois a linha da mediana está no centro da caixa, a média sofre com os valores atípicos e está abaixo da linha da mediana, porém próxima. No mestrado não é possível identificar a linha da mediana. Já o doutorado possui apenas uma linha que significa a simetria e também pouca variação entre o valor máximo e o valor mínimo. Por fim, os *outliers* podem ser observados na Figura 13a e correspondem aos valores atípicos, ou seja, fora do limite superior e inferior. Podemos verificar que eles aparecem para as três formações, porém para mestrado e doutorado aparece abaixo e acima dos limites inferior e superior. É muito provável que os valores atípicos da graduação e do doutorado que se encontram nas posições 92 e 108 respectivamente, correspondem a um erro de inserção de dados.

Uma outra análise possível de ser realizada com estes mesmos dados é através da grande área, que pode ser informada ao inserir a formação acadêmica no currículo da Plataforma Lattes. Assim, além do ano inicial e final, foram recuperadas as grandes áreas para as duas pós-graduações, pois na graduação não é permitido inserir esta informação. Desta forma, verificando o tempo de conclusão da formação acadêmica das doutoras para estes dois níveis de formação, foi possível quantificar por grandes áreas de conhecimento conforme apresentado na Tabela 7.

Tabela 7 – Tempo de conclusão das formações acadêmicas por grandes áreas

		Mestrado		
Grandes áreas	**Total**	**Média**	**Mediana**	**Desvio padrão**
Ciências Agrárias	9.432 (7,51%)	3,17	3	0,77
Ciências Biológicas	15.432 (12,29%)	3,31	3	0,91
Ciências da Saúde	15.263 (12,16%)	3,24	3	1,02
Ciências Exatas e da Terra	9.195 (7,33%)	3,39	3	0,98
Ciências Humanas	19.527 (15,56%)	3,61	3	1,32
Ciências Sociais Aplicadas	8.039 (6,40%)	3,50	3	1,25
Engenharias	5.517 (4,40%)	3,33	3	0,93
Linguística, Letras e Artes	7.739 (6,17%)	3,57	3	1,23
Outra	299 (0,24%)	3,13	3	0,73
		Doutorado		
Grandes áreas	**Total**	**Média**	**Mediana**	**Desvio padrão**
Ciências Agrárias	7.918 (6,31%)	4,78	5	0,91
Ciências Biológicas	15.359 (12,24%)	5,13	5	0,98
Ciências da Saúde	15.716 (12,52%)	4,62	5	1,08
Ciências Exatas e da Terra	8.911 (7,10%)	5,34	5	1,17
Ciências Humanas	19.855 (15,82%)	5,15	5	1,21
Ciências Sociais Aplicadas	7.372 (5,87%)	5,10	5	1,28
Engenharias	5.539 (4,41%)	5,34	5	1,09
Linguística, Letras e Artes	7.717 (6,15%)	5,16	5	1,18
Outra	442 (0,35%)	5,04	5	0,80

Fonte: Elaboração do autor

Como é possível preencher mais de uma área de conhecimento para mestrado e doutorado, para essa análise foi selecionado somente a primeira grande área preenchida pela doutora, considerando ser a mais relevante dentre as áreas cadastradas. Alguns campos do currículo da Plataforma Lattes não possuem preenchimento obrigatório e os dados referentes as grandes áreas de conhecimento que não continham informações foram caracterizados como “Não informado” e esses registros não foram utilizados. Ressaltando que, para o mestrado 35.072 (27,94%) registros foram categorizados como Não informado e para o doutorado 36.686 (29,22%) registros agrupados nesta mesma categoria. As informações apresentadas na Tabela 7 referentes a média, mediana e desvio padrão são cálculo estatísticos relevantes, pois permitem verificar a tendência central e a variabilidade dos dados para cada uma das grandes áreas. Sendo assim, podemos definir esses termos como: média sendo o valor que demonstra a concentração dos dados de uma distribuição; a mediana pode ser considerada como uma tendência central que indica exatamente o valor central de determinados dados; o desvio padrão demonstra a regularidade referente a um conjunto de dados de modo a apontar o grau de oscilação destes em comparação com a média dos valores do conjunto.

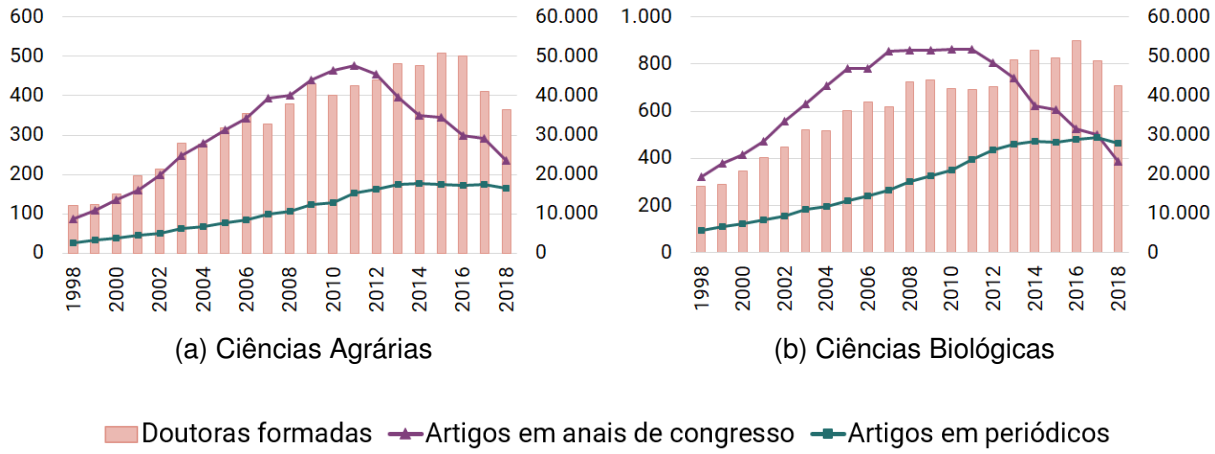
Observando a Tabela 7, podemos verificar que a mediana referente ao tempo de formação acadêmica em todas as grandes áreas não teve alteração para os dois níveis de pós-graduação. Para o mestrado, quanto aos dados da média, o valor ultrapassou a mediana em todas as grandes áreas, sendo a menor média de tempo de formação 3,13 e a maior 3,61. Como a média sofre influência dos valores atípicos, buscou no conjunto de dados a quantidade desses registros, totalizando 5.287 *outliers*, no qual a menor duração da formação acadêmica corresponde a 1 ano e a maior corresponde a 23 anos, este com grande área correspondendo a Linguística, Letras e Artes. O maior desvio padrão (1,32) corresponde a grande área Ciências Humanas, indicando o quanto os dados desta grande área estão afastados da média. Ao analisar a Figura 13b para o mestrado, não é possível identificar a linha da mediana. Porém, ao relacionarmos a Figura 13b com a Tabela 7, através do cálculo da mediana, percebe-se que os dados são assimétricos positivos, e que a mediana está conectada ao primeiro quartil.

Já a pós-graduação Doutorado apresentou maior variação entre as grandes áreas para a média, sendo as duas grandes áreas Ciências Agrárias e Ciências da Saúde possuem a média menor que a mediana, respectivamente 4,78 e 4,62. Este fato ocorreu, pois, existem uma quantidade maior de *outliers* superiores a 5 e menos dados menores que 5. Já para as outras grandes áreas que possuem a média maior que a mediana, ocorre o contrário. Um dado interessante para o Doutorado corresponde aos valores que influenciaram a média (*outliers*), compreendendo uma quantidade de 56.746 registros, no qual a menor duração de formação corresponde a 1 e a maior corresponde a 25. O maior desvio padrão que indica o quanto os dados do conjunto estão afastados da média corresponde a grande área Ciências Sociais Aplicadas (1,28), sendo que a mesma não possui a maior média.

Vale ressaltar que, a Plataforma Lattes não disponibiliza os meses de início e de conclusão das formações na consulta pública do currículo e por isso, a fórmula utilizada para o cálculo das formações contempla de modo geral o tempo de conclusão do mestrado e do doutorado, podendo apresentar uma variação acima do padrão devido a ausência dessa informação nos currículos. Outra questão importante a se destacar, que pode inferir no tempo de formação do mestrado e doutorado das mulheres, refere-se ao período de afastamento relativo a licença-maternidade. No momento da coleta dos dados, essa informação não era inserida nos currículos, porém conforme enunciado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq (2021), a próxima versão do currículo disponibilizado pela Plataforma Lattes, permitirá o registro dos períodos de licença-maternidade das pesquisadoras. Essa informação será importante e poderá ser realizado um estudo mais detalhado sobre esse tópico no tempo de formação das doutoras.

Como a formação acadêmica é um processo que se inicia antes da graduação, as práticas cotidianas realizadas na universidade são influenciadas pelas experiências adquiridas anteriormente, através da escola, da família, dos profissionais e da sociedade (BESUTTI; REDANTE; FÁVERO, 2017). Assim, a formação acadêmica também pode ser considerada um dos fatores que contribui amplamente para o seu desempenho como pesquisadores, no qual possui a produtividade acadêmica como consequência científica de sua pesquisa e os impactos sociais que ela proporciona (SOARES; NOVA, 2017). Neste contexto, foi realizado um levantamento contemplando os 20 anos mais recentes do RCDP para as nove grandes áreas, comparando o ano da formação acadêmica das doutoras com a perspectiva de produção científica (Figura 14 e Figura 15), do ano 1998 até 2018. Não foi incluído o ano de 2019, pois os dados foram coletados neste ano e como a quantidade de artigos estava abaixo de 50% com relação a 2018, não seriam relevantes para o estudo.

Figura 14 – Publicações por grandes áreas e formação das doutoras

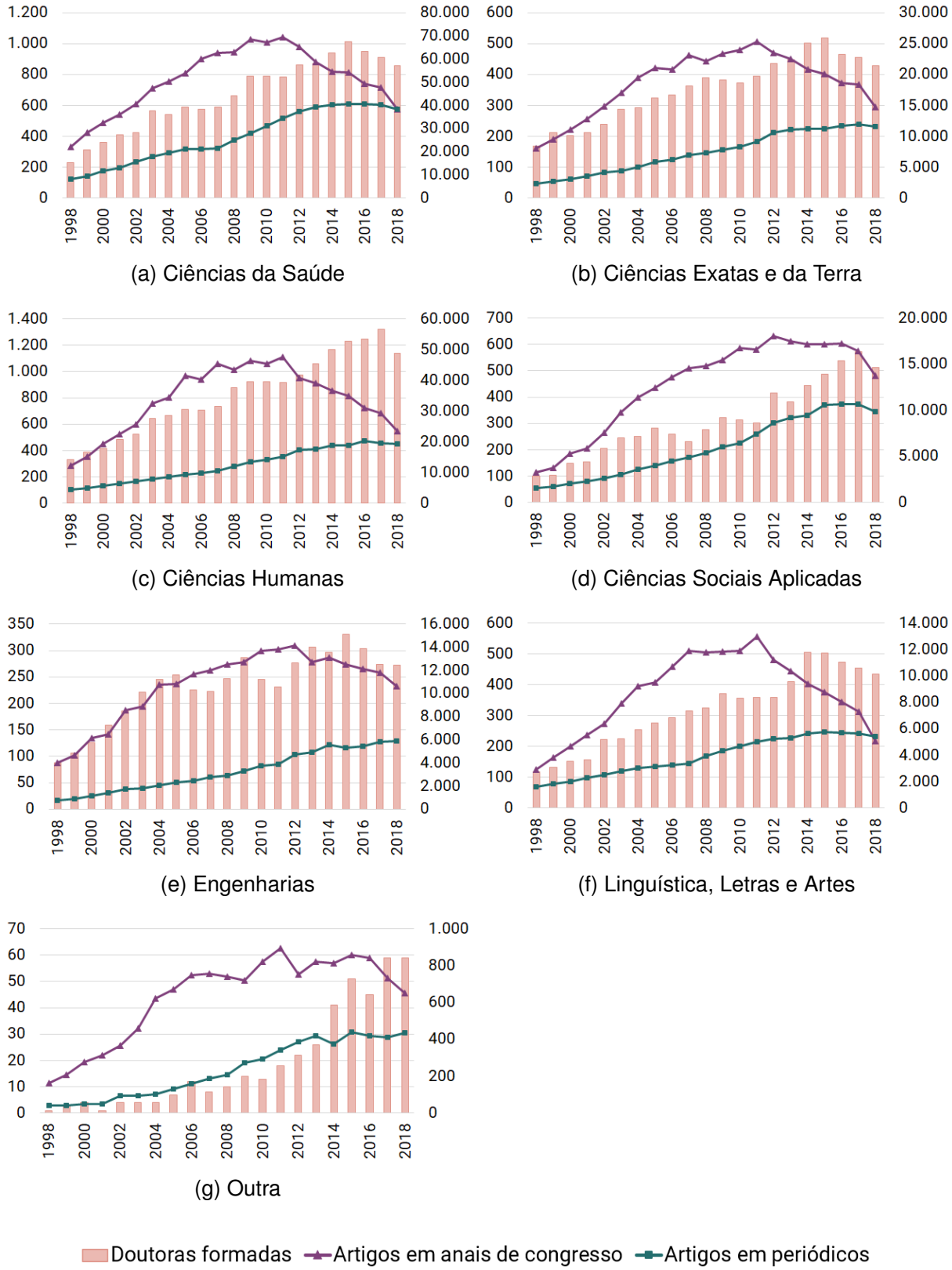


(a) Ciências Agrárias

(b) Ciências Biológicas

Fonte: Elaboração do autor

Figura 15 – Publicações por grandes áreas e formação das doutoras (continuação)



Fonte: Elaboração do autor

Assim, a Figura 14 e Figura 15 apresentam os dados por anos relativos ao total de doutoras pesquisadoras formadas na pós-graduação doutorado confrontando com as publicações de artigos em anais de congresso e artigos em periódicos, no qual, percebe-se que para os artigos em periódicos a tendência de crescimento segue, em sua grande maioria, a mesma disposição da formação das doutoras. Já para os artigos em anais de congresso, a tendência diverge quanto a formação das doutoras, principalmente Figura 15d e Figura 15f.

Relacionando o início do período de queda dos artigos em anais de congresso no ano de 2011 da Figura 8 (Seção 5.2) com as Figura 14 e Figura 15, percebe-se que as duas grandes áreas Ciências Sociais Aplicadas (Figura 15d) e Engenharias (Figura 15e), divergem do restante ao iniciar a queda dos artigos no ano de 2012. É relevante observar que, para todas as grandes áreas com início do período de queda de publicações em 2011/2012, tem discrepância com o crescimento da formação das doutoras pesquisadoras destes anos até os anos de 2015/2016. Para as grandes áreas Ciências Biológicas (Figura 14b), Ciências da Saúde (Figura 15a) e Linguística, Letras e Artes (Figura 15f), o número total de artigos em periódicos superaram o número total dos artigos em anais de congresso, demonstrando a diminuição considerável deste tipo de publicação. Na grande área Ciências Humanas (Figura 15c), entre os anos de 2005 e 2011, teve um período de instabilidade quanto as publicações, enquanto que a formação das doutoras continuou crescente e constante.

