

**UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE EDUCAÇÃO E CIÊNCIAS HUMANAS
PROGRAMA DE PÓS GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO**

VANESSA PAULA ALVES DE MOURA

**A CONTRIBUIÇÃO DA ASSOCIAÇÃO ENTRE GENEALOGIA ACADÊMICA E
BIBLIOMETRIA PARA A AVALIAÇÃO DE PROGRAMAS DE PÓS-
GRADUAÇÃO**

**SÃO CARLOS
2020**

VANESSA PAULA ALVES DE MOURA

**A CONTRIBUIÇÃO DA ASSOCIAÇÃO ENTRE GENEALOGIA ACADÊMICA E
BIBLIOMETRIA PARA A AVALIAÇÃO DE PROGRAMAS DE PÓS-
GRADUAÇÃO**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de São Carlos como requisito parcial para a obtenção do título de mestre em Ciência da Informação.

Área de concentração: Conhecimento, Tecnologia e Inovação.

Linha de pesquisa 1: Conhecimento e Informação para a Inovação

Orientador: Prof. Dr. Leandro Innocentini Lopes de Faria.

**SÃO CARLOS
2020**



UNIVERSIDADE FEDERAL DE SÃO CARLOS
Centro de Educação e Ciências Humanas
Programa de Pós-Graduação em Ciência da Informação

Folha de Aprovação

Defesa de Dissertação de Mestrado da candidata Vanessa Paula Alves de Moura, realizada em 28/07/2020.

Comissão Julgadora:

Prof. Dr. Leandro Innocentini Lopes de Faria (UFSCar)

Prof. Dr. Roniberto Morato do Amaral (UFSCar)

Prof. Dr. Ricardo Arcanjo de Lima (EMBRAPA)

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.
O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Ciência da Informação.

À Gabi, Isa e meu sobrinhos na
esperança de que trilhem suas
vidas no caminho do bem e do
conhecimento.

AGRADECIMENTOS

A Deus que me concedeu a vida e tudo que eu precisava para chegar aonde estou.

A todos os professores e colegas do Departamento de Ciência de Informação da UFSCar pelo companheirismo e colaboração.

Ao Prof. Dr. Leandro Innocentini Lopes de Faria, meu orientador, pela confiança, paciência, apoio e oportunidade.

Ao Prof. Dr. Jesus P. Mena-Chalco, que além de toda ajuda, e provavelmente sem querer, me ensinou, na prática, o que é humildade. Bem como sua equipe, em especial, ao doutorando Rafael Jeferson Pezzuto Damaceno, que compartilhou comigo, sem medir esforços, o seu conhecimento.

A todos da equipe do Grupo de Estudos do Núcleo de Informação Tecnológica (NIT Materiais) pela recepção, apoio, carinho e grandes contribuições para a conclusão desta pesquisa.

À empresa Uppercase, em especial ao Aurivan Galdino, que flexibilizou o meu trabalho, para que eu pudesse dar andamento ao meu projeto de mestrado.

A todos da Fatec-Jahu pelo incentivo e apoio para a realização deste projeto, em especial ao Flávio Luiz Grava Scalco e à amiga Maria Aparecida Zem Lopes.

À minha mãe Paulina (in memoriam), que em tão pouco tempo que esteve em minha vida, me ensinou os principais valores quais ainda me norteiam e que me fizeram chegar até aqui.

Ao meu namorado Paulo, que sempre me apoia, me motiva, me ensina, me surpreende e que me deu a força diária para a realização deste projeto com toda paciência e amor do mundo.

Às minhas filhas amadas, Isabelle e Gabrielle que, mesmo não tendo a noção das minhas dificuldades e desafios diários, são os principais motivos que me fazem respirar fundo e ir adiante no caminho da busca pelo conhecimento.

As minhas irmãs Nilcéia e Cláudia pelas críticas construtivas e apoio.

Ao meu irmão Émerson pelas palavras de carinho e incentivo que me dão suporte.

A todos, minha eterna gratidão.

“ Outra parte das sementes caiu em terra boa e deu fruto, brotando e crescendo: rendeu trinta, sessenta e até cem por um.”

(Bíblia Sagrada, Marcos 4:8)

RESUMO

A medição da produção científica é um objeto de estudos muito explorado na Ciência da Informação, por meio da Bibliometria, que permite avaliar o desempenho dos pesquisadores e os impactos de suas produções, tendo como resultado os indicadores bibliométricos, que são úteis para compreender a realidade e agir em diversos cenários. A evolução tecnológica permitiu o aprimoramento da Bibliometria e beneficiou também outros estudos, como a Genealogia Acadêmica, que de forma mais recente, vem sendo utilizada para analisar o desempenho dos pesquisadores e os impactos de suas atuações para a Ciência, através da relação entre orientador e seus orientados. Um dos estudos mais recentes no Brasil sobre a Genealogia Acadêmica, tem como base de dados a Plataforma Lattes e a extração desses dados é feita por um *software* chamado AGLattes, desenvolvido por pesquisadores da área da Ciência da Computação. Apesar dos dados obtidos serem ricos e valiosos, há poucas pesquisas que os utilizam com o propósito de complementar os indicadores bibliométricos. Diante disso, o presente estudo tem como objetivo investigar se a aplicação da Genealogia Acadêmica e da análise bibliométrica em conjunto pode trazer informações relevantes sobre a contribuição de um Programa de Pós-Graduação na formação de pesquisadores e na produção de conhecimento. Para a execução da pesquisa, foi escolhido o Programa de Pós-Graduação em Engenharia de Materiais da Universidade Federal de São Carlos - PPGCEM/UFSCar - que tem como origem o primeiro curso de Engenharia de Materiais no Brasil. A análise foi feita a partir de todos os docentes que passaram e estão no programa desde a sua fundação, um total de 58 docentes. Como esses docentes atuam em outros programas, foi desenvolvido um procedimento que viabilizou chegar somente nos pesquisadores que têm a sua origem no PPGCEM/UFSCar, eliminando a contaminação dos dados de outros programas e permitindo obter informações de 993 descendentes diretos e 1205 indiretos. Para extração da produção científica dos Currículos Lattes desses descendentes foi utilizada a ferramenta Lattes Machine e posteriormente a mineração de dados e análises bibliométricas foram realizadas no *software* VantagePoint. Os resultados são ilustrados por indicadores bibliométricos construídos a partir da associação da Genealogia Acadêmica com a Bibliometria, que fornecem informações valiosas sobre o programa, sobre os recursos humanos formados e de todos seus descendentes, como por exemplo, onde trabalham, como estão distribuídos pelo mundo e no Brasil, a área de formação, além de informações sobre suas produções científicas que podem ser utilizadas na elaboração de estratégias, na autoavaliação dos programas de pós-graduação visando a garantia da qualidade e obtenção de recursos das agências de fomento, neste caso a CAPES. O estudo também contribui para constatar a importância histórica e atual do PPGCEM/UFSCar e de seus pesquisadores para a área do programa e para a Ciência.

Palavras-Chave: Bibliometria, Indicadores Bibliométricos, Genealogia Acadêmica, Plataforma Lattes, Engenharia de Materiais.

ABSTRACT

The measurement of scientific production is an object of studies much explored in Information Science, through Bibliometrics, which allows to evaluate the performance of researchers and the impacts of their productions, and have the results bibliometric indicators, which are useful to understand reality and act in different situations. Technological evolution allowed the improvement of Bibliometrics and has also benefited other studies, such as Academic Genealogy, which, more recently, has been used to analyze the performance of researchers and the impacts of their actions on Science, through the relationship between supervisor and its oriented. One of the most recent studies in Brazil on Academic Genealogy, is based on the Lattes Platform and the extraction of these data is done by software called AGLattes, developed by researchers in the field of Computer Science. Although the data obtained are rich and valuable, there is little research that uses them for the purpose of complementing bibliometric indicators. Therefore, this study aims to investigate whether the application of Academic Genealogy and bibliometric analysis together can bring relevant information about the contribution of a Postgraduate Program in the training of researchers and in the production of knowledge. For the execution of the research, the Postgraduate Program in Materials Engineering of the Federal University of São Carlos - PPGCEM/UFSCar - was chosen, which originated in the first course of Materials Engineering in Brazil. The analysis was made from all teachers which are in the program since its foundation, a total of 58 teachers. As these professors work in other programs, a procedure was developed that made it possible to reach only researchers who have their origin in PPGCEM/UFSCar, eliminating the contamination of data from other programs and allowing to obtain information from 993 direct and 1205 indirect descendants. To extract the scientific production of the Lattes Curriculum of these descendants, the software Lattes Machine was used and later data mining and bibliometric analyzes were performed using the VantagePoint software. The results are illustrated by bibliometric indicators built from the association of Academic Genealogy with Bibliometrics, which provide valuable information about the program, the human resources trained and all their descendants, such as where they work, how they are distributed around the world and in Brazil, the training area, in addition to information about its scientific productions that can be used in the elaboration of strategies, in the self-assessment of graduate programs aiming at guaranteeing the quality and obtaining resources from development agencies, in this case the CAPES. The study also contributes to verify the historical and current importance of PPGCEM/UFSCar and its researchers for the program area and for Science.

Keywords: Bibliometrics, Bibliometric Indicators, Academic Genealogy, Lattes Platform, Materials Engineering.