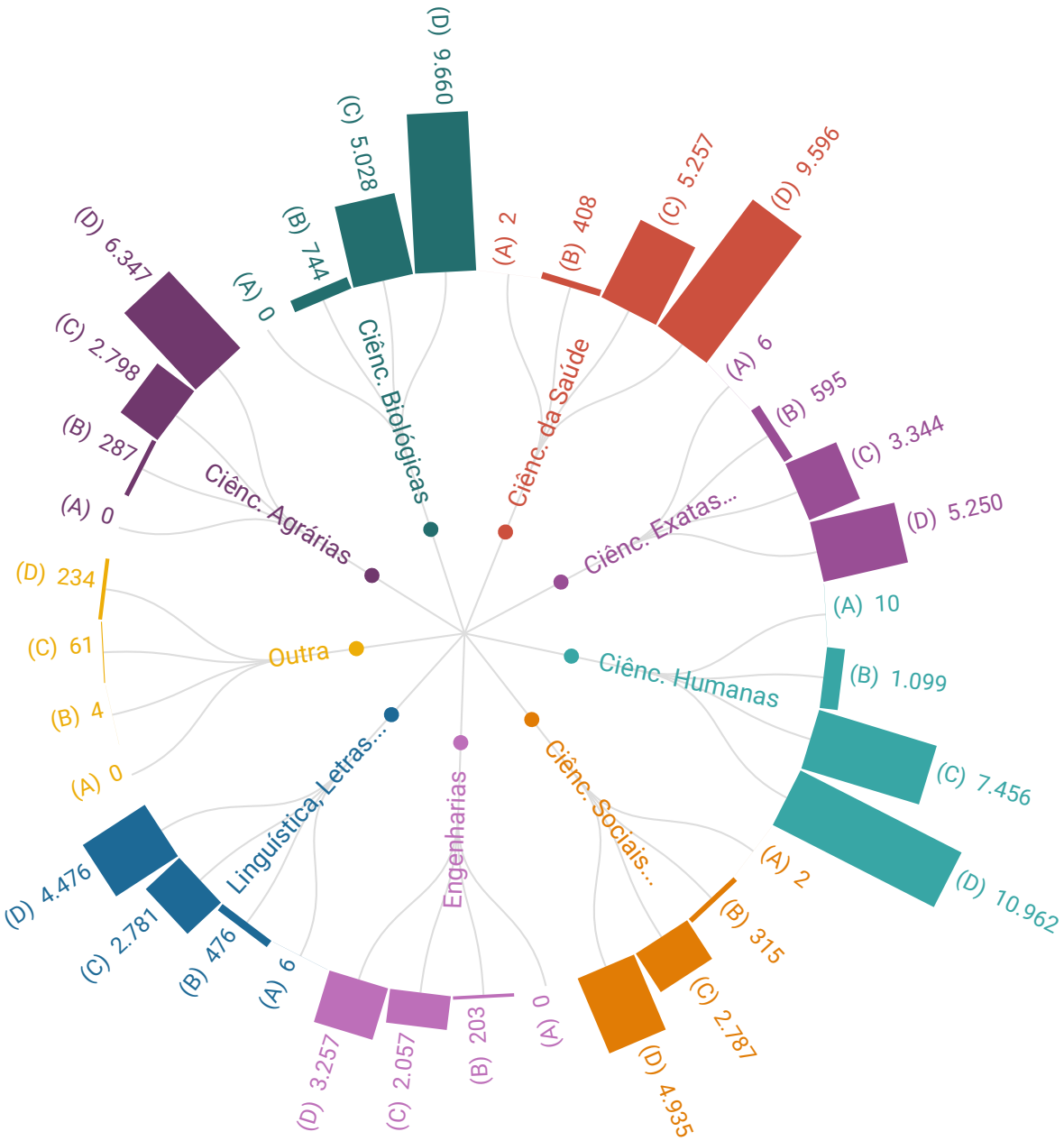
Devido ao seu crescimento e após décadas de evolução e reformas, a pós-graduação brasileira elevou o país a uma categoria relevante na produção científica internacional (ALMEIDA; GUIMARÃES, 2013). A comunicação científica associada a formação acadêmica é, portanto, a disseminação de dados e de novas descobertas, a interação do meio científico e o reconhecimento pelos seus pares. Assim, caracterizar a formação acadêmica e as trajetórias dos docentes é importante, visto que, é possível verificar o impacto na qualidade dos programas de pós-graduação aos quais estão associados (SILVA et al., 2016). Existem dois tipos de pós-graduações, os cursos *stricto sensu* e *lato sensu*. Conforme enunciado por Martins (2018, p. 16), a categoria *stricto sensu* inclui:

> ... o mestrado e o doutorado, cujo objetivo seria de natureza acadêmica, de pesquisa e de cultura, tendo como compromisso o avanço do saber. Esses cursos, na medida em que estariam ligados à essência da universidade, deveriam se constituir atividades regulares e permanentes e conferir diplomas de mestre e de doutor, sendo que esse último corresponderia ao nível mais elevado na hierarquia dos cursos superiores.

Assim, as pós-graduações *stricto sensu* mestrado e doutorado são títulos acadêmicos concedidos ao indivíduo por uma instituição de ensino superior, com o intuito de certificar academicamente a aptidão do candidato para desenvolver uma pesquisa e aprofundar o conhecimento em determinado campo da ciência, sendo possível seguir a mesma área de formação de conhecimento ou não. Neste contexto, procurando entender a evolução

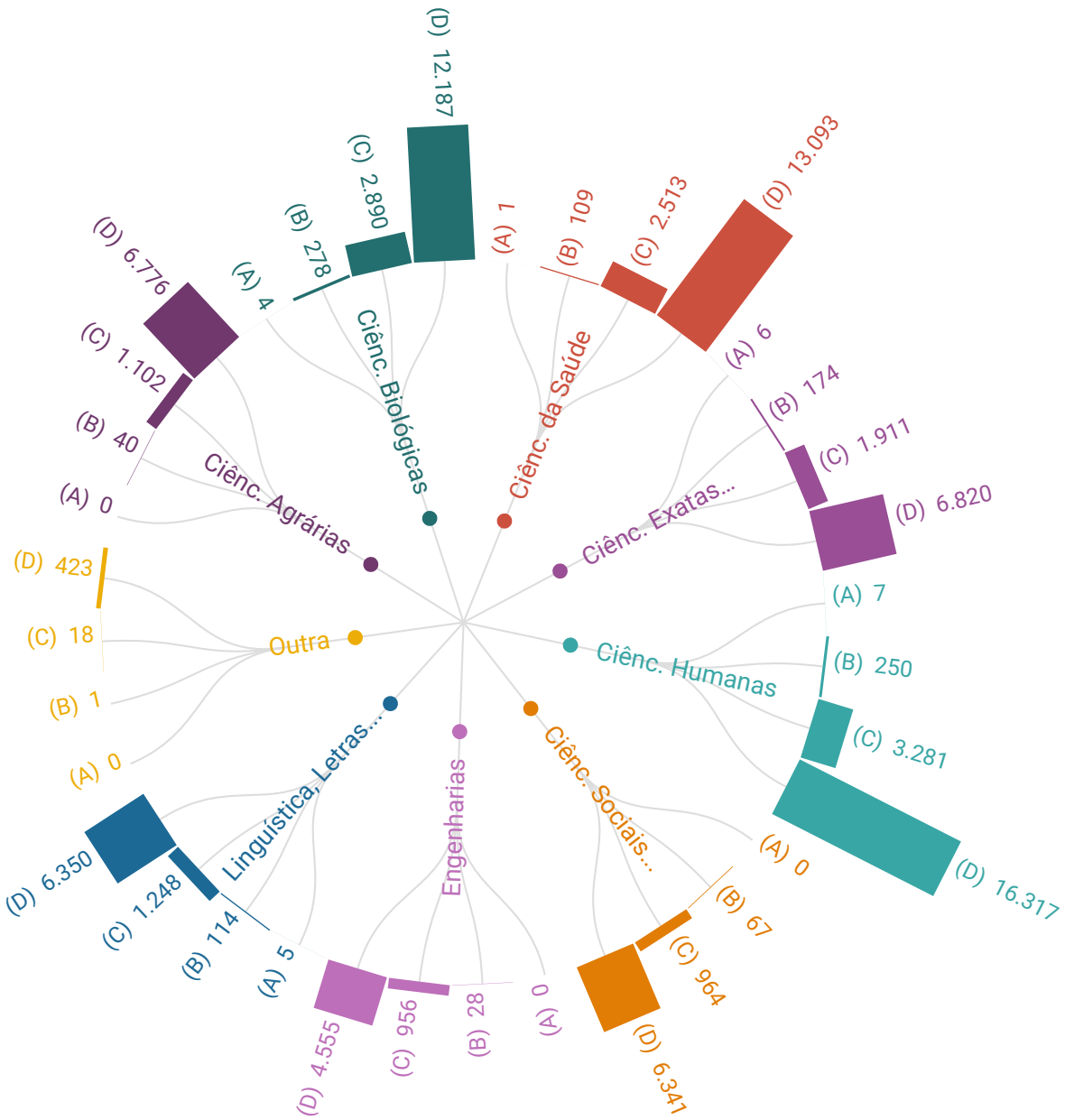
através das grandes áreas, sumarizou-se o ano de conclusão de cada uma das formações acadêmicas das doutoras pesquisadoras para cada área de conhecimento, tanto para o mestrado (Figura 16) e quanto para o doutorado (Figura 17). É importante destacar que, para os currículos do RCDP, o primeiro ano de formação do mestrado corresponde a 1949 e para o doutorado 1948, e o ano mais recente de formação para estas duas pós-graduações equivale a 2019. Neste contexto, para melhorar a análise dos dados, os anos de formações foram distribuídos e agrupados em quatro períodos, que correspondem a: **A**: 1948 a 1965; **B**: 1966 a 1983; **C**: 1984 a 2001; **D**: 2002 a 2019.

Figura 16 – Evolução da formação acadêmica através das grandes áreas: mestrado



Fonte: Elaboração do autor

Figura 17 – Evolução da formação acadêmica através das grandes áreas: doutorado



Fonte: Elaboração do autor

Observando a Figura 16, Ciências Humanas é a grande área que permaneceu em primeiro lugar quanto ao total de doutoras formadas para mestrado em todos os períodos. Referente ao período **A**, além desta área, outras que são relevantes são Ciências Exatas e da Terra e Linguística, Letras e Artes. Neste primeiro período, apenas cinco das nove grandes áreas tiveram doutoras formadas. Já no período **B**, possui formação em todas as grandes áreas, sendo Ciências Biológicas uma área com crescimento considerável para este período. Para os períodos **C** e **D**, além da grande área Ciências Humanas, as mais representativas quanto ao total de doutoras formadas são Ciências Biológicas e Ciências da Saúde que alternam as posições entre os períodos. Quanto ao crescimento

entre períodos, do período **B** para **C** as grandes áreas apresentaram as seguintes taxas de crescimento: Outra (14,25%), Ciências da Saúde (11,88%), Engenharias (9,13%), Ciências Agrárias (8,75%), Ciências Sociais Aplicadas (7,85%), Ciências Humanas (5,78%), Ciências Biológicas (5,76%), Linguística, Letras e Artes (4,84%), Ciências Exatas e da Terra (4,62%). Já as taxas de crescimento das grandes áreas do período **C** para o **D** correspondem a: Outra (2,84%), Ciências Agrárias (1,27%) e Ciências Biológicas (0,92%), Ciências da Saúde (0,83%), Ciências Sociais Aplicadas (0,77%), Linguística, Letras e Artes (0,61%), Engenharias (0,58%), Ciências Exatas e da Terra (0,57%), Ciências Humanas (0,47%).

Já verificando a Figura 17, o primeiro período também possui formação em apenas cinco áreas, porém não correspondem as mesmas grandes áreas do mestrado para este mesmo período. Ciências Humanas foi a área com mais formações para os períodos **A**, **C** e **D**. Quanto ao crescimento entre períodos, do período **B** para **C** as grandes áreas apresentaram as seguintes taxas de crescimento: Engenharias (33,14%), Ciências Agrárias (26,55%), Ciências da Saúde (22,06%), Outra (17%), Ciências Sociais Aplicadas (13,39%), Ciências Humanas (12,12%), Ciências Exatas e da Terra (9,98%), Linguística, Letras e Artes (9,95%), Ciências Biológicas (9,4%). Já as taxas de crescimento das grandes áreas do período **C** para o **D** correspondem a: Outra (22,5%), Ciências Sociais Aplicadas (5,58%), Ciências Agrárias (5,15%), Ciências da Saúde (4,21%), Linguística, Letras e Artes (4,09%), Ciências Humanas (3,97%), Ciências Biológicas (3,22%), Engenharias (3,76%), Ciências Exatas e da Terra (2,57%).

A Plataforma Lattes pode ser considerada a memória do pesquisador disponibilizada livremente na Web. Durante toda sua trajetória acadêmica e profissional, ele informa todos os registros de suas produções bibliográficas, técnicas e outros tipos de produções, além de contemplar os dados de formação acadêmica, áreas de atuação, atuação profissional, e diversos dados complementares. Segundo Brito, Quoniam e Mena-Chalco (2016), a Plataforma Lattes é um importante sistema de informação curricular, no qual pode ser utilizada principalmente para avaliações isoladas de grupos ou instituições. Estes 71 anos de história das pós-graduações mestrado e doutorado apresentam a evolução das grandes áreas tendo como referência o ano de formação registrado no currículo da Plataforma Lattes. O período **D** que é o mais recente, 2002 a 2019, possui as grandes áreas Ciências Humanas, Ciências Biológicas e Ciências da Saúde como as mais representativas. Mas um dado relevante correspondeu ao crescimento das grandes áreas Engenharias e Ciências Sociais Aplicadas, demonstrando como as doutoras pesquisadoras ampliaram suas áreas de formação. Assim, observando o cenário atual que elas ocupam e seus diversos aspectos, é possível observar sua evolução em todas as grandes áreas, principalmente uma “presença maior de mulheres em áreas mais interdisciplinares, relacionadas a aplicação da computação ou com aspecto social claro” (LORENS et al., 2020, p. 9).

5.5 Publicações das doutoras após a conclusão do doutorado

O mestrado e doutorado são níveis independentes de formação que compreendem a pós-graduação *stricto sensu*, no qual possuem como uma de suas finalidades estimular e produzir conhecimento científico. A pesquisa cientifica é o elemento indispensável ao desenvolvimento da pós-graduação e encontra-se diretamente associada a formação acadêmica, por isso pode ser considerado um dos requisitos base para a produção acadêmica e científica dos pesquisadores e seus docentes (ALMEIDA; GUIMARÃES, 2013). Segundo a CAPES (2014), a pós-graduação *stricto sensu* deve contemplar as seguintes características: "ser de natureza acadêmica e de pesquisa e, mesmo quando voltado para setores profissionais, ter objetivo essencialmente científico.", sendo que estas formações são partes integrantes do complexo universitário e tem como propósito essencial a universidade.

A pós-graduação possui como premissa a formação de pesquisadores em campos específicos e menor abrangência quanto à formação pedagógica dos professores, ou seja, o objetivo real é preparar pesquisadores (MACHADO ; LAMPERT; FALAVIGNA, 2017). No decorrer de sua formação acadêmica, os pesquisadores desenvolvem diversas pesquisas e o meio de publicação mais utilizado por eles para divulgar seus trabalhos, provenientes de dissertações e teses, são através dos artigos, seguidos de capítulos de livros e por último livros (SACARDO; HAYASHI, 2011). A pesquisa acadêmica-científica caracteriza-se por ser um processo que possui objetivo fundamental de produzir conhecimentos, utilizando a coleta, análise e interpretação dos dados, com base em fundamentos teóricos (MARTINS; VARANI, 2012).

Assim, foi realizado um levantamento da produção bibliográfica, utilizando o ano de conclusão do doutorado como ano inicial para sumarizar as produções dos seguintes meios de publicação: artigos em anais de congresso, artigos em periódicos, livros e capítulo de livros. Quantificando esses dados, as publicações das doutoras cadastradas nos currículos do RCDP para os artigos em anais de congresso após a conclusão do doutorado correspondem a 87.965 (70,08%). No entanto, 6.760 doutoras não possuem publicação em nenhum período de sua carreira e 30.790 doutoras possuem publicações, porém nenhuma ocorreu após a conclusão do doutorado. Já as publicações das doutoras cadastradas nos currículos do RCDP para os artigos em periódicos após a conclusão do doutorado correspondem a 101.153 (80,59%). Entretanto, 6.558 doutoras não possuem publicação em nenhum período de sua carreira e 17.804 doutoras possuem publicações, mas nenhuma ocorreu após a conclusão do doutorado.

As publicações das doutoras cadastradas nos currículos do RCDP para livros após a conclusão do doutorado correspondem a 37.572 (29,93%). No entanto, 77.563 doutoras não possuem publicação em nenhum período de sua carreira e 10.380 doutoras possuem publicações, porém nenhuma ocorreu após a conclusão do doutorado. Por fim, as publicações das doutoras cadastradas nos currículos do RCDP para capítulos de livros após

a conclusão do doutorado correspondem a 62.462 (49,76%). Entretanto, 45.832 doutoras não possuem publicação em nenhum período de sua carreira e 17.221 doutoras possuem publicações, porém nenhuma ocorreu após a conclusão do doutorado.

Esses dados nos revelam um número muito representativo de doutoras no RCDP que não possuem publicações em sua carreira e principalmente após a conclusão do doutorado. A publicação em livros possui o menor percentual e o principal meio de publicação das doutoras pesquisadoras após a conclusão do doutorado são artigos, corroborando com o estudo de Sacardo e Hayashi (2011). Vale ressaltar que, uma hipótese para o baixo número de publicação das doutoras pode ser um indicativo de um número insuficiente de bolsas de pós doutorado e de concursos para docentes, no qual as doutoras tendem a migrar para a iniciativa privada e possivelmente não vão publicar.

A pós-graduação brasileira é responsável por um número considerável de produções científicas e através das publicações dos resultados de suas pesquisas em diversos meios tem promovido mudanças no cenário da pesquisa científica proporcionando maior representatividade no cenário internacional (QUIRINO, 2018). Assim, objetivando identificar um padrão ou um período de maior produtividade das doutoras pesquisadoras, foram sumarizadas as produções por ano e calculados a média, mediana e desvio padrão para cada ano a partir do ano de conclusão do doutorado, utilizando como dados as informações das doutoras que possuem publicações após a conclusão do doutorado, (Tabela 8 / Tabela 9).

Tabela 8 – Publicações após a conclusão do doutorado por ano

Anos após conclusão	**Artigos em anais de congresso**			**Artigos em periódicos**		
	Média	**Mediana**	**Desvio padrão**	**Média**	**Mediana**	**Desvio padrão**
1º	2,06	1	3,03	1,41	1	1,47
2º	3,01	2	4,15	1,37	1	1,81
3º	3,48	2	4,75	1,44	1	1,90
4º	3,73	2	5,13	1,52	1	2,03
5º	3,99	2	5,44	1,64	1	2,21
6º	4,09	2	5,55	1,73	1	2,35
7º	4,27	2	5,82	1,85	1	2,45
8º	4,34	2	5,71	1,97	1	2,61
9º	4,47	3	5,77	2,10	1	2,76
10º	4,52	3	5,83	2,20	1	2,83
11º	4,64	3	6,00	2,33	1	2,93
12º	4,70	3	6,18	2,43	2	3,00
13º	4,73	3	6,09	2,52	2	3,06
14º	4,80	3	6,16	2,60	2	3,15
15º	4,88	3	6,19	2,71	2	3,28
16º	4,88	3	6,20	2,80	2	3,26
17º	4,92	3	6,58	2,90	2	3,47
18º	4,93	3	6,44	2,97	2	3,51
19º	4,88	3	6,30	3,03	2	3,68
20º	4,87	3	6,15	3,10	2	3,71
21º	4,93	3	6,40	3,15	2	3,71
22º	4,85	3	6,29	3,19	2	3,79

Tabela 9 – Publicações após a conclusão do doutorado por ano (continuação)

Anos após conclusão	Artigos em anais de congresso			Artigos em periódicos		
	Média	Mediana	Desvio padrão	Média	Mediana	Desvio padrão
23º	4,79	3	6,23	3,30	2	3,87
24º	4,84	3	6,41	3,26	2	3,72
25º	4,77	3	6,09	3,24	2	3,82
26º	4,68	3	6,02	3,19	2	3,50
27º	4,80	3	6,03	3,12	2	3,44
28º	4,73	3	6,48	3,24	2	3,73
29º	4,57	3	5,49	3,17	2	3,68
30º	4,62	3	5,93	3,19	2	3,88
31º	4,25	2	5,08	3,17	2	3,50
32º	5,02	3	7,01	3,32	2	3,81
33º	4,86	3	6,42	3,30	2	4,03
34º	4,66	3	6,05	3,33	2	4,14
35º	4,55	2	5,99	3,43	2	3,96
36º	4,56	3	5,52	3,36	2	4,15
37º	4,77	3	6,64	2,95	2	3,43
38º	4,62	3	5,33	3,18	2	3,86
39º	4,55	2	5,67	3,14	2	3,46
40º	4,82	2	7,21	3,03	2	3,41
41º	3,72	2	4,76	3,20	2	3,34
42º	3,16	1	3,90	2,99	2	3,36
43º	3,41	2	4,97	3,06	2	2,73
44º	3,74	1	6,47	2,77	2	2,77
45º	5,35	3	6,07	2,48	2	2,61
46º	3,43	2,5	3,25	2,39	1	3,03
47º	2,91	1	4,23	2,13	1	2,22
48º	2,83	1	3,31	2,88	2	2,83
49º	7,00	6	7,78	2,17	1	2,67
50º	3,33	1	4,93	3,22	2	2,96
51º	2,00	2	0,00	3,58	3	2,75
52º	1,50	1,5	2,12	3,70	2	4,55
53º	6,00	6	5,66	3,80	2	3,68
54º	0,00	0	0	5,00	4,5	5,22
55º	2,00	2	0	4,67	4	3,39
56º	0,00	0	0	1,33	0,5	1,75
57º	3,00	3	0	0,75	1	0,50
58º				1,67	2	1,53
59º				0,67	1	0,58
60º				1,00	1	0

Fonte: Elaboração do autor

Observando as Tabela 8 e Tabela 9, o primeiro ano após a conclusão do doutorado, referente aos artigos em anais de congresso, possui as menores média, mediana e desvio padrão. Indicando que, dentre os anos iniciais após a conclusão do doutorado o primeiro ano corresponde ao ano que as doutoras pesquisadoras são menos produtivas. Do segundo ao oitavo ano, a média é maior que a tendência central dos dados e isto deve-se ao fato que, como mais da metade do conjunto publicou entre 1 e 10 artigos, temos doutoras que publicaram mais de 80 artigos para cada ano, interferindo diretamente no grau de oscilação dos dados. A partir do nono ano, pode-se considerar um equilíbrio nas publicações das

doutoras, não tendo grandes alterações na média e desvio padrão.

Já para os artigos em periódicos, do primeiro ao décimo primeiro ano, as doutoras são menos produtivas. A tendência central igual a 2 perdura do décimo segundo ano ao quadragésimo quinto ano, no qual, a média e o desvio padrão não apresentam alterações consideráveis, demonstrando um equilíbrio de produção das doutoras no decorrer destes anos. Os dados das publicações das doutoras do RCDP, além de demostrar a diferença de produtividade entre os artigos em anais de congresso e os artigos em periódicos, nos mostram que nos anos iniciais as doutoras são menos produtivas, e alguns anos posterior ao ano de conclusão do doutorado a quantidade de publicações tende a aumentar.

O progresso significativo na atividade científica e tecnológica no Brasil acompanha a tendência mundial, em consequência do esforço aplicado em facilitar o desenvolvimento tecnológico e estimular a inovação. Porém o principal objetivo da pesquisa científica e tecnológica aparece bastante desvinculada dos problemas reais da nação e poderia ser mais aproveitada se aplicada para melhorar a qualidade de vida da sociedade, “razão pela qual ocorre uma defasagem entre o que a pesquisa científica gera e o que pode ser aproveitado socialmente” (SOARES, 2018, p. 309). O estudo da ciência atrelada ao gênero promove descobertas científicas através de novas perspectivas, objetivando aprimorar a eficiência experimental e possibilitar igualdade social, e deste modo, possa garantir que os resultados das pesquisas sejam aplicados a toda sociedade (TANNENBAUM et al., 2019).

5.6 Colaboração científica das doutoras

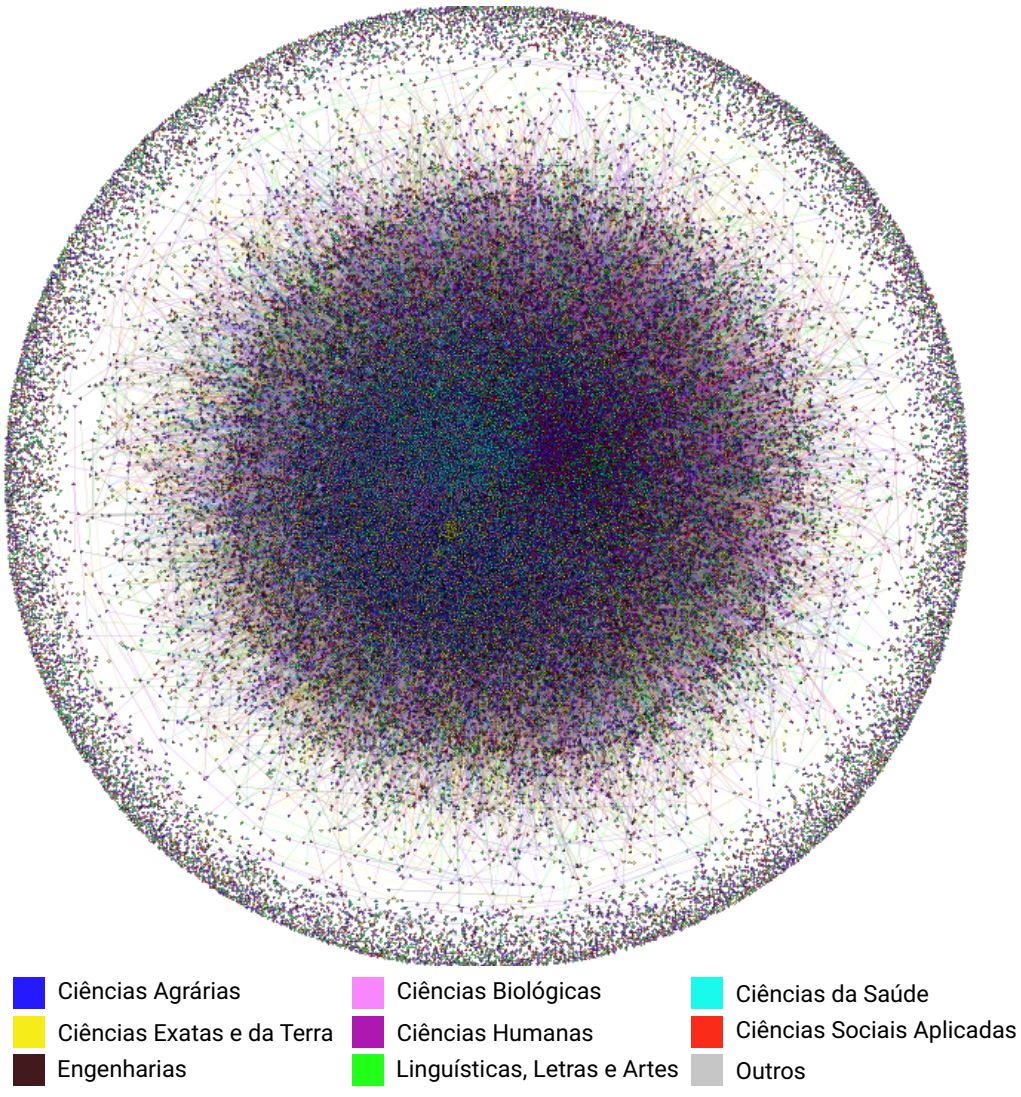
Diversos estudos utilizam dados da produção científica para analisar e entender a colaboração científica em todos os seus níveis, como entre indivíduos, instituições, países, áreas de conhecimento (SONNENWALD, 2007). Existem várias formas de identificar a relação entre dois pesquisadores, como: participação em projeto; publicação de coautoria; atividade de consultoria; participação em avaliações; bancas examinadoras; prêmios; e outras produções científicas, por exemplo, patentes (STRÖELE et al., 2011). Os pesquisadores ao se envolverem em um projeto científico colaborativo têm acesso a mais recursos materiais e informacionais, pois a colaboração científica potencializa tanto o crescimento profissional, quanto o avanço do conhecimento (HILÁRIO; GRÁCIO; GUIMARÃES, 2018).

Segundo Grácio (2018, p. 27), a coautoria pode ser considerada um produto resultante da “colaboração entre um conjunto de pesquisadores”, que interagiram no decorrer do desenvolvimento da pesquisa científica. As colaborações científicas através das coautorias, tem recebido interesse de pesquisadores, entidades avaliadoras e de fomento em C&T (BALANCIERI, 2004), pois, além de apresentar a qualidade de produções bibliográficas, as coautorias permitem visualizar a estrutura e a dinâmica das colaborações entre os pesquisadores. A colaboração científica realizada entre pesquisadores e publicações é

representada pela rede, no qual os vértices correspondem aos autores das publicações e as arestas aos relacionamentos entre os autores, ou seja, a relação de coautoria.

Conforme apresentado na Tabela 4 (Seção 5.2), as doutoras pesquisadoras possuem um número considerável de produções bibliográficas e técnicas, bem como, demonstrando esses dados na Figura 11 (Seção 5.3), foi possível apresentar os dois principais meios de divulgação de conteúdo por grandes áreas. Como cada produção possui um ou mais autores, uma mesma doutora pode participar de outras produções, formando então, as conexões entre autores. Assim, para criar a rede de todas as doutoras pesquisadoras (Figura 18), dois arquivos foram recuperados do RCDP, o primeiro contém os nós que correspondem às doutoras e o segundo contém as arestas que correspondem às conexões entre os autores, para cada produção de artigos em anais de congresso e artigos em periódicos.