LISTA DE QUADROS

Quadro 1: Os dez países que mais investem em P&D.....	27
Quadro 2: Países que investem maior % de seu PIB em P&D.....	27
Quadro 3: Crescimento do número de Currículos Lattes.....	46
Quadro 4: Relação dos PPG da UFSCar e suas respectivas notas.	71
Quadro 6: Síntese das ferramentas utilizadas.....	75
Quadro 5: Sementes do PPGCEM/UFSCar 1978 - 2019	80
Quadro 7: Formação por orientador e tempo de atuação.....	93
Quadro 8: Acadêmicos que aderiram ao programa após formação	94
Quadro 9: Principais Revistas Científicas.....	109

LISTA DE FIGURAS

Figura 1: Investimento em P&D por país	26
Figura 2: Currículos indexados na Plataforma Lattes em 2019	47
Figura 3: Pôster personalizado de GA oferecidos pelo PGM	55
Figura 4: Página inicial do Portal do PGM	56
Figura 5: Resultados da pesquisa feita no PGM	57
Figura 6: Página Inicial da Plataforma Acácia	58
Figura 7: Resultados da pesquisa feita na Plataforma Acácia.....	59
Figura 8: Exemplo de grafo	60
Figura 9: Métricas aplicadas em uma árvore genealógica.....	62
Figura 10: Genealogia acadêmica da ESPN	64
Figura 11: Grafo da GA de Naercio Aquino Menezes Filho.....	65
Figura 12: Rede de colaboração do PPGCEM/UFSCar	72
Figura 13: Distribuição de notas dos Programas de Pós-Graduação	39
Figura 14: Formação de mestres e doutores pelo PPGCEM/UFSCar.....	79
Figura 15: Fluxograma do procedimento da pesquisa.....	81
Figura 16: Método de classificação dos descendentes	84
Figura 17: Solução do problema no AGLattes.....	85
Figura 18: Formação de recursos humanos pelo PPGCEM/UFSCar.....	90
Figura 19: Descendentes diretos por ano e por titulação	91
Figura 20: Atuação profissional dos descendentes	92
Figura 21: Atuação por Instituição	95
Figura 22: Atuação por Departamento.....	96
Figura 23: Localização dos descendentes por país.....	97
Figura 24: Localização dos descendentes por regiões brasileiras	99
Figura 25: Cidades com maior concentração de descendentes	100
Figura 26: Principais áreas de formação de mestrado	101
Figura 27: Principais áreas de formação de doutorado	102
Figura 28: Tipos de publicações.....	104
Figura 29: Quantidade de publicações por ano e por áreas	105
Figura 30: Quantidade de publicações por área e subáreas	106
Figura 31: Quantidade de publicações por especialidade	107
Figura 32: Publicações por Qualis Periódico	108

Figura 33: Publicações em coautoria com orientador.....	111
Figura 35: Análise da relação por meio das publicações científicas.....	112

LISTA DE ABREVIATURAS E SIGLAS

CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico
DEMa	Departamento de Engenharia de materiais
ENSP	Escola Nacional de Saúde Pública Sergio Arouca, da Fundação Oswaldo Cruz
FAP	Fundações Estadual de Amparo à Pesquisa
FAPESP	Fundação de Amparo à Pesquisa do Estado de São Paulo
FAPERJ	Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro
GA	Genealogia Acadêmica
NIT	Grupo de Estudos do Núcleo de Informação Tecnológica
MEC	Ministério da Educação
PGM	Projeto de Genealogia Matemática
PPG	Programa de Pós-Graduação
PPGCEM/UFSCar	Programa de Pós-Graduação em Ciência e Engenharia de Materiais da Universidade Federal de São Carlos
P&D	Pesquisa e Desenvolvimento
UFSCar	Universidade Federal de São Carlos

SUMÁRIO

1	INTRODUÇÃO.....	15
1.1	OBJETIVO.....	19
2	REVISÃO DA LITERATURA.....	22
2.1	CIÊNCIA E A SOCIEDADE	22
2.2	A COMUNIDADE E A COMUNICAÇÃO CIENTÍFICA.....	28
2.3	A CIÊNCIA DA INFORMAÇÃO	34
2.4	A AVALIAÇÃO DA PÓS-GRADUAÇÃO NO BRASIL E A CAPES .	36
2.5	BIBLIOMETRIA	43
2.5.1	Base de dados: Plataforma Lattes.....	45
2.5.2	Análise da produção científica a partir de Indicadores	48
2.6	GENEALOGIA ACADÊMICA.....	50
2.6.1	Estudos correlatos de Genealogia Acadêmica.....	63
2.6.1.1	Genealogia Acadêmica da ENSP	63
2.6.1.2	Genealogia Acadêmica da FEAUSP	64
2.6.1.3	Limitações da GA aplicada ao estudo de Programas de pós-graduação	66
2.7	A ENGENHARIA E A IMPORTÂNCIA DO PPGCEM/UFSCAR	68
3	MÉTODO E DESENVOLVIMENTO	74
3.1	ABORDAGEM E TIPOLOGIA DA PESQUISA	74
3.2	FERRAMENTAS	75
3.3	PROCEDIMENTO DE COLETA DOS DADOS E DELIMITAÇÃO DA AMOSTRA	79
3.4	NOVO PROCEDIMENTO DE GENEALOGIA ACADÊMICA E BIBLIOMETRIA.....	80
3.4.1	Triagem dos descendentes diretos	82
3.4.2	Triagem dos demais descendentes do PPGCEM/UFSCar..	83

3.4.3	Análise bibliométrica da segunda aplicação da Genealogia Acadêmica	86
4	RESULTADOS E DISCUSSÃO	87
4.1	INDICADORES DE FORMAÇÃO DE RECURSOS HUMANOS	88
4.1.1	Contribuição na formação de mestres e doutores	89
4.1.2	Formação por ano e por titulação	90
4.1.3	Indicador de Atuação Profissional	91
4.1.4	Formação por orientador	92
4.1.5	Aderência ao programa após a formação	94
4.1.6	Atuação por instituição	95
4.1.7	Indicador de atuação por departamento	96
4.1.8	Localização dos descendentes por país	96
4.1.9	Localização dos descendentes por regiões brasileiras	98
4.1.10	Cidades com maior concentração de descendentes	99
4.1.11	Áreas de formação de mestres e doutores	101
4.2	INDICADORES DE PUBLICAÇÕES CIENTÍFICAS	102
4.2.1	Tipos de publicações	103
4.2.2	Publicações de todos os tipos por ano e áreas	104
4.2.3	Publicações da área do PPGCEM/UFSCar	105
4.2.4	Publicações por especialidade	106
4.2.5	Periódicos com maiores publicações	107
4.2.6	Autores com maior publicação por áreas	109
4.2.7	Publicações em coautoria com o orientador	110
4.2.8	Análise da relação por meio das publicações científicas	111
5	CONSIDERAÇÕES FINAIS	113
	REFERÊNCIAS	116
	APÊNDICE A – EXECUÇÃO DO SOFTWARE AGLATTES	127

1 INTRODUÇÃO

Considerada um dos meios mais importantes para a disseminação do conhecimento, as publicações científicas também norteiam outros estudos, como a medição e a avaliação do desempenho de pesquisadores, realizadas através do número de publicações produzidas, nas citações recebidas e nas publicações resultante da coautoria (FARIA, 2001).

As publicações científicas são importantes por vários motivos. São através delas que os pesquisadores apresentam os resultados de suas pesquisas, principalmente para a comunidade científica. Outro ponto importante, é que os Programas de Pós-Graduação (PPG) são avaliados por métricas específicas e esses resultados norteiam as avaliações pelos órgãos de fomento e de forma significativa, garantem a continuidade de cada programa, o que faz com que os pesquisadores busquem aumentar a sua produtividade (DROESCHER e SILVA, 2014).

As atividades de pesquisa e desenvolvimento tecnológico, sem um vínculo com Instituições de Ciência e Tecnologia e sem a adaptação aos instrumentos de governança e financiamento gerenciado pelos órgãos públicos brasileiros, se tornam uma tarefa árdua (TORRES, 2018).

Logo, estudos sobre as métricas são necessários, pois por meio desses indicadores os programas podem se autoavaliar, medindo o desempenho de cada pesquisador e a sua contribuição para a Ciência, de forma que as pesquisas não sejam prejudicadas.

Neste contexto, se destaca a Bibliometria, uma das técnicas para mensurar a Ciência e que, apesar de ser uma prática antiga, ganhou grande interesse de estudos após as possibilidades do uso do computador. A tecnologia permitiu a evolução dessa técnica, o que marcou como um todo a Ciência da Informação (ARAÚJO, 2006a).

Atualmente, há diversos meios de analisar a Ciência, seja por análise bibliométrica da produção científica de pesquisadores, grupos de pesquisas ou instituições, os quais direcionam a forma pela qual esses resultados podem ser representados, de maneira que possam ser analisados tanto de forma qualitativa

como quantitativa. Os indicadores bibliométricos norteiam tanto os pesquisadores e os programas de pesquisas, como toda a evolução da Ciência de um país, o que é uma preocupação constante do meio acadêmico e de forma crescente, vem despertando o interesse da comunidade empresarial e de outros segmentos da sociedade (DROESCHER e SILVA, 2014).

No entanto, apesar da grande importância das métricas de produção, atualmente, a CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) - responsável pela consolidação do regulamento da pós-graduação no Brasil - entende que o principal “produto” de um PPG não são os artigos, nem os livros, nem as teses e dissertações, mas os mestres e doutores formados pelo programa e espera-se que darão continuidade na pesquisa e desta forma, também na área de estudo (RIBEIRO, 2007).

Para a CAPES (2019) um egresso de pós-graduação é todo portador de diploma de pós-graduação, mas sob a ótica de um PPG um egresso também significa a continuidade em suas pesquisas e descobertas, a renovação de seu quadro de docentes, a continuidade do programa através de suas gerações e a garantia de que as ideias, contribuições e visões de seus pesquisadores continuem influenciando o pensamento científico. Assim, não só os dados dos egressos se mostram importantes, mas também de toda a sua descendência através da propagação do conhecimento científico que se dá pela relação orientador e orientado.

A CAPES desempenha um papel fundamental na expansão e consolidação da pós graduação *stricto sensu* (mestrado e doutorado) em todo território brasileiro. Através de seu sistema de avaliação, busca manter o padrão de excelência acadêmica para os mestrados e doutorados nacionais. Os resultados da avaliação servem de base para a formulação de políticas para a área de pós-graduação, bem como para o dimensionamento das ações de fomento (bolsas de estudos, auxílios e apoios) (CAPES, 2019).

Desta forma, todos os cursos de Pós-Graduação *stricto sensu* devem atender o padrão CAPES, através de uma avaliação externa. Este padrão foi estabelecido em comum acordo com a comunidade acadêmica das diferentes áreas de conhecimento. Esta avaliação é realizada com êxito, principalmente no que diz respeito a regulação, porém a CAPES entende que a evolução desse processo deve ir além e reconhece que a autoavaliação dos PPG's pode trazer mais subsídios para

o desenvolvimento do sistema com qualidade. Acredita-se, que a autoavaliação favorece a construção da identidade, heterogeneidade e envolvimento dos programas avaliados, para além dos padrões mínimos garantidos pela avaliação externa e tem como resultado a melhoria dos próprios programas (CAPES, 2019).

Portanto, para a CAPES, a finalidade dos programas de pós-graduação não é apenas a produção de conhecimento, mas também a formação discente. Atualmente parte dos dados sobre egressos dos cursos de mestrado e doutorado também são avaliados pela CAPES, utilizando as bases de dados Rais-Caged, Coleta-Capes e Plataforma Sucupira (CAPES, 2017).

Fundamentadas nos princípios mencionados, em 2018 iniciaram as ações que ocasionaram mudanças nos instrumentos da avaliação da CAPES, com o objetivo de valorizar e aumentar a qualidade da formação de mestres e doutores. Uma dessas ações foi a diminuição dos quesitos e itens da ficha da avaliação, destacando os quesitos que ressaltam a qualidade dos programas, com a finalidade de proporcionar maior ênfase à formação e avaliação de resultados do que nos processos. A autoavaliação e o planejamento estratégico do PPG em relação ao seu programa, que consistiam em boas práticas na avaliação da CAPES, também passaram a ser considerados no processo de julgamento (CAPES, 2019).

Essa avaliação dos PPG's realizada pela CAPES ocorre periodicamente a cada quatro anos e para cada um desses quesitos as áreas devem propor as definições e indicadores que atendam às particularidades de cada área, acadêmica ou profissional (CAPES, 2019).

Agora são três quesitos de avaliação: Programa, Formação e Impacto na Sociedade. No quesito Programa, serão avaliados o funcionamento, estrutura e planejamento do PPG considerando os perfis e objetivos. No quesito Formação, serão analisados aspectos como qualidade das teses, dissertações, produção intelectual de alunos e professores, das atividades de pesquisas e avaliação dos egressos. E por fim, o quesito Impacto na Sociedade, serão considerados o caráter inovador da produção intelectual, os efeitos econômicos e sociais do programa, internacionalização e visibilidade (CAPES, 2019).

Contudo, se tratando de uma autoavaliação de um PPG, ter informações sobre seus egressos se torna muito difícil. Essa dificuldade aumenta quando a avaliação é de egressos internacionais (LIMA, 2015).

Geralmente quando se pretende avaliar os egressos, em uma autoavaliação,

as instituições de ensino obtém esses dados por meio de consulta a documentos administrativos, prontuários dos estudantes e acesso ao sistema de informação, como pode ser constatado em Pizzi e Santos (2017); ou ainda através de coletas manuais de dados dos currículos lattes como mostrou Noronha et. al. (2009) e que posteriormente, essas coletas foram aprimoradas através de *software* específico, o *script lattes*, como foi explorado por Batista (2019).

Porém, esse universo dos egressos de um PPG é muito amplo, de forma que o impacto de sua formação pode difundir-se a níveis difíceis de se conhecer. Assim, essa pesquisa busca meios de conhecer esse universo de forma automatizada e tem como amostra de estudos o Programa de Pós-Graduação em Engenharia de Materiais da Universidade Federal de São Carlos.

Diante disto, e considerando a dificuldade e o tempo elevado em obter dados dos egressos dos programas, visto que após a formação, muitos deles não mantêm vínculo com o programa, essa pesquisa buscou por estudos sobre Genealogia Acadêmica (GA), que segundo Damaceno et. al. (2019) tem o objetivo de compreender a Ciência através da relação orientador e orientado.

Damaceno et. al. (2019) explicam que as árvores genealógicas permitem obter uma melhor compreensão da configuração e evolução dos pesquisadores e dos respectivos grupos, auxiliando agências de financiamento e instituições acadêmicas para avaliar a formação dos recursos humanos, além de complementar a avaliação dos resultados da pesquisa. Embora as árvores acadêmicas serem objetos de estudos há mais de 20 anos, ainda há pouco uso das informações obtidas através dos dados das relações de conselho-orientador ou a supervisão acadêmica de pesquisadores ou entre professores e alunos.

Um dos estudos recentes no Brasil é a Plataforma Acácia, que tem como base de dados os Currículos Lattes dos pesquisadores. O *software* responsável pela GA, denomina-se AGLattes. Pesquisas sobre GA, utilizando o *software* AGLattes, têm sido mais realizadas com o propósito de se conhecer a GA de um único pesquisador a fim de constatar a importância e impacto para a Ciência individualmente.

O *software* apresenta algumas limitações para a sua aplicação, sendo uma delas, se tratando de estudos em PPG, a propagação da análise para pessoas que não são do programa, a qual foi contornada nesta pesquisa através de um novo procedimento.

Dado o exposto, por meio dos estudos realizados, do uso da tecnologia disponível e pela metodologia utilizada foi possível associar os resultados da GA com a Bibliometria, de modo a verificar se esses resultados trazem informações relevantes e se podem contribuir no atendimento aos quesitos da avaliação da CAPES e autoavaliação de um PPG por meio de indicadores.

A partir dos docentes do PPGCEM/UFSCar e pela associação da GA com a Bibliometria, foi possível chegar a informações que levariam muito tempo para serem obtidas manualmente e que podem, de fato, contribuir com os PPG's no atendimento aos quesitos da avaliação da CAPES, bem como serem utilizados na autoavaliação. O *software* AGLattes mapeou 993 egressos, os quais estão descritos neste trabalho como “descendentes diretos”, pois se formaram no programa e 1205 descendentes indiretos, que foram formados pelos descendentes diretos do programa.

Ademais, foi possível avançar os estudos além dos dados dos egressos, ampliando a busca para se conhecer a produção científica de todos os descendentes do programa, utilizando a ferramenta Lattes Machine, que permitiu também, conhecer mais a fundo a relação entre orientador e orientado, como por exemplo, a longevidade desta relação.

A mineração e análise de dados foi feita na ferramenta computacional VantagePoint e os resultados foram ilustrados por meio de indicadores na ferramenta Microsoft Excel, que evidenciam a atuação do PPGCEM/UFSCar durante os seus 40 anos de existência e o quanto os seus descendentes estão contribuindo, através de suas pesquisas e orientações, para a continuidade do programa e para o avanço da Ciência na área de Engenharia de Materiais e Metalúrgica.

1.1 OBJETIVO

Considerando a importância das métricas para a avaliação e atuação dos Programas de Pós-graduação, esta pesquisa tem por objetivo geral investigar se a aplicação da GA e da análise bibliométrica em conjunto pode trazer informações relevantes sobre a contribuição de um Programa de Pós-Graduação na formação

de pesquisadores e na produção de conhecimento.

A presente pesquisa tem como objetivos específicos:

- Explorar o processo de recuperação de dados a partir da Plataforma Lattes, através do *software* AGLattes, com foco na análise de PPG.
- Consolidar um procedimento útil de uso das ferramentas de GA e Bibliometria para a autoavaliação da atuação de Programas de Pós-Graduação.
- Estudar o Programa de Pós-Graduação em Engenharia de Materiais da Universidade Federal de São Carlos quanto à formação de recursos humanos e produção de conhecimento.
- Construir indicadores que contribuam para o atendimento dos quesitos da avaliação da CAPES e autoavaliação do Programa de Pós-Graduação em Engenharia de Materiais da Universidade Federal de São Carlos a partir da análise dos dados obtidos pela GA, quanto à formação de recursos humanos e produção de conhecimento.

Este estudo apresenta-se dividido em cinco capítulos, sendo o primeiro a introdução.