Figura 18 – Rede geral das doutoras contendo todos os nós

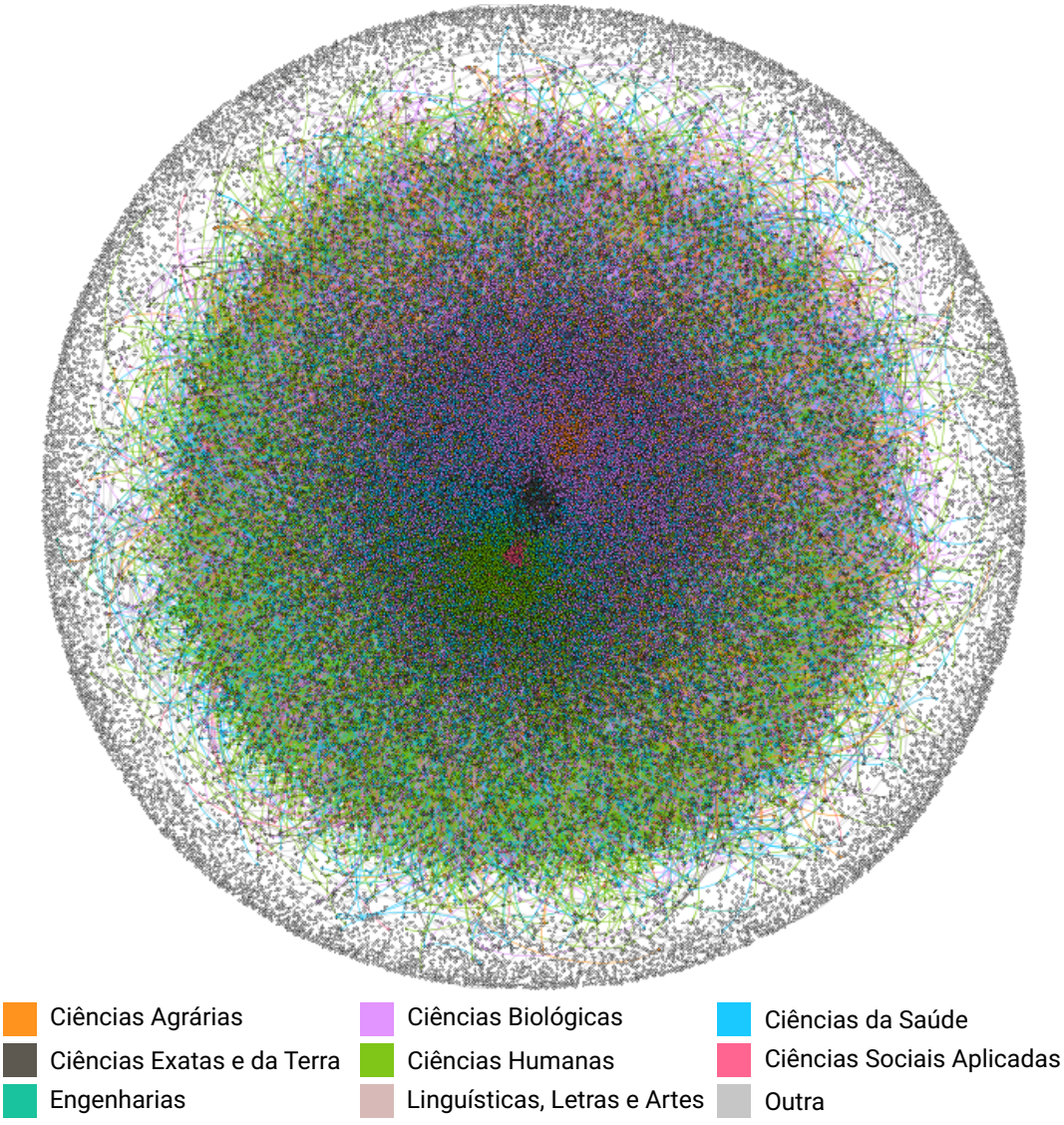


Fonte: Elaboração do autor

Como é possível observar na Figura 18, a rede possui uma concentração dos nós na componente gigante (centro da rede) em que esta concentração se caracteriza pelas colaborações, em geral de doutoras que atuam em uma mesma área de atuação ou nas mesmas instituições de pesquisa. Além disso, também é possível perceber nas bordas da rede um conjunto de nós com poucas conexões ou isolados, resultando em que as doutoras pesquisadoras não tem trabalhado em colaboração com outras no conjunto analisado.

É possível ainda, analisar somente a componente gigante da rede, conforme apresentado na Figura 19. Nesta rede, são excluídos todos os nós que não estão vinculados a maior componente identificada na rede.

Figura 19 – Componente gigante da rede contendo todos os nós

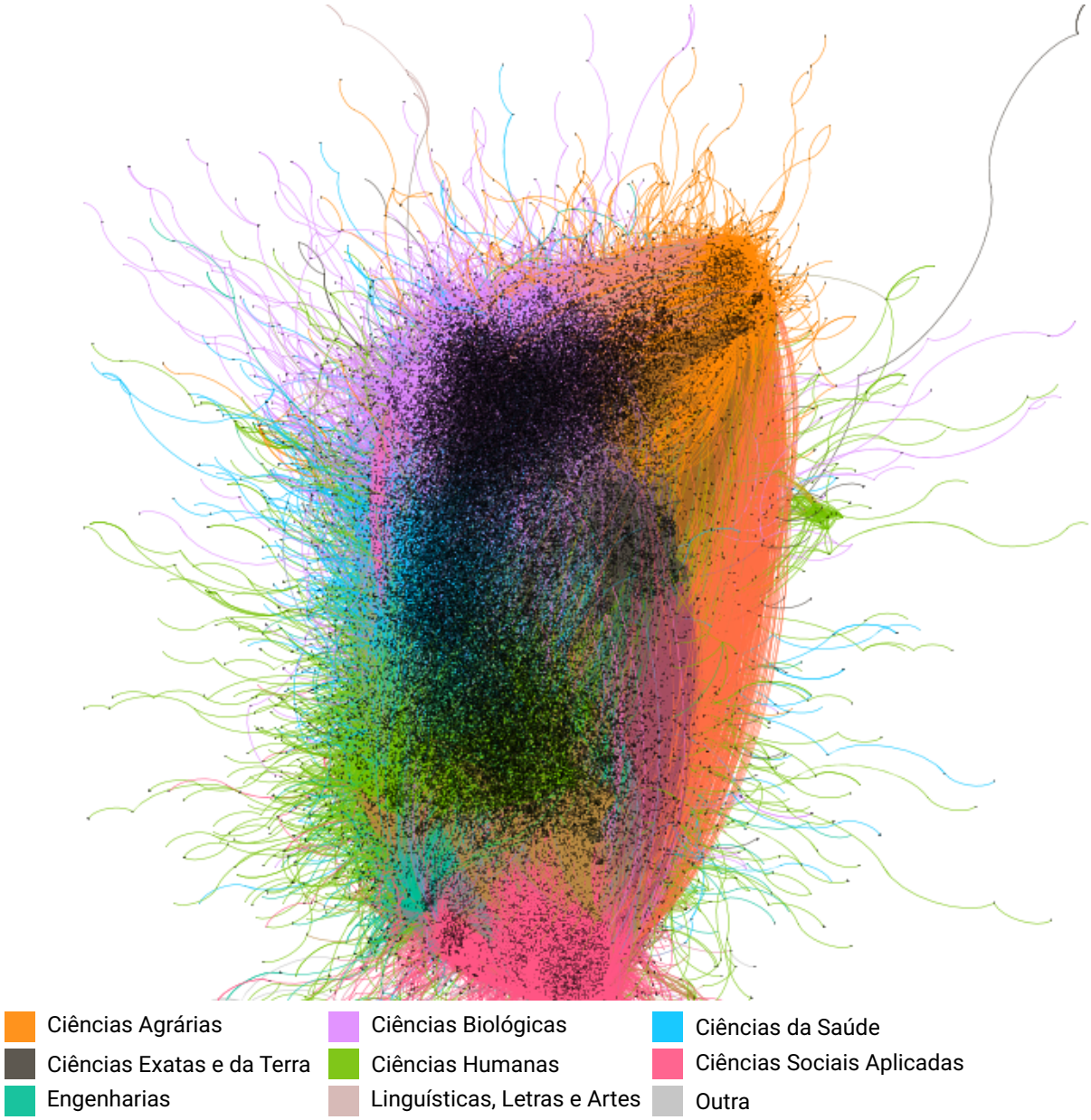


Fonte: Elaboração do autor

Como pode ser observado na Figura 19, na componente gigante identificada, ainda é possível observar alguns nós com uma menor quantidade de conexões nas bordas da rede,

bem como, uma maior concentração no núcleo, resultado de nós altamente conectados. É possível perceber ainda, uma maior concentração de nós de algumas grandes áreas. Tais nós, influenciado por algumas de suas características e consequentemente pelas suas colaborações, tendem a formar núcleos. Então complementando a Figura 19, é possível caracterizar as grandes áreas por *clusters* e assim identificar a concentração de nós por grandes áreas, conforme apresentado na Figura 20.

Figura 20 – *Clusters* identificados



Fonte: Elaboração do autor

Ao analisar os *clusters* identificados na Figura 20, percebe-se que eles tendem a concentrar doutoras de uma mesma área, e que alguns destes *clusters* possuem uma maior densidade. Destaca-se os *clusters* das grandes áreas de Ciências Exatas e da

Terra, Ciências Biológicas e Ciências Sociais Aplicadas. Um estudo considerando métricas de análise de redes é apresentado na próxima seção, em que, o conjunto de doutoras pesquisadoras que possuem bolsa de produtividade em pesquisa do CNPq é analisado sob diversas perspectivas.

Gephi foi a ferramenta escolhida e utilizada para gerar as redes e para analisar os dados, sendo possível através dela explorar, entender e visualizar todos os tipos de grafos e redes. Ao utilizar esta ferramenta é possível escolher diversos algoritmos de *layout* que permitem produzir uma imagem de fácil visualização e experiência para o usuário. Também é possível gerar diversas métricas na estrutura de estatísticas, que podem ser utilizadas para análise de redes sociais e redes sem escala, como centralidade de intermediação, proximidade, diâmetro, coeficiente de agrupamento, dentre outras. Mais uma opção disponibilizada pela ferramenta corresponde aos filtros, utilizados para selecionar nós e/ou arestas com base na estrutura ou nos dados da rede. Quanto a edição e visualização dos dados, o Gephi permite visualizar e editar todos os dados importados através da tabela de dados. Assim, nas redes geradas, quanto a aparência, foi configurado as cores dos nós por grandes áreas de atuação das doutoras. Para caracterizar as redes foram adotados métodos que posicionam os nós mais conectados no centro da rede e os nós menos conectados são deslocados para as extremidades.

A colaboração científica tem grande importância no cenário científico nacional e internacional, pois favorece o crescimento e o desenvolvimento dos estudos da ciência em diversas áreas do conhecimento. Neste cenário, “os indicadores bibliométricos de colaboração são utilizados para medir a produção científica realizada em colaboração e caracterizar as estruturas da cooperação existentes na ciência” (SANTIN; VANZ; CAREGNATO, 2019, p. 195). Podem ser utilizados diversos métodos e indicadores no estudo da colaboração científica, como também podem envolver aspectos que vão desde a contagem básica de autores, até representação utilizando estatísticas mais complexas. Assim, analisar a perspectiva da produção científica e compreender como as doutoras pesquisadoras contribuem para o crescimento científico nacional e internacional é um importante aspecto da análise da participação científica feminina.

5.7 Produção bibliográfica das bolsistas de produtividade em pesquisa do CNPq

As bolsas de produtividade em pesquisa do CNPq têm a finalidade de valorizar a produção bibliográfica e científica do pesquisador de excelência, que se destaca entre seus pares, de acordo com critérios estabelecidos pelo CNPq (normativos) e pelos Comitês de Assessoramento do CNPq - CAs (específicos). Os pesquisadores são classificados em categorias e níveis. Para a categoria 1, o pesquisador será enquadrado em quatro diferentes

níveis (A, B, C ou D), tendo como base comparativa seus pares e nos dados dos últimos 10 (dez) anos, que demonstre capacidade de formação contínua de recursos humanos. Já para a categoria 2 não há especificação de nível, a avaliação é realizada através da produtividade do pesquisador, no qual são valorizados os trabalhos publicados e orientações, ambos referentes aos últimos 5 (cinco) anos. A bolsa tem duração de acordo com sua categoria ou nível: 1A de 60 (sessenta) meses de duração; 1B, 1C e 1D com 48 (quarenta e oito) meses; e categoria 2 possui 36 (trinta e seis meses) de duração (CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO - CNPq, 2020).

O arcabouço LattesDataXplorer gerou um arquivo contendo um total de 14.475 registros de pesquisadoras que possuem bolsas de produtividade no CNPq dividido entre as categorias/níveis. A partir dessa lista, foi recuperado do RCDP todas as doutoras que são bolsistas de produtividade em pesquisa e obteve-se um total de 5.085 (35,13%), conforme apresentado na Tabela 10. É importante ressaltar que, os valores apresentados na coluna Total RCDP correspondem a todas as doutoras que participam ou participaram em algum momento de um grupo de pesquisa do DGPB.

Tabela 10 – Doutoras bolsistas de produtividade em pesquisa do CNPq

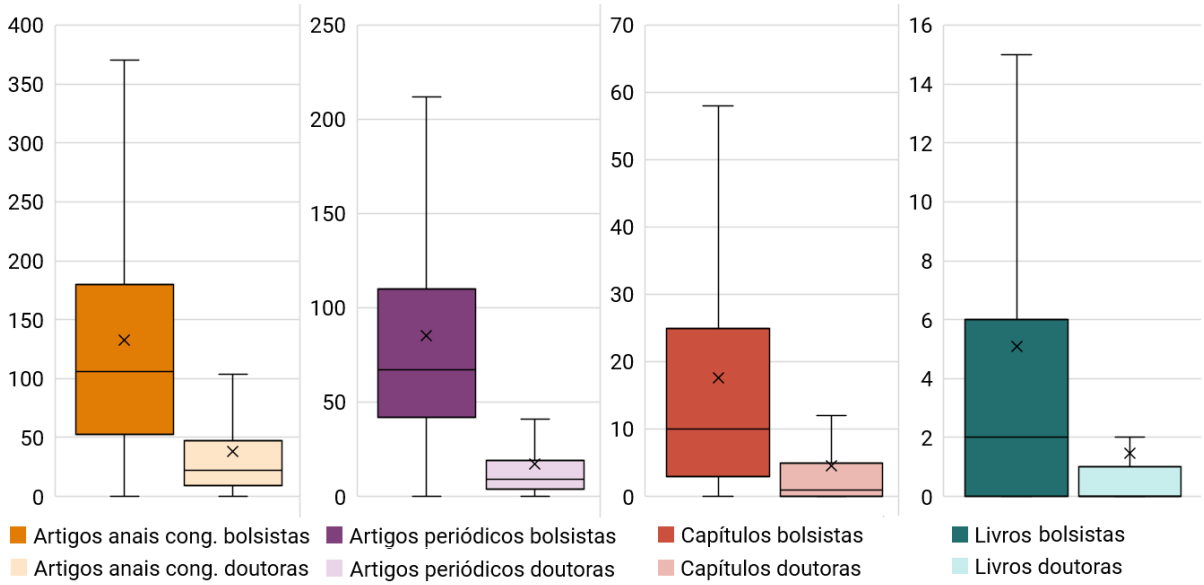
Categoria/Nível	Total Geral	Total RCDP
1A	1.180	300
1B	1.267	392
1C	1.432	508
1D	2.425	842
2	8.171	3.043

Fonte: Elaboração do autor

Os valores apresentados na Tabela 10 demonstram uma diferença considerável entre o total geral e o total RCDP. Uma hipótese para explicar essa divergência refere-se a limitação desta pesquisa, em que, não são todas as doutoras que estão inseridas em um grupo de pesquisa. Logo, pode ser que estes totais não contemplam todas as mulheres que são bolsistas.