No segundo capítulo, são explanados os temas centrais da pesquisa, permitindo compreender a Ciência, o Conhecimento e como os investimentos contribuem para seus avanços, passando pela explosão informacional, que ocorreu após a Segunda Guerra Mundial, tendo como um de seus resultados uma nova área do conhecimento, a Ciência da Informação. Desta forma, neste mesmo capítulo busca-se entender como a Ciência é medida e avaliada no Brasil por meio das agências de fomento. Em seguida, o estudo avança na compreensão da Bibliometria e da GA, enunciando brevemente a história da Engenharia no Brasil, para então, chegar no PPGCEM/UFSCar, momento em que se pode conhecer a trajetória desse programa.

Com isso, pode-se entender o contexto da pesquisa e adentrar nos procedimentos utilizados para atingir os objetivos deste estudo e que são descritos no terceiro capítulo.

Posteriormente, no quarto capítulo, são apresentados os resultados, por meio de indicadores, que foram divididos em Indicadores de Formação de Recursos Humanos e Indicadores de Publicações Científicas, onde se pode confirmar que a

associação entre GA e a Bibliometria traz informações relevantes sobre a atuação de um Programa de Pós-Graduação e que podem contribuir para sua autoavaliação e no atendimento dos quesitos da CAPES.

Por fim, no último capítulo, são apresentadas as considerações finais.

2 REVISÃO DA LITERATURA

De acordo com Echer (2001) a revisão da literatura é imprescindível para a elaboração de um trabalho científico e que se inicia antes da definição do tema e vai até quando o pesquisador se sente familiarizado com os textos. À medida que o pesquisador avança na revisão da literatura se torna seletivo e se aprofunda em suas buscas. Ter o conhecimento sobre o assunto a ser trabalhado facilita a finalização do trabalho.

Neste capítulo encontra-se todo o conteúdo estudado para o embasamento teórico sobre o tema e a metodologia a ser utilizada. Foram utilizadas diversas fontes como artigos científicos, sites, dissertações, teses e livros.

2.1 CIÊNCIA E A SOCIEDADE

O homem, desde a sua origem, buscou conhecer os objetos que o cercam. A princípio foi motivado pela necessidade de controle dos fenômenos naturais para garantir a própria sobrevivência biológica e posteriormente, pelo próprio conhecimento adquirido, se tornando um sistema de autoalimentação. Acredita-se que o conhecimento alcançado através da Ciência é de maior confiabilidade quando comparado a exemplo com o conhecimento empírico, pois teorias, métodos, técnicas e produtos tem uma maior aprovação geral quando considerados científicos (CHIBENI, 2019).

Diante disso, muitos utilizam o termo “científico” com o objetivo de se valorizar algo, como é o caso das indústrias para valorizarem os processos de produtos, bem como os testes os quais são submetidos e muitas atividades de pesquisas que também se autoqualificam como “científicas” com o objetivo de afirmar-se. A consolidação da Ciência, deve-se na sua maior parte, principalmente, ao grande sucesso prático alcançado pela física, pela química e pela biologia, assumindo de forma implícita ou explícita, que ao se referir como Ciência, se entende que trata-se de um “método” especial, uma “receita”, quando seguida, resulta em conhecimento certo e seguro (CHIBENI, 2019).

O surgimento da Ciência no sentido usual do termo moderno envolveu muito mais do que a questão do método de investigação da natureza, onde o

conhecimento científico não se resume à observação sistemática e registro de fenômenos apenas pela teoria. A Ciência moderna vai além do desenvolvimento de novos procedimentos de investigação, como a descoberta de novos fenômenos e, principalmente, pelo desenvolvimento de novas teorias capazes de explicá-los (CHIBENI, 2019).

Alguns pesquisadores são afetados diretamente pela necessidade da sociedade e sob uma ótica diferente consegue encontrar soluções, muitas vezes simples, para problemas de grande dimensão. Como o pesquisador, em anatomia patológica, Ignaz Semmelweis, considerado um pioneiro nos procedimentos antissépticos, que diante de um grande número de mortes de parturientes por febre puerperal, propôs como solução a lavagem das mãos pelos profissionais que tinham contato com essas mulheres (RODRIGUES e MATTOS, 2015).

Percebe-se então, que identificar e entender os fenômenos da natureza é a principal finalidade da Ciência, sendo considerada um instrumento indispensável na incessante busca de respostas que satisfaçam as mais diversas necessidades de toda a sociedade (DROESCHER e SILVA, 2014).

A compreensão do mundo de forma mais profunda é uma necessidade humana. Isso, somado à necessidade de troca de informações, tem como resultado à elaboração de sistemas mais estruturados de organização do conhecimento (ARAÚJO, 2006b).

E desde o seu surgimento, cada vez mais, a Ciência desempenha um papel extremamente importante para a sociedade, criando conhecimento e melhorando a qualidade de vida das pessoas na medida que possibilita avanços nos campos da saúde, da alimentação, do ambiente, da tecnologia entre outros (UNESCO, 2019).

A definição de conhecimento pode ser determinada, inicialmente, através de três dimensões, onde a primeira trata-se de um mecanismo de compreensão e transformação do mundo, qual remete o ser humano à reflexão sobre o mundo cultural em que vive. A segunda dimensão, como uma necessidade para a ação, já que todos os atos são acompanhados de pensamentos, de reflexões sobre o observado, o sentido e o vivido. A necessidade de conhecer e compreender o mundo em que se vive, além de simplesmente vivê-lo, como uma ação diferenciada de pensar, é o movimento humano de dar significado a tudo, onde o conhecimento é o resultado da prática com o mundo e só tem sentido pleno na relação com a realidade (LUCESI et al., 2012).

O conhecimento clareia a ação humana sobre o mundo e as coisas. O fato de agir com pensamento e compreensão faz parte, exclusivamente, da essência humana, de forma, que é impossível o homem se relacionar com o mundo sem uma adequação cognitiva sobre a realidade que o cerca, momento em que o conhecimento é libertador, desde que este atue de acordo com as necessidades e vontades de todos os sujeitos envolvidos em seu processo, caso contrário, quando o conhecimento opera a serviço de determinados grupos com interesses próprios, passa a ser um instrumento de opressão (TOZONI-REIS, 2010).

A compreensão da Ciência torna-se mais fácil após a delimitação das outras formas de conhecimento como o religioso, o artístico e do senso comum. Esses se diferem do conhecimento científico, que por sua vez busca compensar as limitações dos anteriores através de métodos, técnicas e comprovações (ARAÚJO, 2006b).

O autor supracitado explica a diferença entre os conhecimentos:

- O conhecimento religioso oferece como desafio compreender uma verdade que já está pronta, através de um conjunto de conhecimentos que já se apresenta organizado, sistematizado, com regras, hierarquias e leis, ao invés de conhecer e produzir verdades sobre o mundo, evoluindo muito lentamente porque não permite a verificação devido sua transcendência, ou seja, que excede os limites normais, sendo entendido como ato da iluminação divina.
- O conhecimento artístico, por sua vez, consiste numa forma de conhecimento não objetiva, qual não se propõe a ser uma verdade e nem propõe explicações. Apesar de possuir métodos e técnicas, como o conhecimento científico, eles se diferem pois o conhecimento artístico acontece de forma espontânea, dinâmica e aberta através da arte com infinitas interpretações.
- O conhecimento de senso comum, busca contemplar as formas de conhecimento filosófico, que estuda a natureza e os limites das diferentes manifestações do conhecimento humano: A ideologia, que tem como forma de conhecimento composta de enunciados que justificam relações de poder, sendo definida como uma dimensão do conhecimento que pode estar presente em todas as formas de conhecimento, onde todo conhecimento pode funcionar como discursos ideológicos em determinados contextos.

No conhecimento científico há uma relação direta entre a Ciência e as atividades de pesquisas. A Ciência como busca de conhecimento e sistematização dos fatos, proporciona ao homem condições mais precisas para o entendimento das questões cotidianas, já a pesquisa é um processo que auxilia a Ciência em entender os fatos que acontecem diante das relações sociais quais estão em constante mudanças (SOUZA, 2011).

A pesquisa é uma ação de conhecimento da realidade, um processo minucioso e sistemático de investigação, que tem por objetivo conhecer a realidade ou alguns aspectos da realidade ainda desconhecidos, ou seja, é um processo de produção de conhecimentos para a compreensão de uma dada realidade, de forma que, as explicações sobre o termo “pesquisa” estão sempre associadas a “conhecimento”(TOZONI-REIS, 2010).

Por meio da pesquisa, os pesquisadores vão, progressivamente, resolvendo problemas que tem como objetivo comum, melhorar a qualidade de vida das pessoas, possibilitando a transferência do saber. Desta forma, a pesquisa deve ser vista como um bem incontável, como uma motivação para o conhecimento diferencial e como a forma mais confiável de crescimento social (CORTELLI, 2018).

Para tanto, se faz necessário investimentos em pesquisas de todos os campos do conhecimento, inclusive em áreas que, em algum momento, podem não despertar interesse da sociedade, já que a natureza dos desafios da sociedade muda constantemente e é necessário que a sociedade esteja preparada para enfrentá-los (GLOBAL RESEARCH COUNCIL, 2019).