De acordo com estudo realizado por Gonçalves (2020), um de seus resultados apontam que a probabilidade de publicações realizadas pelos pesquisadores bolsistas é superior em 23 pontos percentuais aos pesquisadores não-bolsistas. Neste contexto, como as doutoras bolsistas possuem um número considerável de publicações, sendo 14% para artigos em anais de congresso, 20% para artigos em periódicos, 14% para livros e 15,7% para capítulos de livros, realizar um comparativo entre as produções delas com as de todas as doutoras é relevante. Esses percentuais levam em consideração o total de produções do RCDP, e assim, a Figura 21 apresenta a produção bibliográfica das bolsistas e das doutoras. Os dados foram apresentados através do *boxplot* e para melhorar a visualização, os *outliers* foram ocultados por causa dos inúmeros valores discrepantes.

Figura 21 – Produção bibliográfica das bolsistas e das doutoras



Fonte: Elaboração do autor

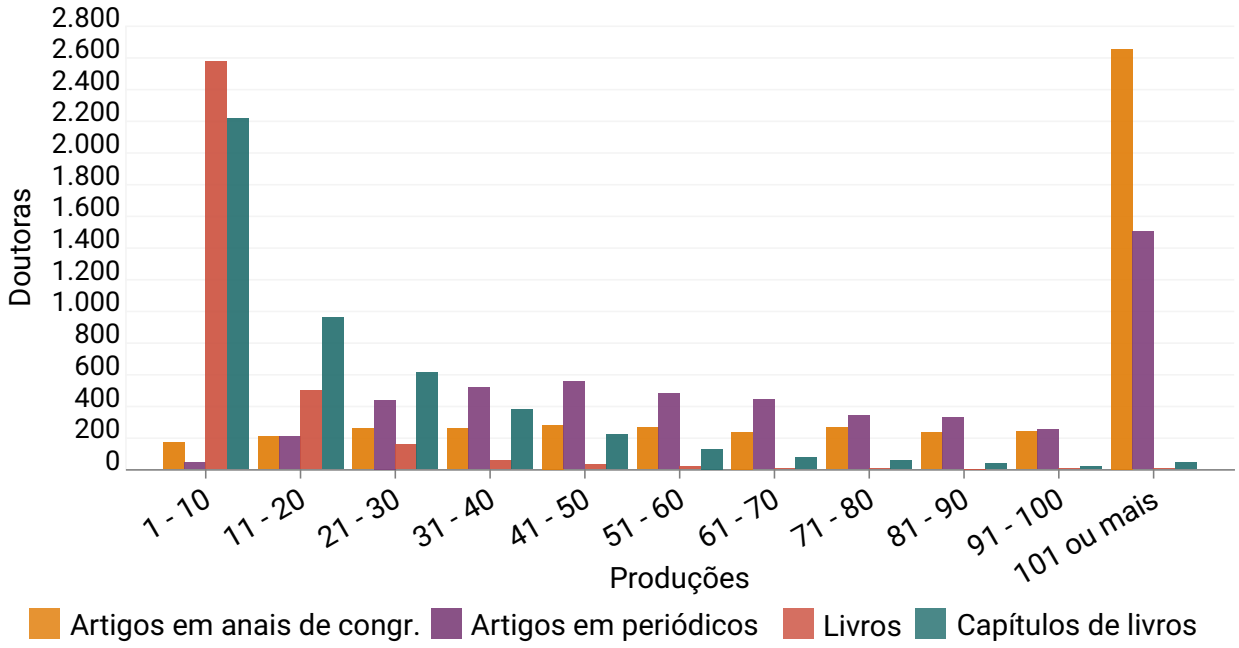
Um tópico importante na Figura 21 corresponde a mediana, como a mesma não possui influência dos *outliers*, todas as médias apresentam valores maiores que a mediana. Para os artigos em anais de congresso das bolsistas, a mediana corresponde a 106, tendo a amplitude variando entre 0 e 370 artigos, já os valores centrais dos dados compreendem 53 a 180 artigos. Comparando aos dados de todas as doutoras podemos verificar que a mediana apresentou um valor bem inferior, correspondendo a 22, amplitude menor e os limites do primeiro e terceiro quartil corresponde a 9 e 47 respectivamente. É muito representativo a diferença entre as medianas, tendo como o limite superior das doutoras abaixo da mediana das bolsistas. Já para os artigos em periódicos, é muito representativo as informações quanto aos valores centrais dos dados e a amplitude das doutoras que estão abaixo do primeiro quartil das bolsistas. A mediana das bolsistas corresponde a 67, com valores centrais entre 42 a 110, e amplitude de 0 a 212.

Já para livros, tanto para as bolsistas quanto para as doutoras, apresentam o limite inferior zero, este fato é explicado pelo número considerável de doutoras (64,28%) e bolsistas (34,43%) que não possuem nenhuma publicação em livros. Já a mediana das bolsistas corresponde a 2, com limite superior de 6. Por fim, para capítulo de livros, as bolsistas apresentam mediana de 10, variando entre 3 a 25 capítulos de livros publicados. As bolsistas possuem números consideravelmente maiores se comparado ao das doutoras.

Complementando a análise da Figura 21, foi realizado um levantamento destas produções e distribuídas por conjuntos, possibilitando comparar a produção de todas as doutoras com as bolsistas. Assim, a Figura 22 apresenta o eixo vertical contendo o total de bolsistas e o eixo horizontal os conjuntos com intervalo de 10 anos para cada uma das

produções (artigos em anais de congresso, artigo em periódicos, livros, capítulos de livros). O último conjunto contém todas as bolsistas com produção igual ou superior a 101.

Figura 22 – Quantitativo das produções bibliográficas das bolsistas por conjuntos



Fonte: Elaboração do autor

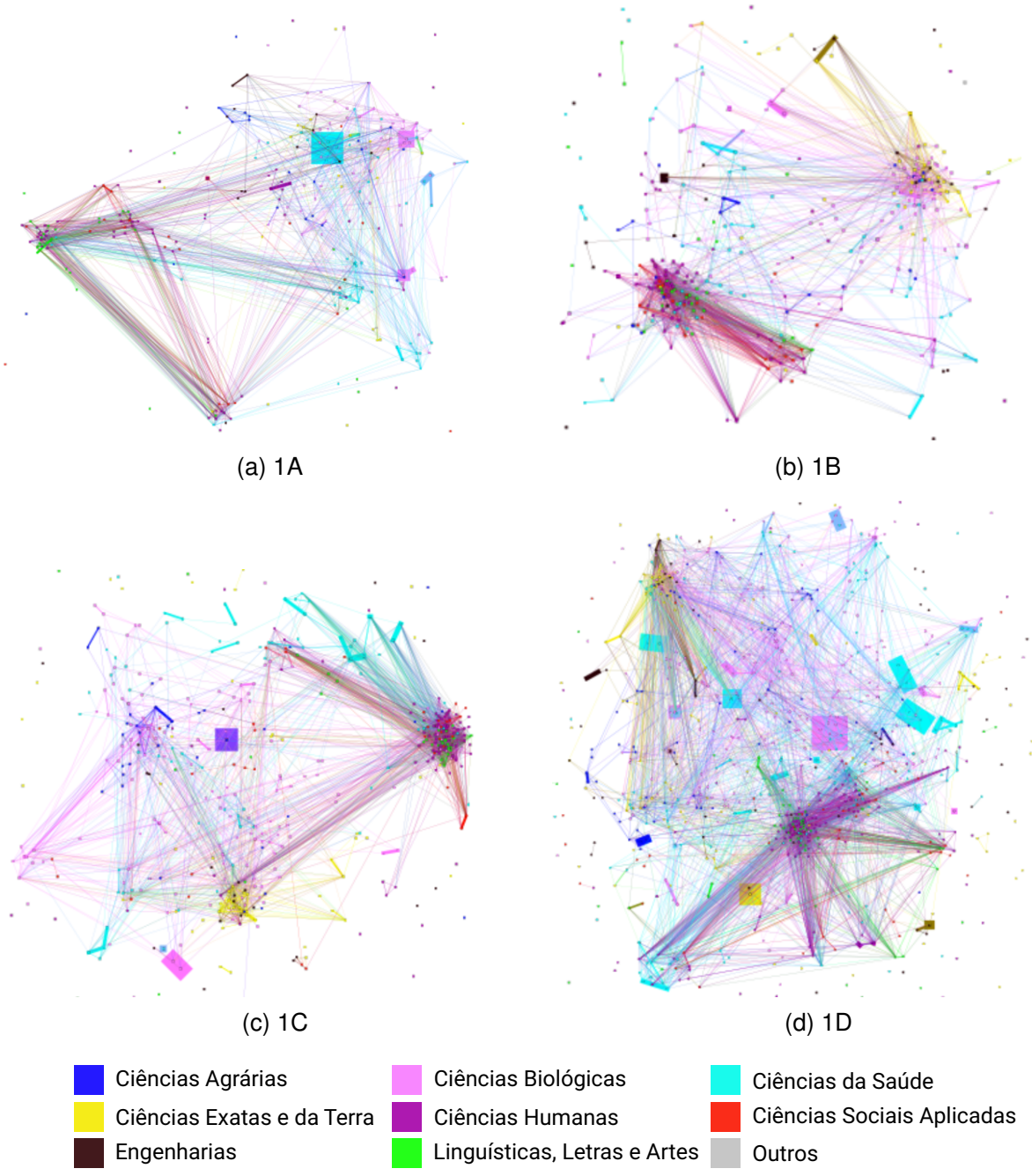
Verificando o quantitativo das produções das bolsistas por conjunto da Figura 22, as publicações de artigos em anais de congresso nos três primeiros conjuntos (1 a 30 artigos) têm um percentual baixo de 12,37%, enquanto que na Figura 9 (Seção 5.2) as doutoras pesquisadoras publicaram no mesmo conjunto 55,79%. Porém, ao compararmos o percentual das bolsistas quanto aos artigos publicados no conjunto de 101 ou mais, o valor corresponde a 52,17%, enquanto que das doutoras para o mesmo conjunto corresponde a 8,34%, ou seja, o total de produções de artigos em anais de congresso de cada bolsista é consideravelmente superior ao de cada doutora. Este mesmo tópico pode ser observado para os artigos em periódicos, ou seja, as bolsistas também possuem uma quantidade consideravelmente superior de artigos publicados, no qual para o conjunto de 101 ou mais das bolsistas o percentual corresponde a 29,49%, enquanto que para as doutoras corresponde a 1,99%.

Ainda observando a Figura 22, para os livros e capítulos de livros, a maior concentração de produções corresponde ao primeiro conjunto tanto para bolsistas quanto para todas as doutoras. Neste tópico é relevante comparar o percentual de produções, no qual as bolsistas também apresentam quantidade superior para as duas produções. As bolsistas possuem percentual de livros igual a 50,60% e capítulos de livros igual a 43,54%, enquanto que as doutoras apresentam 35,53% e 51,48% respectivamente. É importante ressaltar que, das 13 doutoras hiper produtivas que possuem cada uma mais de 500 artigos em

periódicos, 10 delas são bolsistas. Já para artigos em anais de congresso, das 17 doutoras hiper produtivas que possuem cada uma mais de 750 artigos, 8 delas são bolsistas.

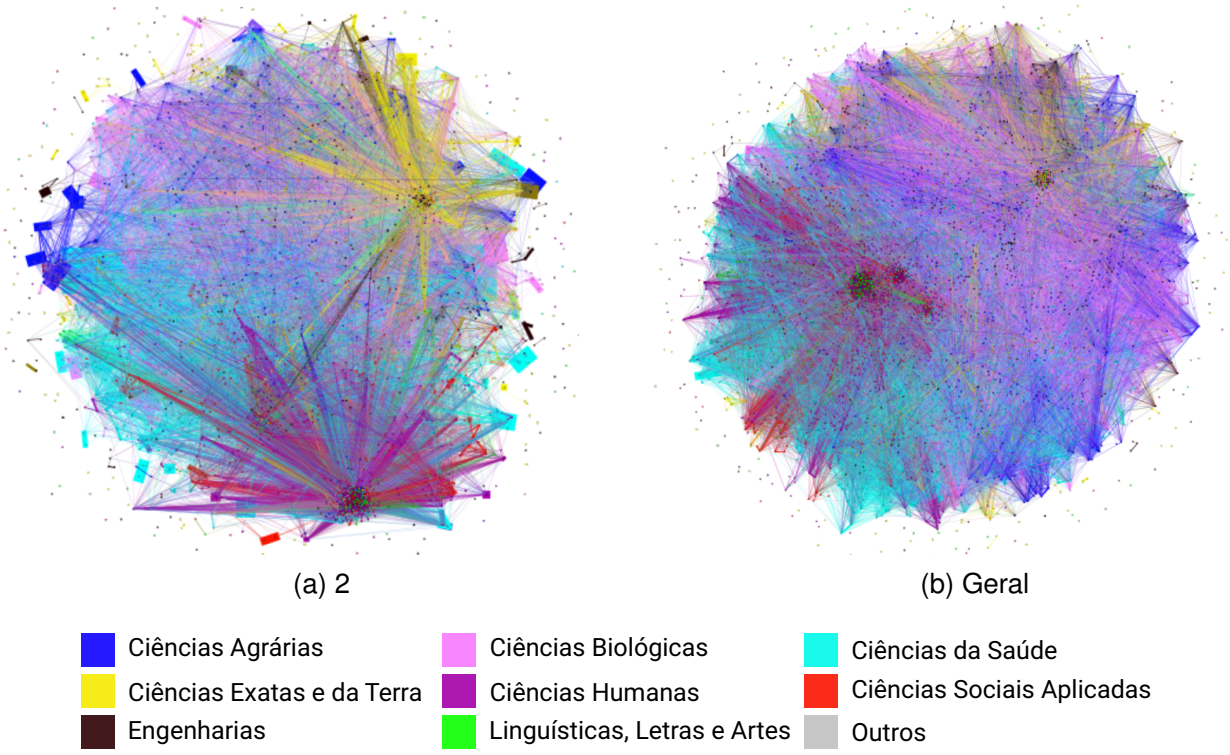
Outro aspecto relevante na análise das bolsistas corresponde as redes de colaboração. Neste contexto, as redes com todos os nós e arestas para cada uma das categorias/níveis de todas as bolsistas de produtividade em pesquisa do CNPq recuperados do conjunto RCDP, são apresentadas nas Figura 23 e Figura 24.