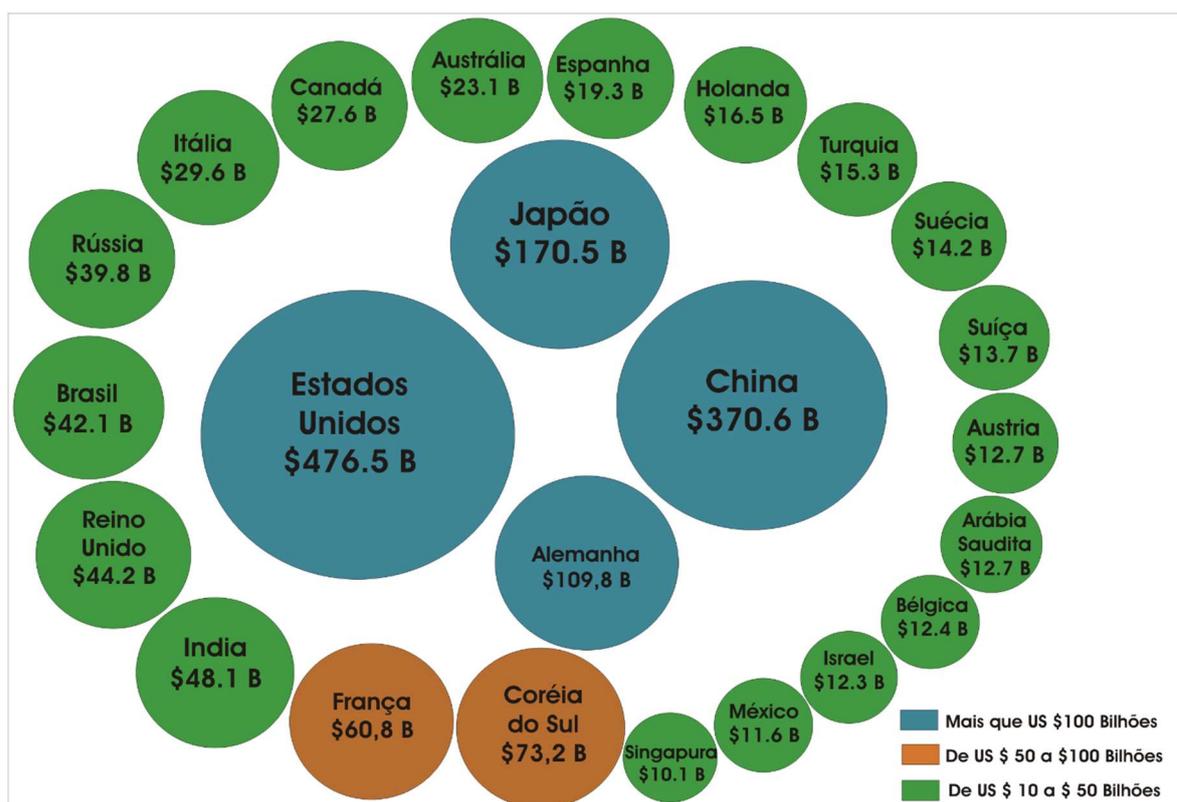
Qualquer economia desenvolvida pode ter a inovação como uma vantagem competitiva, porém a inovação é um resultado complexo e difícil de mensurar, tendo muitas variáveis diferentes que a levam em consideração em nível nacional. Os gastos com pesquisa e desenvolvimento (P&D) é uma dessas variáveis, representando tempo, capital e esforço investidos na pesquisa e no design dos produtos do futuro e estes nem sempre se correlacionam diretamente com os resultados da inovação (DESJARDINS, 2018).

O desenvolvimento de um país está relacionado com os gastos em P&D. Este estímulo para o desenvolvimento, envolvem uma série de atividades, incluindo resultados de pesquisa básica e aplicada, lançamentos de novos produtos, bem como a formação de pesquisadores e profissionais qualificados e essas atividades podem ser realizadas por empresas, universidades e outras instituições científicas

(MARQUES, 2019).

O infográfico desenvolvido para a plataforma howmuch.net, que tem como base de dados os resultados de uma série de pesquisas regionais menores realizadas pelo Instituto de Estatística da Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO), demonstram esses investimentos. Essas pesquisas comparam os números de investimento em P&D, no ano de 2018, para grande parte dos países do mundo, como pode ser observado na Figura 1. Neste levantamento estão inclusos todo tipo de gasto, desde pesquisa em inteligência artificial a invenção de novos equipamentos, até construção de equipamentos de caça de ponta (AMOROS et al., 2018).

Figura 1: Investimento em P&D por país



Fonte: Adaptado de Amoros et al. (2018).

Apesar da importância das atividades em P&D para um país e para a sociedade como um todo, de acordo com o Quadro 1, apenas dez países são responsáveis por cerca de 80% (oitenta por cento) dos investimentos em pesquisa

e desenvolvimento no mundo inteiro.

Quadro 1: Os dez países que mais investem em P&D

Posição	Países	US \$ (bilhões)
1º	Estados Unidos	476,5
2º	China	370,6
3º	Japão	170,5
4º	Alemanha	109,8
5º	Coréia do Sul	73,2
6º	França	60,8
7º	Índia	48,1
8º	Reino Unido	44,2
9º	Brasil	42,1
10º	Rússia	39,8

Fonte: Adaptado de Amoros et al. (2018).

Alguns países se destacam quando é feita uma análise dos seus gastos em P&D de acordo com o respectivo PIB (Produto Interno Bruto) como pode ser observado no Quadro 2, que demonstra um ranking dos 10 países que mais alocam a maior parte de suas economias em P&D.

Quadro 2: Países que investem maior % de seu PIB em P&D

Posição	Países	P&D (% PIB)
1º	Coréia do Sul	4,30%
2º	Israel	4,20%
3º	Japão	3,40%
4º	Suíça	3,20%
5º	Finlândia	3,20%
6º	Áustria	3,10%
7º	Suécia	3,10%
8º	Dinamarca	2,90%
9º	Alemanha	2,90%
10º	Estados Unidos	2,70%

Fonte: Adaptado de Desjardins (2018)

De acordo com Financiadora de Estudo e Projetos - FINEP (2019) Israel se

destaca em P&D, sendo considerado o país mais inovador do mundo. Situado em um território relativamente pequeno, onde a vegetação predominante é desértica, sem fontes de água suficientes para a sua população, o que torna esse país tão inovador é o contínuo investimento em P&D, onde se concentram inúmeros cientistas, engenheiros, patentes médicas e profissionais em tecnologia, o que o tornou líder mundial nas tecnologias de dessalinização, as quais são exportadas para mais de 40 países. Além disso, Israel conta com uma boa cooperação entre instituições de pesquisar públicas e empresas.

Na contramão de Israel, está o Brasil, que apesar de estar entre os 10 países que mais investiu em P&D no ano de 2018, de acordo com Grossmann (2019), vem reduzindo drasticamente seus investimentos neste setor e em 2019, quando houve uma redução de 42% (quarenta e dois por cento) no orçamento, deixou a comunidade científica brasileira bastante indignada.

A pesquisa é a principal atividade responsável pelo desenvolvimento da humanidade nos mais diversos aspectos, mas que infelizmente os nossos governos não dão a devida importância. Por mais difícil que esteja a situação de um país, não é a pesquisa que tem que ser sacrificada, pois é ela que pode tirar a sociedade de uma situação precária. Se hoje vivemos em um mundo repleto de tecnologia é porque houve pesquisa (TSALLIS, 2019, p.01).

O investimento público em Ciência e tecnologia deve ser de forma consistente e permanente, já que estes elementos asseguram o desenvolvimento do país. Um dos grandes problemas para o apoio ao desenvolvimento científico e tecnológico no Brasil ainda é a flutuação dos investimentos nesse setor. A maior parte das atividades científicas brasileiras só terá seu apoio garantido e continuado, somente, diante da estabilidade dos investimentos a nível nacional, proveniente do governo central (SILVA; TUNDISI, 2018).

2.2 A COMUNIDADE E A COMUNICAÇÃO CIENTÍFICA

Um grupo social formado por indivíduos que tem como profissão a pesquisa científica e tecnologia, forma uma comunidade científica. Dentro dessas comunidades, o principal sistema é o de doação onde, geralmente, o pesquisador

transfere para a sua comunidade as informações de sua pesquisa, sem qualquer interesse monetário. Em contrapartida, a comunidade científica, oferece a esse pesquisador a confirmação de que ele é um cientista, tornando-o reconhecido conforme os trabalhos desenvolvidos (LE COADIC, 1996).

Segundo o mesmo autor, cada pesquisador destas comunidades científicas trabalha em determinadas instituições de natureza social e econômica, como academias, sociedades científicas, associações de pesquisadores, laboratórios e universidades, que têm como função predominante a comunicação científica.

A invenção da imprensa por Johannes Gutemberg, por volta do ano de 1430, no século XV foi uma das grandes revoluções na modernidade, pois acelerou a produção de livros, um processo que era totalmente manual e artesanal e que demorava um grande tempo (FERNANDES, 2019).

Porém a invenção da imprensa só foi possível pela invenção e refinamento das técnicas de fabricação do papel, como o de farrapos na China em 105 d.C., feito com fibras vegetais e trapos velhos que era mais barato que as outras opções, como os papeis de seda (BACELAR, 1999).

Durante o Renascimento, meados do século XIV e o fim do século XVI, houve ampla repercussão das obras de grandes filósofos, como Aristóteles. Essas ideias, que estavam em manuscritos, foram registradas inicialmente em papiro, pergaminho e posteriormente em papel, iniciando então, o registro das informações (BARROS, 2018).