Figura 23 – Redes de colaboração das bolsistas por categorias/níveis



Fonte: Elaboração do autor

Figura 24 – Redes de colaboração das bolsistas por categorias/níveis (continuação)



(a) 2 (b) Geral

Fonte: Elaboração do autor

Para gerar cada uma das redes, os arquivos de nós e arestas respectivos de cada categoria/nível foram importados no Gephi e posteriormente realizadas as configurações. Quanto a aparência para cada uma das redes, selecionou-se e padronizou as cores dos nós por grandes áreas, no qual foi possível diferenciar cada grande área pela sua cor específica. Quanto ao *layout*, foi aplicado a distribuição OpenOrd com parâmetros padrões. Para melhorar a visualização, alterou-se os parâmetros de espessura das arestas para 0,5 e desabilitou a opção curvo.

Analisar a colaboração científica entre os autores das produções e o relacionamento entre eles é um processo que pode ser evidenciado através da caracterização das redes de colaboração. As investigações acerca da coautoria são utilizadas para medir a força da colaboração científica entre os indivíduos quantificando as publicações (EOM; FORTUNATO, 2011), bem como analisar fatores, comportamentos, desempenhos e motivações nas colaborações científicas. Assim, é possível observar nas Figura 23 e Figura 24 a diferença das redes entre cada uma das categorias/níveis, bem como as aglomerações e conexões por cada grande área. Neste contexto, as estatísticas destas redes caracterizadas, bem como as métricas adotadas em todos níveis/categorias analisados, são apresentadas na Tabela 11, no qual é possível observar numericamente e compreender a diferença entre os níveis das doutoras que possuem bolsa de produtividade pelo CNPq.

Tabela 11 – Métricas adotadas

Métricas	**1A**	**1B**	**1C**	**1D**	**2**	**Geral**
Total de nós	300	392	508	842	3.043	5.085
Total de arestas	1.575	1.927	2.850	5.399	57.072	195.570
Grau médio dos nós	10,5	9,83	11,22	12,82	37,51	76,37
Grau ponderado médio	37,95	29,42	38,78	53,23	107,56	261,93
Diâmetro da rede	7	9	10	10	12	9
Densidade do grafo	0,035	0,025	0,022	0,015	0,12	0,015
Modularidade	0,72	0,64	0,69	0,78	0,64	0,65
Coeficiente de *clustering* médio	0,43	0,47	0,47	0,47	0,45	0,44
Componentes conectados	40	53	78	115	207	169

Fonte: Elaboração do autor

Observando a Tabela 11, o total de nós aumenta entre cada uma das categorias/níveis hierárquico, conforme esperado. Já o total de arestas para 1A, 1B, 1C e 1D correspondem a aproximadamente cinco vezes o total para cada um de seus nós respectivos. Enquanto que na categoria 2, o total de aresta é bem maior que nos outros níveis, representa mais de 18 vezes o total de nós.

O grau médio dos nós define o peso dos nós de acordo com a quantidade de suas conexões, ou seja, número de conexões que, em média, os nós de uma rede possuem. Já o grau ponderado médio é similar ao grau médio, porém, para a sua medida, utiliza-se dos pesos das arestas em seu algoritmo para então definir o peso dos nós. Logo, não avalia apenas se há conexão entre os nós, mas também a intensidade dessas conexões. Exemplificando, se a doutora A possui 50 conexões com a doutora B, e a doutora C possui 50 conexões com 10 doutoras diferentes, A e C possuem o mesmo grau, porem a doutora C possui um valor maior em seu grau ponderado médio, por conta da diversidade de suas relações. Assim, observando a Tabela 11, é possível perceber a diferença entre eles. Para o grau médio os valores são inferiores aos valores do grau ponderado médio, pois o conjunto analisado possui uma grande diversidade entre seus relacionamentos.

Na Tabela 11 também é possível verificar o diâmetro da rede que está entre 7 e 12. O diâmetro da rede mede a distância média do grafo entre todos os pares de nós, ou seja, maior distância existente entre dois nós da rede. Como o diâmetro é a maior distância entre dois nós quaisquer da rede, ou seja, o quão separados estão os dois nós mais distantes, essa medida é útil para comparar redes e avaliar, em uma certa medida, a maior distância que seus nós precisariam percorrer para se conectarem. Comparando com o diâmetro de todos os doutores do trabalho de Dias (2016), o período de 2010 a 2014 possui valor 15, enquanto que as bolsistas possui a maior distância de 9 passos necessários para se conectarem.

Densidade do grafo mostra a quão próxima a rede está de ser completa, ou seja, mostra a taxa de quantas conexões existem no grafo em relação a todas as conexões

possíveis, considerando se todos os nós estiverem ligados a todos diretamente. Um grafo completo é aquele em que todos os seus componentes se encontram conectados, ou seja, para cada nó, existe uma aresta conectando esse nó aos demais. Assim, observando a Tabela 11, para todas as categorias/níveis os valores estão próximos de 0 conforme esperado. A rede é incompleta, pois a possibilidade de um nó ligar a todos os outros nós é mínima, logo, quanto maior a rede, mais baixo este valor será.

A modularidade define os *clusters* da rede de acordo com a força de suas conexões, ou seja, o algoritmo vai realizar um cálculo que detecta as comunidades. Assim, é uma medida que mostra o quão bem uma rede poderia ser decomposta em comunidades modulares. Observando a Tabela 11, percebe-se que os valores de modularidade são baixos, ou seja, um pequeno valor de modularidade indica uma rede com uma estrutura de comunidades internas menos sofisticadas.

Por fim, o coeficiente de *clustering* médio determina o coeficiente médio em que os nós de uma rede aglomeram entre si, ou seja, estão mais ou menos conectados formando grupos de conexões diretas. Todas as categorias/níveis da Tabela 11 possuem valores próximos uns aos outros, mostrando o grau de articulação existente entre seus nós. Já o componente conectado pode ser entendido como um conjunto de nós no qual há pelo menos um caminho de conexão entre todos eles. Assim, na Tabela 11, o número de componentes é pequeno para todas as categorias/níveis.

A relação da mulher na ciência é um tópico que vem sendo estudado através de várias abordagens diferentes e são vários aspectos que podem ser observados ao analisar a rede de colaboração das doutoras pesquisadoras. Os estudos que tratam da colaboração científica evidenciam a grande importância deste tema para o cenário científico nacional e internacional, além de favorecer o crescimento e o desenvolvimento dos estudos da ciência em diversas áreas do conhecimento. Além disso, como a ciência é um campo neutro para as questões de gênero, esses estudos contribuem para melhorar a visão do atual cenário que a mulher se encontra na ciência. Segundo Liu, Song e Yang (2020), a participação feminina na pesquisa traz frutos interessantes para o campo científico e se os países trabalharem juntos para adotar políticas apoiando as famílias e fornecendo mais oportunidades para as mulheres, elas serão fortalecidas e terão sua autoestima aumentada, proporcionando equilíbrio nos campos da pesquisa científica.

6 Conclusão

A crescente participação das mulheres na ciência tem proporcionado diversos estudos que abordam esta temática e os diversos aspectos da participação delas na área científica e tecnológica. Este trabalho abordou a perspectiva da ciência da informação, expondo a participação científica e tecnológica das doutoras pesquisadoras e apresentando os aspectos de sua trajetória acadêmica no âmbito da ciência e tecnologia. Os dados utilizados foram provenientes dos currículos cadastrados na Plataforma Lattes, sendo os mesmos selecionados pelo critério de formação acadêmica/titulação doutorado concluído, como também utilizados os dados do Diretório dos Grupos de Pesquisa no Brasil para selecionar estes currículos através do gênero feminino. Desta forma, após a coleta e seleção dos dados, foi possível preparar e analisar os 125.515 currículos do conjunto selecionado. É importante destacar que o conjunto analisado possui um único currículo cadastrado para cada indivíduo, pois o mesmo vincula o login ao CPF do proprietário do currículo.

Os dados foram agrupados objetivando traçar o perfil de orientações das doutoras, analisar a representatividade da produção bibliográfica e técnica, fazer uma distribuição das doutoras em suas áreas de atuação, compreender a evolução da formação acadêmica, investigar as suas publicações após a conclusão do doutorado, analisar como ocorre a colaboração científica do conjunto selecionado, realizar uma análise daquelas que possuem bolsas de produtividade em pesquisa do CNPq, sendo possível assim realizar diversas análises de forma temporal.

As orientações acadêmicas correspondem a um tópico relevante pois ocorrem em diferentes níveis de formação como, nos cursos técnicos de nível médio, graduação, iniciação científica, especialização, residência, mestrado e doutorado. Os doutores além de serem responsáveis pela formação dos alunos nos maiores níveis de capacitação, no decorrer de sua trajetória acadêmica eles realizam as orientações dos alunos em seus diferentes níveis. Assim, as orientações foram classificadas em dois tipos, as que foram finalizadas que correspondem as orientações concluídas e as que estão em desenvolvimento que correspondem as orientações em andamento. As orientações se caracterizam como um processo interativo entre o orientador e o orientando a fim de construir o conhecimento, que resulta em benefício para ambos e abrange diversas dimensões como sociais e afetivas, além da produtiva. Logo, ao apresentar o perfil das orientações das doutoras pesquisadoras, foi possível extrair as seguintes conclusões: quanto ao quantitativo das orientações concluídas e das orientações em andamento para cada nível de formação acadêmica/titulação, foi possível perceber que no decorrer de suas carreiras, as doutoras tendem a orientar mais alunos que possuem nível mais alto de formação e diminuem a orientação dos alunos que possuem o nível de formação mais baixo; quanto a distribuição

das bolsas de estudo para as formações acadêmicas/titulações, foi notável o percentual de queda no total de bolsas comparando as orientações concluídas com as orientações em andamento, principalmente para mestrado e doutorado; quanto as orientações concluídas e orientações em andamento para todos os níveis de formação acadêmica/titulação, as grandes áreas mais representativas se intercalam entre Ciências Biológicas, Ciências Humanas e Ciências da Saúde; quanto ao percentual de crescimento das orientações concluídas no decorrer de dez anos para todas as formações acadêmicas/titulações, foi possível perceber o crescimento efetivo nos anos de 2010 e 2011, enquanto que nos outros anos, ao menos um ano de alguma formação apresentou uma redução; quanto a sumarização das orientadoras e coorientadoras, as doutoras possuem uma representação importante como orientadoras principais nos níveis de mestrado e doutorado.

A produção científica é o resultado da investigação científica e se enquadra em uma das diversas atividades de pesquisa do docente-pesquisador. Como os doutores são responsáveis pela maioria das produções científicas cadastradas nos currículos da Plataforma Lattes, este é um tópico expressivo para o estudo. Logo, ao apresentar a produção bibliográfica e técnica das doutoras pesquisadoras, foi possível extrair as seguintes conclusões: quanto a produção de conteúdo, os que possuem maior representatividade são artigos em periódicos e artigos em anais de congresso; quanto a sumarização das produções bibliográficas e técnicas por ano, o crescimento dos artigos em anais de congresso foi expressivo, tendo ápice no ano de 2011 e queda significativa após esse ano, enquanto que para artigos em periódicos, o mesmo apresentou crescimento até o ano 2017 com queda após este ano e para os outros tipos de produção científica não houve alterações consideráveis; quanto a distribuição das produções por conjunto, mais da metade das doutoras possuem um número baixo de produções, entre 1 a 10 artigos, para os artigos em periódicos e para os artigos em anais de congresso, enquanto que para o conjunto que compreende 101 ou mais, um pequeno número de doutoras possuem uma quantidade significativa de produções; quanto a atualização dos currículos, quanto mais produções científicas a doutora possuir, mais ativa e maior será o percentual de atualização de seu currículo.

As áreas de atuação dos indivíduos são cadastradas na Plataforma Lattes utilizando a classificação de áreas e subáreas do CNPq. A grande área cadastrada que possui a maior ocorrência dentre os registros em cada currículo foi utilizada para realizar uma distribuição das doutoras em suas grandes áreas de atuação. Logo, ao distribuir as doutoras pesquisadoras em suas grandes áreas, foi possível extrair as seguintes conclusões: quanto ao quantitativo por grandes áreas, a área que possui mais doutoras corresponde a Ciências Humanas, seguido de Ciências da Saúde e Ciências Biológicas; quanto a sumarização de artigos em periódicos e em anais de congresso por grandes áreas, a área que possui mais produções corresponde a Ciências da Saúde, seguido de Ciências Biológicas e Ciências Humanas; quanto a sumarização de artigos em periódicos e em anais de congresso por grandes áreas per capita, é possível visualizar a produção sem influência do número total de

doutoras de algumas áreas e apresenta um cenário mais interessante, no qual a área que possui mais produções corresponde a Ciências Agrárias, seguido de Ciências da Saúde, Engenharias, Ciências Biológicas, Ciências Exatas e da Terra, Ciências Humanas, Ciências Sociais Aplicadas, Linguística, Letras e Artes, e Outra.

Nos currículos de cada indivíduo são informados os registros de suas trajetórias acadêmicas e profissionais. Para todas as formações acadêmicas/titulações foram recuperados do RCDP os dados referentes ao ano inicial e ano final de cada doutora para cada uma de suas titulações registradas, e posteriormente, realizado o cálculo do tempo de conclusão para cada formação acadêmica, através da fórmula (AnoFinal-AnoInicial)+1. Logo, utilizando esses dados foi possível compreender a evolução da formação acadêmica das doutoras e extrair as seguintes conclusões: quanto ao tempo de conclusão das formações acadêmicas, a graduação possui uma variação maior, enquanto que mestrado e doutorado possuem pouca variação dos dados; quanto ao tempo de formação acadêmica por grandes áreas do conhecimento, as que possuem as maiores médias de formação para o mestrado são Ciências Humanas, Linguística, Letras e Artes, Ciências Sociais Aplicadas e Ciências da Saúde, e para ao doutorados são Ciências Exatas e da Terra, Engenharias, Linguística, Letras e Artes, Ciências Humanas; quanto a evolução das grandes áreas por períodos, Ciências Humanas foi a área que permaneceu em primeiro lugar quanto ao total de indivíduos formados para todos os períodos no mestrado, enquanto que no doutorado, Ciências Humanas foi a área com mais formações para $3/4$ dos períodos, e a taxa de crescimento entre períodos demonstrou como as doutoras pesquisadoras ampliaram suas áreas de formação.