Historicamente, a comunicação científica teve origem no século XVII, quando houve a publicação das primeiras revistas científicas, no ano de 1665. Neste mesmo período surgiram os fenômenos que influenciaram a consolidação de uma sociedade leitora e a institucionalização da Ciência (BIBLIOTECA DA FEUP, 2019).

A escrita e suas decorrências foi a responsável por trazer o homem ao patamar em que ele se encontra hoje, como um ser social, científico e pensador. A escrita é um dos maiores instrumentos de divulgação e perpetuação de conhecimento. A propagação massiva de conhecimento permitiu ao homem que alcançasse a sabedoria necessária à sua sobrevivência (LUCESI et al., 2012).

Os pesquisadores podem tornar acessíveis os resultados do seu trabalho de investigação nas mais diversas áreas do conhecimento, através de uma publicação científica, que consiste na divulgação dos resultados da investigação, que podem ser comunicados sob diferentes formas de publicações (BIBLIOTECA DA FEUP,

2019).

De acordo com Medeiros (2013), os pesquisadores precisam escrever para transmitir o resultado de suas atividades, de suas pesquisas, de seus conhecimentos. São consideradas publicações científicas:

- **Artigo científico:** Trata de problemas específicos e apresenta o resultado de estudos e pesquisas, permitindo que as experiências sejam repetidas. Geralmente, é publicado em revistas, jornais ou outro periódico especializado. O seu conteúdo é variado e apresentam, na maioria das vezes, abordagens atuais e as vezes temas novos. Podem ser classificados como analíticos, classificatórios ou argumentativos e a sua redação leva em conta o público a que se destina.
- **Comunicação científica:** Pode ser definida como a informação que se apresenta em congressos, simpósios, reuniões, academias, sociedades científicas. Tem como finalidade divulgar e comunicar a descoberta e os resultados da pesquisa. Já os requisitos básicos para a divulgação científica são o conhecimento daquilo que se comunica, a precisão terminológica, a acessibilidade da linguagem e a adaptação à audiência.
- **Ensaio:** Trata-se de uma exposição metódica dos estudos realizados e das conclusões originais a que se chegou após apurado exame de um assunto. Pode ser formal e informal.
- **Informe científico:** Se apresenta como o mais breve dos trabalhos científicos, já que se restringe à descrição dos resultados alcançados pela pesquisa, de forma parcial ou total. A sua escrita deve ser compreendida e a parte experimental ser repetida, se for do interesse do investigador.
- **Monografia:** É uma dissertação que trata de um assunto particular, de forma sistemática e completa, onde a revisão da literatura é suficiente e tem como resultado a assimilação de conteúdo, de confecção de fichamentos e sobretudo, de reflexão, o que não exclui a capacidade investigativa de conclusões ou afirmações dos autores consultados. Comumente, o termo monografia é utilizado para identificar trabalhos científicos realizados por graduandos e apresentados ao final de um curso. Todavia, isto é um equívoco, já que tanto os trabalhos científicos a nível de mestrado e doutorado constituem-se numa monografia, porém com exigências diferentes. Qualquer trabalho científico é dissertativo e, se

apresentado para a obtenção do grau de mestre ou de doutor, deverá ser monográfico.

- **Dissertação:** No campo universitário, dissertação assumiu sentido mais restrito, que é a defesa de um assunto para se obter o título de mestre. É um trabalho mais de natureza reflexiva que, propriamente, de descobertas ou de ideias originais onde espera-se a sistematização do conhecimento, contudo, com observações e interpretações pessoais, limitando-se ao estado da arte.
- **Tese:** A tese é o trabalho científico apresentado para obtenção do grau de doutor, onde o autor deve demonstrar a capacidade de fazer avançar a área de estudo a que se dedica. O trabalho deve ser original e deve apresentar uma descoberta ou contribuição para a Ciência de forma que ultrapasse o estado da arte.
- **Paper:** O paper é uma síntese de pensamentos aplicados a um tema específico. Diferente de um relatório, em um paper espera-se que o autor faça uma avaliação ou interpretação de fatos ou de das informações que foram recolhidas, com expressão de pensamento, de forma original.
- **Relatório de Pesquisa:** Descreve de forma objetiva os fatos que ocorreram na pesquisa, onde o pesquisador faz análise desses fatos, para chegar a conclusões ou tomar decisões. O relatório deve responder as seguintes questões: O quê? Por quê? Para quê? Para quem? Onde? Como? Com que? Quanto? Quando? Quem? Com quanto?

Depois que uma pesquisa é publicada, um outro pesquisador pode acessar essas informações e utilizá-las como subsídios para construir um novo conhecimento e assim sucessivamente. É essa retroalimentação contínua, que faz a Ciência evoluir. Porém, para que esse processo aconteça, de forma que, um novo conhecimento gerado seja utilizado, é necessário à sua validação por pares. A avaliação é um princípio fundamental para a Ciência, sendo essa a responsável pela aceitação de artigos para publicação, a aprovação de projetos, a concessão de apoio financeiro, a criação e manutenção de um curso ou Programa de Pós-Graduação, entre outros (BARROS, 2018).

Atualmente os artigos científicos ganham destaque, tendo grande relevância para a Ciência e são publicados em periódicos científicos nas diferentes disciplinas

e áreas temáticas, constituindo-se como o principal meio de comunicação dos resultados de pesquisa científica. Os primeiros periódicos científicos surgiram a mais de 350 anos e são um dos principais meios de comunicação dos resultados de pesquisa científica (SPINAK e PACKER, 2015).

E assim, ao longo da história, propiciada pelos próprios cientistas como Johannes Gutemberg, a comunicação passou de oral a escrita. E a consequência disso foi a multiplicação da informação, ou seja, um conhecimento inscrito (gravado) sob a forma escrita (impressa ou numérica), oral ou audiovisual e que cresce exponencialmente (SPINAK e PACKER, 2015).

O livre fluxo de informação e sua justa distribuição tem sido um sonho de diversos homens em diversas épocas. Mas foi após a segunda guerra mundial que um grande volume de informação foi liberado (BARRETO, 2007).

A informação, mantida secreta no período de guerra, seria colocada à disposição do mundo e junto com o grande volume de informação, vieram os entraves, como a formação dos recursos humanos adequados para lidar com o volume de informação, os equipamentos de armazenamento e recuperação da informação existente e a estrutura teórica existente, que não explicava ou solucionava as práticas de informação da época (BARRETO, 2007).

Esses problemas da informação em Ciência e tecnologia e os possíveis obstáculos foram levantados pelo Doutor Vannevar Bush em 1945. Busch foi designado pelo presidente Roosevelt, como o responsável pelo Comitê Nacional de Pesquisa, nos anos de 1938 a 1942 e uma de suas missões foi congregar cerca de 6.000 cientistas americanos e europeus para direcioná-los ao esforço da guerra (BARRETO, 2007).

Vannevar Bush era engenheiro elétrico e administrador. Desenvolveu o Analisador Diferencial e supervisionou a mobilização governamental de pesquisas científicas durante a Segunda Guerra Mundial. Ele e seu laboratório de pesquisa se tornaram os principais designers e construtores de computadores analógicos. Como organizador de pesquisas científicas e engenheiro americano, teve participação no direcionamento do financiamento do governo para a investigação científica, que apresentavam custos exorbitantes e que antes disso, muitos experimentos eram impraticáveis. Posteriormente, esse sistema de financiamento e direcionamento da pesquisa científica, através das forças armadas, ficou conhecido como sistema do Pentágono, ou complexo industrial militar (DENNIS, 2019).

Em 1945 o Doutor Vannevar Bush era uma das pessoas mais bem posicionadas do mundo para especular o futuro da Ciência e da tecnologia. Sob sua supervisão, os dois maiores projetos científicos foram desenvolvidos, o radar e a bomba atômica. Ao final da guerra, ele definiu a estruturação do sistema de pesquisa norte-americano, que teve e ainda tem, enorme impacto sobre a organização da atividade científica em muitos países (CRUZ, 2011).

Em uma de suas obras, Bush, especula sobre o que a Ciência e a tecnologia poderiam trazer à humanidade em tempos de paz, depois do que havia sido feito em tempo de guerra, tendo como foco, como os avanços da pesquisa poderiam vir a modificar a forma de se pensar e organizar o conhecimento (CRUZ, 2011).

Desta forma, tem sua atenção voltada para os instrumentos de registro e transmissão de informação, que ele considera ser um dos principais desafios para os cientistas: Como ler e entender tantos artigos e relatórios e acessar tantas informações e ali selecionar o que é relevante, já que os instrumentos que ele tinha à disposição ainda eram o papel, lápis e fichários (CRUZ, 2011).

No mesmo ano, o Doutor Vannevar Bush indicou uma mudança no paradigma para a área de informação em Ciência e tecnologia, indicando uma explosão informacional, que envolvia seus profissionais, seus materiais de trabalho e a falta de condições teóricas para embasar a representação da informação para processamento, armazenagem e recuperação. Diante disso, surge a formação do profissional de informação (BARRETO, 2007).

Bush introduziu a noção de associação de conceitos ou palavras para organização da informação, pois este seria o padrão que o cérebro humano utiliza para transformar informação em conhecimento. Indicou que os sistemas de classificação e indexação existentes à época, eram limitativos e não intuitivos. Os processos para armazenar e recuperar informação, deveriam ser operacionalizados por associação de conceitos “como nós pensamos” (BARRETO, 2007, p.10).

Em 1946 foi realizada em Londres a Conferência Científica do *Royal Society Empire*, onde foi discutido a importância da informação e que alavancou a realização da Conferência de Informação Científica da *Royal Society*, que foi realizada em 1948, com duração de dez dias úteis e reuniu 340 cientistas. As propostas para resolver os problemas da organização e acesso à informação surgiram de vários cientistas de quase todas as áreas do conhecimento. A nova

área foi criada com o nome de: Ciência da Informação (BARRETO, 2007).

2.3 A CIÊNCIA DA INFORMAÇÃO

Durante todo o decorrer da história, nas mais diversas áreas, o homem tem buscado formas para classificar, registrar, organizar e difundir a informação. Para tanto, foi criada uma área específica para tratar de problemas relativos à informação, enquanto um fenômeno social. Desta forma, surge a Ciência da Informação, com a perspectiva de atender a necessidade de se organizar o conhecimento e a disseminação da informação (SILVA e FREIRE, 2012).

O campo da Ciência da Informação é interdisciplinar e, assim como outros, originou-se, a partir revolução científica e técnica que se seguiu à Segunda Guerra Mundial. Porém, este campo originou-se em processo de emergência, diante da explosão informacional pós-guerra e ainda está em evolução. Todos os esforços, de cientistas e engenheiros de todo o mundo para conter a grande quantidade de informação, foram financiados por vários programas estratégicos e por empresas privadas, o que teve como um dos resultados o desenvolvimento da moderna indústria da informação (SARACEVIC, 1996).

Diferentemente, a Ciência da Informação não surgiu a partir do desenvolvimento de outros campos de estudos como a psicologia ou a união de dois campos como a bioquímica, mas a partir da necessidade de aplicação no campo prático, como a documentação e a recuperação de informação (WERSIG; NEVELING, 1975).

A introdução de novas tecnologias, principalmente o processamento eletrônico de dados foi o que alavancou o seu surgimento diante das necessidades, porém as contribuições para o nascimento da “Ciência da informação” vieram de muitas disciplinas diferentes, já que eram muitas pessoas envolvidas, de diversas áreas de aplicação, para resolver o problema da explosão informacional, que entre algumas dessas se destacam, a Ciência da computação, a biblioteconomia, a filosofia e taxonomia, a linguística, a Cibernética e a Matemática. As discussões sobre o novo campo continuaram e posteriormente, outras disciplinas como Ciências Sociais e Teoria da Comunicação se aderiram (WERSIG; NEVELING,

1975).

O referido campo ainda apresenta fraca autonomia, pois não está totalmente estruturado e solidificado, o que ainda permite influências externas e a torna de caráter interdisciplinar, já que necessita de outras áreas para a sua atuação, tendo como desafio, além das abordagens sobre a natureza, manifestações e efeitos da informação e conhecimento, a busca de soluções tecnológicas que garantam a comunicação e uso. Quando se trata da informação como objeto, inclui todo o ciclo (produção, organização, armazenamento, representação, disseminação, recuperação, acesso e o uso) da informação (NHACUONGUE e FERNEDA, 2015).

O advento da internet acelerou a comunicação entre os cientistas, de forma que o acompanhamento e o registro do avanço da Ciência foram promovidos pelas facilidades das tecnologias de informação e comunicação (SPINAK e PACKER, 2015).

Houve um tempo em que a informação era privilégio de poucos, protegidas por muros, vigiada constantemente e onde o saber ficava sob custódia. Depois de muito tempo, o avanço da tecnologia quebrou as barreiras ao acesso à informação. Através da internet grande maioria das informações é liberada por completa em sua linguagem natural (BARRETO, 2007).

Mas, segundo o mesmo autor, o ideal do acesso ao conhecimento livre e para todos não surgiu com a internet, já que o livre fluxo de informação e sua distribuição por igual, foi um sonho de muitos homens em diversas épocas. Tanto que Paul Otlet, advogado e documentalista, o qual sonhava em facilitar o acesso do maior número de pessoas à informação graças a um complexo conjunto de bibliotecas conectadas por canais telegráficos e telefônicos.

Barreto (2007) explica que essa rede de conhecimento iniciou através da enciclopédia que se configura bem o sentido de rede de conhecimento distributivo. A primeira rede do saber acumulado foi a Encyclopédie, que teve seu trabalho iniciado em 1750 e os volumes finais publicados em 1772, na França, compreendendo 28 volumes, 71.818 artigos e 2.885 ilustrações.

De acordo com o mesmo autor, Paul Otlet foi um pioneiro da organização das redes mundiais de tratamento e difusão da informação registrada, tendo grande contribuição da documentação e posteriormente para a Ciência da informação, área que sistematiza teorias, métodos e técnicas para organizar o conhecimento registrado e redistribuí-lo. Muitos outros estudiosos, se embasaram em Otlet para

fundamentar suas ideias.

Para Otlet os métodos da documentação são os mesmos das outras Ciências, sendo a documentação uma Ciência tão abstrata e própria para constituir sistemas documentais por meio do raciocínio como as matemáticas. A aplicação da matemática à documentação é feita por meio da estatística e da bibliometria (SANTOS, 2007).

2.4 A AVALIAÇÃO DA PÓS-GRADUAÇÃO NO BRASIL E A CAPES

Para se fazer Ciência é necessário recursos e investimentos, de forma que a pesquisa brasileira é financiada por diferentes sistemas e instituições de fomento, que estão ligadas direta ou indiretamente aos ministérios brasileiros (MARQUES, 2019).

De acordo com o Sistema de Bibliotecas da Universidade de São Paulo (2018) entre as principais organizações de financiamento à pesquisa brasileira estão:

- Conselho Nacional de Desenvolvimento Científico (CNPq): Com sua sede em Brasília, financia projetos de pesquisa, bolsas e programas de cooperação internacional.
- Fundações Estaduais de Amparo à Pesquisa (Agências estaduais que constituem as FAPs): Financiam projetos de pesquisas, bolsas e programas de cooperação internacional. Estão entre elas a FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) e a FAPERJ (Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro) dirigidas pelo CONFAP (Conselho Nacional das Fundações Estaduais de Amparo à Pesquisa) que agrupa todas as 27 fundações públicas de amparo à pesquisa, sendo que apenas o Estado de Roraima não possui uma fundação pública de amparo à pesquisa.
- Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES): Agência de financiamento e de avaliação do Ministério da Educação. Financia bolsas e programas de cooperação internacional.
- Financiadora de Estudos e Projetos (FINEP): Com sede no Rio de Janeiro, direciona investimentos a projetos tecnológicos em parceria

com empresas. Tem como fonte de recursos o FNDCT (Fundo Nacional de Desenvolvimento Científico e Tecnológico) e Fundos Setoriais.

O ato de fornecer recursos, geralmente financeiros, para sustentar uma atividade, programa ou projeto de um pesquisador, grupo ou instituição são obtidos a partir de agências governamentais, conselhos de pesquisa ou entidades filantrópicas. O destino desses recursos é determinado a partir de criteriosas consultas para determinar prioridades, com o objetivo de contemplar novas demandas e campos científicos, de acordo com o poder de investimento e a necessidade de alocação de recursos (DUDZIAK, 2018).

De acordo com o mesmo autor as decisões quanto ao destino desses recursos são difíceis e controversas, visto que as avaliações estão, na maior parte, centradas na competição de projetos e no impacto das citações, resultando no fato de que, algumas áreas recebem mais financiamento que outras e descobertas maiores nem sempre são as mais amparadas financeiramente

Atualmente, a Ciência vive uma crise no cenário brasileiro, onde houve uma redução drástica nos investimentos governamentais em pesquisa, obrigando o pesquisador se tornar mais competitivo para pleitear também recursos internacionais e não ter a sua pesquisa prejudicada. O CNPq e a CAPES, principais agências de financiamento da pesquisa, também passam pela escassez de recursos, afetando a concessão de bolsas e apoios (DUDZIAK, 2018).

De acordo com a portaria Nº 182, de 14 de agosto de 2018 publicada no Diário Oficial da União, as propostas de cursos novos e os Programas de Pós-Graduação *stricto sensu* em funcionamento serão avaliados pela CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) e dependerão do alcance do padrão mínimo exigido para entrada e permanência no Sistema Nacional de Pós-Graduação.

A CAPES foi fundada pelo Ministério da Educação (MEC) em 11 de julho de 1951 e desde então desempenha um papel fundamental na expansão e consolidação da Pós-Graduação *stricto sensu* (mestrado e doutorado) em todos os estados do Brasil, e tem como principal objetivo assegurar a existência de pessoal especializado em quantidade e qualidade suficientes para atender às necessidades dos empreendimentos públicos e privados que visam ao desenvolvimento do país,

tendo o seu papel fortalecido com a tarefa de coordenar a avaliação da pós-graduação. Esse programa, além de contribuir para a criação de mecanismos efetivos de controle de qualidade, aprofunda sua relação com a comunidade científica e acadêmica (CAPES, 2019).

Em 2017 o Sistema Nacional da Pós-Graduação (SNPG) contava com 4.175 programas de pós-graduação, sendo esses 2.128 programas possuem mestrado e doutorado (51%), 1.270 programas dedicados exclusivamente ao mestrado acadêmico (30%), 703 mestrados profissionais (17%) e 74 Programas de Pós-Graduação que possuem apenas o curso de doutorado (2%) (CAPES, 2017).