Os níveis de formação acadêmica mestrado e doutorado fazem parte da pós-graduação *stricto sensu* e possuem, dentre vários propósitos, estimular e produzir conhecimento científico. A pesquisa científica está atrelada a formação acadêmica e ao desenvolvimento de projetos na pós-graduação, sendo ela um requisito básico para produção científica dos pesquisadores. Assim, para investigar as publicações das doutoras após a conclusão do doutorado, foi realizado um levantamento das produções bibliográficas utilizando o ano de conclusão do doutorado como ponto inicial para sumarizar as publicações do conjunto e identificar qual o período que as mesmas tendem a ser mais produtivas. Logo, foi possível extrair as seguintes conclusões: quanto as produções após a conclusão do doutorado, um número muito representativo de doutoras pesquisadoras não possuem publicações em sua carreira e principalmente após a conclusão do doutorado; quanto as publicações de artigos em anais de congresso e artigos em periódicos, nos anos iniciais após a conclusão do doutorado as doutoras são menos produtivas e com o passar dos anos a média de produções tende a aumentar.

Os dados sobre a produção científica têm sido muito utilizados para analisar e entender as colaborações científicas através da compreensão sobre como colaboram:

indivíduos, instituições, países e áreas de conhecimento. A colaboração pode ser entendida como uma cooperação entre os autores para realizar uma atividade e transfigura-se como uma relação que pode ser de participação em projeto, de publicação de coautoria, de consultoria, de avaliação, de bancas examinadoras, dentre outras. A relação adotada para analisar a colaboração científica das doutoras corresponde a publicação de artigos, sumarizando todas as produções bibliográficas de cada doutora, no qual os nós referem-se as doutoras e as arestas representam as todas as conexões que cada uma possui com outros autores para cada produção bibliográfica. Logo, foi possível caracterizar as redes e extrair as seguintes conclusões: quanto a rede geral das doutoras que contém todos os nós, é possível visualizar a concentração dos nós no centro da rede onde se concentram as colaborações e uma menor quantidade de nós nas bordas da rede; quanto a rede geral baseada em *clusters*, destacam-se os agrupamentos das grandes áreas de Ciências Exatas e da Terra, Ciências Biológicas e Ciências Sociais Aplicadas.

As bolsas de produtividade em pesquisa do CNPq são destinadas aos pesquisadores de excelência e tem como propósito valorizar sua produção científica. Para verificar a produção bibliográfica das pesquisadoras que possuem bolsas de produtividade em pesquisa do CNPq, foram selecionadas as bolsistas e o nível de bolsa que as mesmas se enquadram. Logo, foi possível apresentar a produção bibliográfica das bolsistas, e assim extrair as seguintes conclusões: quanto a produção bibliográfica, as bolsistas possuem mais representatividade comparativamente com a produção de todas as doutoras; quanto ao quantitativo da produção bibliográfica das bolsistas fracionado por conjuntos, tanto para os artigos em anais de congresso, quanto para os artigos em periódicos, as bolsistas possuem produção consideravelmente superior se comparado ao total das doutoras; quanto as redes de colaboração por níveis de bolsa, bem como, sumarização das métricas adotadas, a maior concentração nós e arestas é encontrada no nível 2.

Todo o potencial que pode ser extraído dos dados curriculares da Plataforma Lattes para compreensão dos aspectos da participação feminina na ciência sobre o conjunto de dados de forma temporal foi apresentado nesta dissertação. Assim, esta temática, que aborda os aspectos gerais da mulher, particularmente na ciência, é relevante e tem como objetivo expor os aspectos de sua trajetória acadêmica e profissional, suas pesquisas, bem como sua evolução na ciência, permitindo contribuir para gerar indicadores científicos nacionais e para a gestão das informações na área científica e tecnológica, bem como, para formulação de políticas e estratégias que potencialize a atuação das mulheres nos ambientes acadêmicos.

6.1 Publicações

A seguir, são apresentadas as produções científicas elaboradas a partir dos estudos realizados nesta dissertação. Como resultados iniciais realizados a partir dos dados curriculares cadastrados na Plataforma Lattes e seleção dos indivíduos pelo critério de gênero, foram produzidos e publicados os seguintes artigos:

- SANTIAGO, M. O.; DIAS, T. M. R.. Caracterização do conjunto de doutores brasileiros com ênfase em gênero. In: **XIX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2018**, v. 24, n. 2, p. 4741–4749, Outubro 2018. Disponível em: <https://brapci.inf.br/index.php/res/v/102902>.

- SANTIAGO, M. O.; DIAS, T. M. R. Distribuição e análise do conjunto de doutores brasileiros baseado em gênero. In: **Pesquisa Brasileira em Ciência da Informação e Biblioteconomia**, v. 13, n. 2, p. 264–271, 2018. DOI: 10.22478/ufpb.1981-0695.2018v13n2.42883. Disponível em: <https://brapci.inf.br/index.php/res/v/106772>.

Com a alteração no processo de seleção por gênero e acrescentando novas análises, foram publicadas mais três artigos que seguiram o mesmo tema de estudo:

- SANTIAGO, M. O.; DIAS, T. M. R.. Dados de acesso aberto para compreensão da produção científica das mulheres no Brasil. In: **Ciência da Informação**, v. 48, n. 3, p. 486-487, 2019. Suplemento: Trabalhos apresentados na 10ª Conferência Luso-Brasileira de Ciência Aberta - ConfOA / (Pôster) Manaus, disponível em: <http://revista.ibict.br/ciinf/article/view/4930>.

- SANTIAGO, M. O.; DIAS, T. M. R.. Um estudo bibliométrico sobre a produção científica das mulheres no Brasil. In: **II LATmetrics: Métricas Alternativas y Ciencia Abierta en América Latina**, p. 190–200, Novembro 2019. Anais do II LATmetrics: Métricas Alternativas y Ciencia Abierta en América Latina, 2019 disponível em: <https://app.uff.br/riuff/handle/1/14146>.

- SANTIAGO, M. O.; DIAS, T. M. R. ; AFFONSO, F. . Caracterização da produção científica e tecnológica das doutoras no brasil. In: **III Workshop de Informação, Dados e Tecnologia - WIDaT 2019**, v. 1. p. 42-47, Brasília: UNB, 2019. Anais do III Workshop de Informação, Dados e Tecnologia (WIDaT 2019) disponível em: <http://widat2019.fci.unb.br/index.php/anais-widat-2019>

Como consequência da produção anterior, publicada no WIDaT 2019 (III Workshop de Informação, Dados e Tecnologia), os autores receberam um convite para envio de uma versão estendida do artigo a ser publicado na Edição especial da revista Ciência da Informação:

- Perfil das orientações e produções das mulheres fundamentando em dados da Plataforma Lattes. (No prelo)

Posteriormente, foi publicado o seguinte artigo:

- SANTIAGO M..O., DIAS T.M.R., AFFONSO F.. Characterization of Women's Scientific Participation in Brazil. In: **Mugnaini R. (eds) Data and Information in Online Environments. DIONE 2020**. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering. ed. 319. Florianópolis: Springer International Publishing, 2020, v. 01, p. 210-221. DOI: https://doi.org/10.1007/978-3-030-50072-6_16

Como consequência da produção publicada no DIONE 2020, os autores receberam um convite para envio de uma versão estendida do artigo para o periódico Transinformação. No artigo estendido o processo de seleção por gênero foi detalhado, o conjunto de dados foi atualizado, que teve como consequência a atualização dos resultados, bem como exposição de novos aspectos da produção:

- SANTIAGO, M. d. O.; AFFONSO, F.; DIAS, T. M. R. Scientific production of women in Brazil. **Transinformação**, SciELO Brasil, v. 32, 2020. Disponível em: <https://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-37862020000100203&tlng=en>

6.2 Trabalhos futuros

Dando continuidade à investigação que possui a temática das mulheres na ciência e aproveitando todo o potencial e a riqueza dos dados da Plataforma Lattes, algumas extensões para possíveis trabalhos futuros são sugeridos:

- **Visão detalhada das redes de todas as doutoras**: este estudo objetiva apresentar uma visão aprofundada e abrangente sobre a colaboração das doutoras e teria como tópicos de investigação apresentar as características dos nós e arestas, as arestas

que possuem maior peso, colaboração por áreas, instituições de vínculos, dentre outros.

- **Análise da trajetória acadêmica de todas as mulheres que possuem o currículo cadastrado na Plataforma Lattes**: neste estudo seria realizado uma análise para verificar os aspectos que correspondem ao caminho percorrido pelas mulheres desde a iniciação científica até o último grau de formação acadêmica, considerando todo o histórico de formação registrado nos currículos, incluindo também o período de licença-maternidade.

- **Investigação da qualidade da produção científica**: neste estudo seria analisado a produção científica das doutoras, através de um comparativo sobre algumas métricas qualitativas dos periódicos, artigos e impacto das produções.

- **Análise das docentes em programas de pós-graduação**: neste estudo, seria realizado uma correlação dos dados da Plataforma Sucupira com os dados da Plataforma Lattes, para analisar as orientações, impacto das produções científicas, dentre outros aspectos sobre o processo de formação acadêmica nos mais altos níveis de formação.

Referências

ABRAMO, G.; D'ANGELO, C. A.; COSTA, F. D. The effects of gender, age and academic rank on research diversification. **Scientometrics**, Springer, v. 114, n. 2, p. 373–387, 2018. Disponível em: <https://doi.org/10.1007/s11192-017-2529-1>. Acesso em: 15 mar. 2019. Citado na página 18.

ALMEIDA, E. C. E. d.; GUIMARÃES, J. A. **A pós-graduação e a evolução da produção científica brasileira**. São Paulo: Editora Senac São Paulo, 2013. Citado 2 vezes nas páginas 59 e 63.

ARAÚJO, C. A. Bibliometria: evolução histórica e questões atuais. **Em Questão**, Universidade Federal do Rio Grande do Sul, v. 12, n. 1, p. 11–32, 2006. Citado 4 vezes nas páginas 2, 6, 7 e 8.

ARENSBERGEN, P. van; WEIJDEN, I. van der; BESSELAAR, P. Van den. Gender differences in scientific productivity: a persisting phenomenon? **Scientometrics**, Akadémiai Kiadó, co-published with Springer Science+ Business Media BV . . . , v. 93, n. 3, p. 857–868, 2012. Disponível em: <https://doi.org/10.1007/s11192-012-0712-y>. Acesso em: 11 set. 2018. Citado na página 16.

ARRUDA, D. et al. Brazilian computer science research: Gender and regional distributions. **Scientometrics**, Springer, v. 79, n. 3, p. 651–665, 2009. Disponível em: <https://doi.org/10.1007/s11192-007-1944-0>. Acesso em: 24 abr. 2019. Citado na página 1.

BALANCIERI, R. **Análise de redes de pesquisa em uma plataforma de gestão em ciência e tecnologia: uma aplicação à plataforma Lattes**. 117 p. Dissertação (Mestrado em Engenharia de Produção) — Universidade Federal de Santa Catarina, 2004. Citado na página 66.

BELLOTTI, E.; GUADALUPI, L.; CONALDI, G. Comparing fields of sciences: Multilevel networks of research collaborations in italian academia. In: **Multilevel network analysis for the social sciences**. [S.l.]: Springer, 2016. p. 213–244. Disponível em: <https://doi.org/10.1007/978-3-319-24520-1_9>. Acesso em: 26 dez. 2019. Citado na página 52.

BELLOTTI, E.; KRONEGGER, L.; GUADALUPI, L. The evolution of research collaboration within and across disciplines in italian academia. **Scientometrics**, Springer, v. 109, n. 2, p. 783–811, 2016. Disponível em: <https://doi.org/10.1007/s11192-016-2068-1>. Acesso em: 03 mar. 2019. Citado na página 52.

BESUTTI, J.; REDANTE, R. C.; FÁVERO, A. A. Formação e construção da identidade docente a partir da narrativa de histórias de vida. **Educação Por Escrito**, v. 8, n. 2, p. 260–277, 2017. Citado na página 57.

BOURDIEU, P. **Usos sociais da ciência**. São Paulo: Editora Unesp, 2004. Citado na página 42.

BRITO, A. G. C.; QUONIAM, L.; MENA-CHALCO, J. P. Exploração da plataforma lattes por assunto: proposta de metodologia. **TransInformação**, Pontifícia Universidade Católica de Campinas, v. 28, n. 1, p. 77–86, 2016. Disponível em: <http://dx.doi.org/10.1590/

2318-08892016002800006>. Acesso em: 09 mai. 2019. Citado 2 vezes nas páginas 20 e 62.

CAPES. **Mestrado e Doutorado: o que são?** Brasil, 2014. Disponível em: <https://capes.gov.br/avaliacao/sobre-a-avaliacao/mestrado-e-doutorado-o-que-sao>. Acesso em: 20 mar. 2020. Citado na página 63.

CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO - CNPq. **Produtividade em Pesquisa - PQ**. Brasil, 2020. Disponível em: <http://cnpq.br/documents/10157/5f43cefd-7a9a-4030-945e-4a0fa10a169a>. Acesso em: 12 jun. 2020. Citado na página 71.

CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO - CNPq. **CNPq anuncia inclusão do campo licença-maternidade no Currículo Lattes**. Brasil, 2021. Disponível em: <https://www.gov.br/cnpq/pt-br/assuntos/noticias/cnpq-em-acao/cnpq-anuncia-inclusao-do-campo-licenca-maternidade-no-curriculo-lattes>. Acesso em: 16 abr. 2021. Citado na página 57.

CORMEN, T. H. et al. **Introduction to algorithms**. 3. ed. [S.l.]: MIT press, 2009. Citado na página 12.

DAMACENO, R. J. et al. The brazilian academic genealogy: evidence of advisor–advisee relationships through quantitative analysis. **Scientometrics**, Springer, p. 1–31, 2019. Disponível em: <http://https://doi.org/10.1007/s11192-019-03023-0>. Acesso em: 09 mai. 2019. Citado na página 21.

DIAS, T. M. R. **Um Estudo da Produção Científica Brasileira a partir de Dados da Plataforma Lattes**. 181 p. Tese (Doutorado em Modelagem Matemática e Computacional) — Centro Federal de Educação Tecnológica de Minas Gerais, Setembro 2016. Citado 9 vezes nas páginas 1, 6, 7, 20, 25, 26, 44, 45 e 76.

DIAS, T. M. R.; MOITA, G. F. Um retrato da produção científica brasileira baseado em dados da plataforma lattes. **Brazilian Journal of Information Science: research trends**, v. 12, n. 4, p. 62–74, 2018. Citado 2 vezes nas páginas 49 e 52.

DIGIAMPIETRI, L. A. **Análise da rede social acadêmica brasileira**. 160 p. Tese (Livre Docência) — Escola de Artes, Ciências e Humanidades, Universidade de São Paulo, São Paulo, 2015. Citado 3 vezes nas páginas 10, 12 e 19.

DIRETÓRIO DE GRUPOS DE PESQUISA NO BRASIL. **Censos do Diretório de Grupos de Pesquisa**. Brasil, 2020. Disponível em: <http://lattes.cnpq.br/web/dgp/home>. Acesso em: 15 jan. 2020. Citado na página 27.

DOMINGUES, I. O sistema de comunicação da ciência e o taylorismo acadêmico: questionamentos e alternativas. **Estudos avançados**, SciELO Brasil, v. 28, n. 82, p. 225–250, 2014. Citado na página 43.

EDLER, J.; FLANAGAN, K. Indicator needs for the internationalisation of science policies. **Research Evaluation**, Beech Tree Publishing, v. 20, n. 1, p. 7–17, 2011. Citado na página 9.

ELSEVIER. **Gender in the Global Research Landscape**. [S.l.]: Elsevier, 2017. Disponível em: <https://www.elsevier.com/__data/assets/pdf_file/0008/265661/ElsevierGenderReport_final_for-web.pdf>. Acesso em: 01 mar. 2020. Citado na página 50.

EOM, Y.-H.; FORTUNATO, S. **Characterizing and Modeling Citation Dynamics**. 2011. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0024926>. Acesso em: 16 ago. 2020. Citado na página 75.

FELIZARDO, T. M. **Colaboração Científica - Um Estudo das Práticas de Coautoria na Área Operations Research Management Science da Web of Science**. 108 p. Dissertação (Mestrado em Engenharia de Produção e Sistemas) — Centro Federal de Educação Tecnológica Celso Suckow da Fonseca (CEFET/RJ), 2018. Citado na página 12.

FERRAZ, R. R. N.; QUONIAM, L. M.; MACCARI, E. A. A utilização da ferramenta scriptlattes para extração e disponibilização on-line da produção acadêmica de um programa de pós-graduação stricto sensu em administração. **Revista Brasileira de Pós-Graduação**, 11 th International Conference on Information Systems and Technology Management - CONTECSI, v. 11, n. 24, 2015. Disponível em: <http://dx.doi.org/10.21713/2358-2332.2014.v11.445>. Acesso em: 27 mar. 2019. Citado 2 vezes nas páginas 9 e 11.

FERREIRA, L. M.; FURTADO, F.; SILVEIRA, T. S. Advisor-advisee relationship: the multiplier knowledge. **Acta cirurgica brasileira**, SciELO Brasil, v. 24, n. 3, p. 170–172, 2009. Citado na página 41.

FREIRE, G. H. de A.; FREIRE, I. M. Sobre o qualis de periódicos da capes. **Informação & Sociedade**, João Pessoa, v. 27, n. 3, 2017. Citado na página 46.

GABARDO, A. C. **Análise de redes sociais: uma visão computacional**. 1. ed. São Paulo: Novatec Editora, 2015. Citado na página 12.

GIL, A. C. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo: Editora Atlas S.A., 2008. Citado na página 24.

GOMES, J. O. **Uma análise temporal dos principais tópicos de pesquisa da ciência brasileira a partir das palavras-chave de publicações científicas**. 127 p. Tese (Doutorado em Modelagem Matemática e Computacional) — Centro Federal de Educação Tecnológica de Minas Gerais, Dezembro 2018. Citado na página 21.

GONÇALVES, D. C. K. **Diferenciais por áreas de conhecimento no impacto das bolsas de produtividade em pesquisa do CNPq sobre a produção acadêmica dos contemplados: um estudo para o Brasil entre 2013-2016**. 132 p. Dissertação (Mestrado em Políticas Públicas) — Universidade Federal do Paraná, 2020. Citado na página 71.

GRÁCIO, M. C. C. Colaboração científica: indicadores relacionais de coautoria. **Brazilian Journal of Information Science: research trends**, v. 12, n. 2, 2018. Citado na página 66.

GROSSI, M. G. R. et al. As mulheres praticando ciência no brasil. **Revista Estudos Feministas**, JSTOR, v. 24, n. 1, p. 11–30, 2008. Citado 2 vezes nas páginas 19 e 22.

GUEDES, V. L. S.; BORSCHIVER, S. Bibliometria: uma ferramenta estatística para a gestão da informação e do conhecimento, em sistemas de informação, de comunicação e de avaliação científica e tecnológica. In: **Encontro Nacional de Ciência da Informação**. [S.l.: s.n.], 2005. v. 6, n. 1, p. 18. Citado na página 8.

HILÁRIO, C. M.; GRÁCIO, M. C. C.; GUIMARÃES, J. A. C. Aspectos éticos da coautoria em publicações científicas. **Em Questão**, v. 24, n. 2, p. 12–36, 2018. Citado na página 66.

HOPPEN, N. H. F.; VANZ, S. A. de S. O que são estudos de gênero: caracterização da produção científica autodenominada estudos de gênero em uma base de dados multidisciplinar e internacional. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, v. 25, p. 01–30, 2020. Citado na página 2.

KATZ, J. S.; MARTIN, B. R. What is research collaboration? **Research policy**, Elsevier, v. 26, n. 1, p. 1–18, 1997. Citado na página 12.

KRETSCHMER, H.; KRETSCHMER, T. Gender bias and explanation models for the phenomenon of women's discriminations in research careers. **Scientometrics**, Springer, v. 97, n. 1, p. 25–36, 2013. Disponível em: <https://doi.org/10.1007/s11192-013-1023-7>. Acesso em: 15 abr. 2019. Citado na página 3.

LANE, J. Let's make science metrics more scientific. **Nature**, Nature Publishing Group, v. 464, n. 7288, p. 488, 2010. Disponível em: <https://doi.org/10.1038/464488a>. Acesso em: 27 mar. 2019. Citado 2 vezes nas páginas 1 e 19.

LEITE.FILHO, G. A.; MARTINS, G. d. A. Relação orientador-orientando e suas influências na elaboração de teses e dissertações. **Revista de Administração de Empresas**, SciELO Brasil, v. 46, n. SPE, p. 99–109, 2006. Citado na página 42.

LETA, J. Mulheres na ciência brasileira: desempenho inferior? **Revista Feminismos**, v. 2, n. 3, p. 139–152, Set.–Dez. 2014. Citado 4 vezes nas páginas 2, 4, 17 e 26.

LIU, J.; SONG, Y.; YANG, S. Gender disparities in the field of economics. **Scientometrics**, Springer, v. 125, n. 2, p. 1477–1498, 2020. Citado na página 77.

LOPES, E. F. B. et al. A relação entre orientador e orientando no processo de produção científica/the relationship between guiding and guiding in the scientific production process. **Brazilian Journal of Development**, v. 6, n. 1, p. 3854–3868, 2020. Citado 2 vezes nas páginas 36 e 43.

LORENS, A. L. C. et al. Participação feminina em comitês de programa de simpósios da computação. p. 90–99, 2020. Citado na página 62.

MACHADO , T. C.; LAMPERT, E.; FALAVIGNA, G. O professor do ensino superior: pesquisador ou docente? **Revista FACISA ON-LINE**, v. 6, n. 1, 2017. Citado na página 63.

MARINI, G.; MESCHITTI, V. The trench warfare of gender discrimination: evidence from academic promotions to full professor in italy. **Scientometrics**, Springer, v. 115, n. 2, p. 989–1006, 2018. Disponível em: <https://doi.org/10.1007/s11192-018-2696-8>. Acesso em: 15 mar. 2019. Citado na página 18.

MARTINS, C. B. As origens pós-graduação nacional (1960-1980). **Revista Brasileira de Sociologia**, Sociedade Brasileira de Sociologia, v. 6, n. 13, p. 9–26, 2018. Citado na página 59.

MARTINS, M. F.; VARANI, A. Professor e pesquisador: considerações sobre a problemática relação entre ensino e pesquisa. **Revista Diálogo Educacional**, v. 12, n. 37, p. 647–680, 2012. Citado na página 63.

MEDEIROS, T. D.; HOPPEN, N. H. F.; VANZ, S. A. de S. A produção científica sobre estudos de gênero no repositório digital da ufrgs. **BIBLOS**, v. 34, n. 2, 2020. Disponível em: <<https://brapci.inf.br/index.php/res/v/102902>>. Citado na página 2.

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES. **Recursos Humanos - Indicadores sobre o ensino de pós-graduação**. Brasil, 2019. Disponível em: <https://www.mctic.gov.br/mctic/opencms/indicadores/detalhe/Recursos_Humanos/RH_3.5.4.html>. Acesso em: 29 abr. 2020. Citado na página 39.

MOSCHKOVICH, M.; ALMEIDA, A. M. F. Desigualdades de gênero na carreira acadêmica no brasil. **Dados-Revista de Ciências Sociais**, Universidade do Estado do Rio de Janeiro, v. 58, n. 3, 2015. Disponível em: <http://dx.doi.org/10.1590/00115258201558>. Acesso em: 09 mai. 2019. Citado na página 17.

MUGNAINI, R.; DIGIAMPIETRI, L. A.; MENA-CHALCO, J. P. Comunicação científica no brasil (1998-2012): indexação, crescimento, fluxo e dispersão. **Transinformação**, Pontifícia Universidade Católica de Campinas, v. 26, n. 3, p. 239–252, 2014. Citado na página 3.

NEWMAN, M. E. The structure of scientific collaboration networks. **Proceedings of the national academy of sciences**, National Acad Sciences, v. 98, n. 2, p. 404–409, 2001. Citado na página 12.

OLINTO, G. A inclusão das mulheres nas carreiras de ciência e tecnologia no brasil. **Inclusão Social**, Brasília, DF, v. 5, n. 1, p. 68–77, jul./dez. 2011. Citado na página 49.

PERUCCHI, V. **Produção de conhecimento científico e tecnológico nos Institutos Federais de Educação, Ciência e Tecnologia: uma investigação sobre a sua natureza, divulgação e aplicação**. 152 p. Tese (Doutor em Ciência da Informação) — Universidade de Brasília (UnB), Março 2015. Citado na página 48.

PLATAFORMA LATTES. **História do surgimento da Plataforma Lattes**. Brasil, 2019. Disponível em: <http://lattes.cnpq.br/>. Acesso em: 02 abr. 2019. Citado na página 9.

PRITCHARD, A. et al. Statistical bibliography or bibliometrics. **Journal of documentation**, New York, v. 25, n. 4, p. 348–349, 1969. Citado na página 6.

QUIRINO, C. A. d. S. **Espelho, espelho meu, existe alguém mais estressado do que eu? O impacto do mestrado em servidores públicos federais**. 93 p. Dissertação (Mestrado em Administração Pública em Rede Nacional (Profiap)) — Universidade Federal do Vale do São Francisco, 2018. Citado na página 64.

SACARDO, M. S.; HAYASHI, M. C. P. I. Bibliometric balance of scientific literature in physical education and special education originating from theses and dissertations/balanco bibliometrico da producao cientifica em educacao fisica e educacao especial oriunda de teses e dissertacoes/balance bibliométrico de la producción científica en educación física y educación especial oriunda de tesis y disertaciones. **Revista Brasileira de Pos-Graduacao**, Coordenacao de Aperfeicoamento de Pessoal de Nivel Superior, v. 8, n. 15, p. 108–135, 2011. Citado 2 vezes nas páginas 63 e 64.

SANTIN, D. M.; VANZ, S. A. d. S.; CAREGNATO, S. E. A análise de redes de colaboração científica com base em indicadores bibliométricos. In: **Educação superior e conhecimento no centenário da Reforma de Córdoba: novos olhares em contextos emergentes.** Porto Alegre: EDIPUCRS, 2019. p. 189–207. Citado na página 70.

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. Citado na página 20.

SCOTT, J. W. Gênero: uma categoria útil de análise histórica. **Educação & Realidade**, Porto Alegre, v. 20, n. 2, p. 71–99, Jul.–Dez. 1995. Citado na página 2.

SILVA, F. F.; RIBEIRO, P. R. C. Trajetórias de mulheres na ciência:"ser cientista" e "ser mulher". **Ciência & Educação**, v. 20, n. 2, p. 449–466, 2014. Disponível em: <http://dx.doi.org/10.1590/1516-73132014000200012>. Acesso em: 26 mar. 2019. Citado 3 vezes nas páginas 2, 4 e 17.

SILVA, M. R. d.; HAYASHI, C. R. M.; HAYASHI, M. C. P. I. Análise bibliométrica e cientométrica: desafios para especialistas que atuam no campo. **InCID: Revista de Ciência da Informação e Documentação**, v. 2, n. 1, p. 110–129, 2011. Disponível em: <https://doi.org/10.11606/issn.2178-2075.v2i1p110-129>. Acesso em: 4 abr. 2019. Citado na página 6.

SILVA, T. H. et al. The impact of academic mobility on the quality of graduate programs. **D-Lib Magazine**, Corporation for National Research Initiatives, v. 22, n. 9/10, 2016. Disponível em: <http://mirror.dlib.org/dlib/september16/silva/09silva.html>. Acesso em: 21 mar. 2020. Citado na página 59.

SILVA, W. D. d. et al. Anotação semântica automática do currículo lattes utilizando linked open data. **Perspectivas em Ciência da Informação**, v. 23, n. 4, p. 53–72, 2018. Citado na página 21.

SOARES, P. C. Contradições na pesquisa e pós-graduação no brasil. **Estudos avançados**, SciELO Brasil, v. 32, n. 92, p. 289–313, 2018. Citado na página 66.

SOARES, S. V.; NOVA, S. P. d. C. C. Pesquisadores brasileiros que publicam em periódicos internacionais: qual sua formação acadêmica? **Revista Gestão Universitária na América Latina-GUAL**, v. 10, n. 1, p. 125–149, 2017. Disponível em: <https://periodicos.ufsc.br/index.php/gual/article/view/1983-4535.2017v10n1p125>. Acesso em: 21 mar. 2020. Citado na página 57.

SONNENWALD, D. H. Scientific collaboration. **Annual Review of Information Science and Technology**, v. 41, n. 1, p. 643–681, 2007. Citado na página 66.

SPINAK, E. Indicadores cienciométricos. **Ciência da informação**, v. 27, n. 2, p. 141–148, 1998. Citado na página 9.

STRÖELE, V. et al. Identifying workgroups in brazilian scientific social networks. **Journal of Universal Computer Science**, v. 17, n. 14, p. 1951–1970, 2011. Citado na página 66.

SUCUPIRA. **Qualis**. Brasil, 2020. Disponível em: <https://sucupira.capes.gov.br/sucupira/public/index.xhtml#>. Acesso em: 07 abr. 2020. Citado na página 46.

TANNENBAUM, C. et al. Sex and gender analysis improves science and engineering. **Nature**, Nature Publishing Group, v. 575, n. 7781, p. 137–146, 2019. Citado na página 66.

TIAN, S.-w. et al. Evolution model for scientific collaboration network with local-world information. **The Journal of China Universities of Posts and Telecommunications**, Elsevier, v. 21, n. 2, p. 15–31, 2014. Citado na página 12.

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. Citado 3 vezes nas páginas 22, 48 e 52.

VALENTOVA, J. V. et al. Underrepresentation of women in the senior levels of brazilian science. **PeerJ**, e4000, v. 5, p. 1–20, Dec. 2017. Disponível em: <https://doi.org/10.7717/peerj.4000>. Acesso em: 24 abr. 2019. Citado 4 vezes nas páginas 3, 4, 18 e 49.

VANTI, N. A. P. Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. **Ciência da informação**, SciELO Brasil, v. 31, n. 2, p. 152–162, 2002. Citado 2 vezes nas páginas 6 e 7.

VOGEL, M. J. M. **Avaliação da pós-graduação brasileira: análise dos quesitos utilizados pela Capes e das críticas da comunidade acadêmica**. Tese (Doutorado) — Universidade de São Paulo, 2015. Citado na página 46.