Esses programas são avaliados pela CAPES em um intervalo de 4 anos, quais são atribuídas notas de acordo com o seu desempenho. Os objetivos desta avaliação é certificar a qualidade da Pós-Graduação Brasileira, o que é uma referência para distribuição de bolsas e recursos para o fomento à pesquisa e identificar assimetrias regionais e de áreas estratégicas do conhecimento para orientar as ações nesse setor (CAPES, 2017).

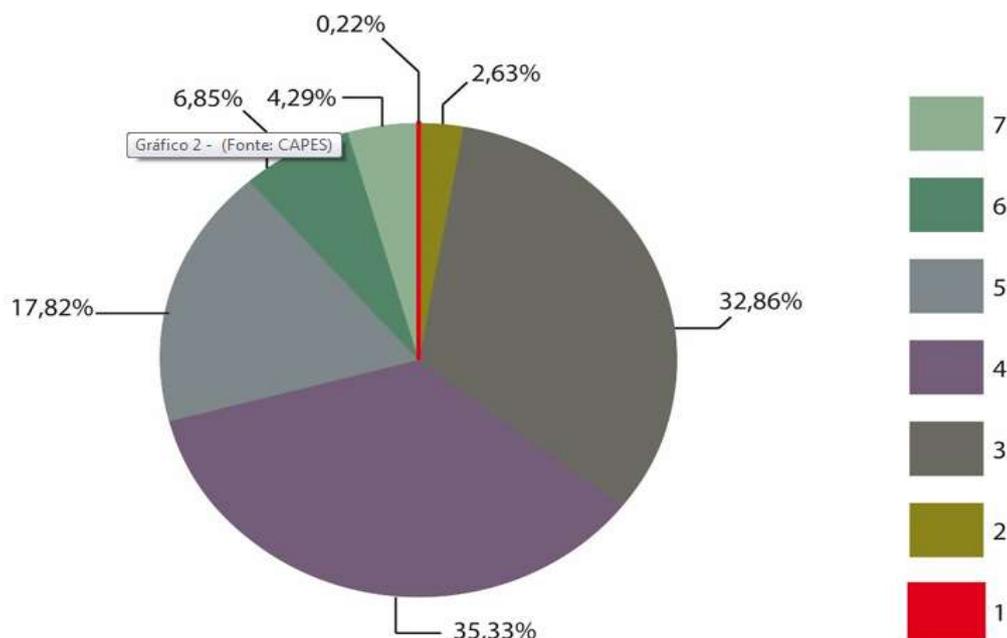
De acordo com a CAPES (2017) os critérios para avaliação periódica estão disponíveis no site da CAPES nos Documentos Orientadores das áreas de avaliação. Após a avaliação periódica, cada programa em funcionamento receberá apenas uma nota, na escala de 1 (um) a 7 (sete). As notas 7 (sete) e 6 (seis) equivalem a padrões internacionais de excelência; já as notas 5 (cinco) e 4 (quatro) significam um desempenho entre muito bom e bom alcançando o nível de excelência nacional e para os programas que atingem um desempenho médio é atribuída a nota 3 (três). A nota inicial de recomendação para os programas abertos recentemente ou que participam pela primeira vez da avaliação é de 3(três) e 4(quatro). Serão desativados os programas que receberem nota inferior a 3 (três). A avaliação é atividade essencial para assegurar e manter a qualidade dos cursos de Mestrado e Doutorado no país.

Dos 4.175 programas, 465 deles possuem notas 6 e 7. A avaliação quadrienal da CAPES de 2017 revelou que os PPG's mantiveram suas qualidades, sendo que a maioria dos 4.175 programas avaliados (67%) tiveram suas notas mantidas; 22% aumentaram as notas e apenas 11% apresentaram queda nos índices (CAPES, 2017).

Por meio da Figura 2, é possível observar a distribuição das notas dos Programas de Pós-Graduação avaliados pela CAPES, onde apenas 4,29%

obtiveram a nota 7 (sete) e a maioria obtiveram nota 4 (quatro).

Figura 2: Distribuição de notas dos Programas de Pós-Graduação



Fonte: (CAPES, 2017)

Para um PPG, manter e melhorar sua nota perante a avaliação da CAPES é de extrema importância para sua sobrevivência. Desta forma, esses programas devem estar atentos quanto suas atuações e produção científica (CAPES, 2017).

Segundo Ribeiro (2007), o sistema de avaliação da agência atua em 45 áreas do conhecimento, apesar de haver uma filosofia comum a todas, cada área tem seus critérios de avaliação, que são principalmente direcionados para:

- **Produção científica:** Esse é o primeiro ponto, que compreende a qualidade da pesquisa e como esta é medida. Publicar uma pesquisa é um meio de difundi-la e de socializá-la. Assim, a produção científica dos PPG's também é avaliada pela sua qualidade e em várias áreas, pelo seu impacto, e não somente pelo "produtivismo". Se tratando de fator de impacto é uma média e mede como uma publicação contribui para constituir uma comunidade científica, não apenas a qualidade de artigos isolados, ou seja, artigos que não são citados. Em princípio, o fator de impacto se mede tomando-se uma revista científica, onde

quanto melhores os artigos maiores são as chances desta revista ser lida. Se a revista A publicou 20 artigos, que foram citados 100 vezes, seu fator de impacto é cinco. Mas, se nas outras publicações saiu apenas uma referência a algum daqueles artigos, e nada mais, o fator é 0,05 (ou seja, um dividido por 20). Se há citações, é sinal de que se está construindo um ambiente de diálogo, uma massa crítica, que permitirá que a área avance.

- Distribuição da produção científica: A distribuição da produção científica é um outro critério avaliado e ocorre quando há um equilíbrio na produção intelectual, ou seja, a produção intelectual de qualidade não está concentrada em poucos membros do programa e há o respeito para a diferença entre seniores e juniores, entre docentes mais maduros e mais novos e onde os alunos são orientados, de fato, por um pesquisador experiente e de forma que todos os docentes do programa respondem pela produção científica e que sejam bons no que fazem.
- Formação de mestres e doutores: Para a CAPES, o principal “produto” de um PPG não são os artigos, nem os livros, nem as teses e dissertações, mas os mestres e doutores formados pelo programa e que, espera-se, darão continuidade na pesquisa.
- Equilíbrio na formação de mestres e doutores: É vista de forma negativa quando pouco membros do programa conduzem as orientações de um curso. A CAPES, espera que o professor que pesquisa seja o mesmo que orienta, de forma que o orientado tenha a experiência real de pesquisa.

Os resultados obtidos da avaliação realizada pela CAPES podem ser usados para diversos fins, como auxiliar na decisão de um estudante em seus futuros cursos, contribuir com as agências de fomento nacionais e internacionais na orientação de suas políticas segundo as notas atribuídas da avaliação e servir de base para uma agenda de redução de desigualdades entre regiões do Brasil ou no âmbito das áreas do conhecimento (CAPES, 2017).

Os incentivos às atividades de pesquisa brasileira ganham notoriedade a partir deste sistema de avaliação o que resulta em uma grande corrida dos PPG's,

na busca de metodologias e ferramentas que possam configurar-se em recurso estratégico para o acompanhamento de sua produção científica e contribuir para sua gestão, principalmente pelo fato de que a produção científica ser um dos itens de maior peso no sistema de avaliação da CAPES. Neste contexto, ressalta a importância da Ciência da Informação (CI), que tem também como foco de seus estudos, o fluxo informacional, desde a produção, disseminação e o uso social da informação (MACIEL, 2018).

Em 2018, iniciaram as ações que ocasionaram mudanças nos instrumentos da avaliação da CAPES com o objetivo de aumentar a qualidade da formação de mestres e doutores. Uma dessas ações foi a diminuição dos quesitos e itens da ficha de avaliação, destacando os quesitos que ressaltam a qualidade dos programas, com a finalidade de proporcionar maior ênfase à formação e avaliação de resultados do que nos processos. A autoavaliação e o planejamento estratégico do PPG em relação ao seu programa, também, passaram a ser considerados no processo de julgamento (CAPES, 2019).

De acordo com a CAPES (2018) agora são três quesitos que compõem a avaliação dos PPG's brasileiros: Programa, Formação e Impacto na Sociedade.

O primeiro deles é o quesito Programa, que tem como pretensão avaliar o funcionamento, estrutura e planejamento do PPG em relação ao seu perfil e seus objetivos. Este quesito é composto por quatro itens:

- Articulação, aderência e atualização o das áreas de concentração, linhas de pesquisa, projetos em andamento e estrutura curricular, bem como a infraestrutura disponível em relação aos objetivos/missão do programa. Perfil do corpo docente, e sua compatibilidade e adequação à Proposta do Programa.
- Planejamento estratégico do programa, considerando também articulações com o planejamento estratégico da instituição, com vistas à gestão do seu desenvolvimento futuro, adequação e melhorias da infraestrutura e melhor formação o de seus alunos, vinculada à produção do conhecimento.
- Os processos, procedimentos e resultados da autoavaliação do programa, com foco na formação discente e produção do conhecimento.

O segundo quesito é Formação, onde o foco se volta para a qualidade dos recursos humanos formados, sendo consideradas a atuação dos docentes e a produção de conhecimento diretamente associada às atividades de pesquisa e de formação do programa. Este quesito é formado por três itens:

- Atuação dos docentes permanentes em relação às atividades de pesquisa e de formação do programa e à produção intelectual.
- Qualidade e adequação das teses, dissertações ou equivalente em relação às áreas de concentração e linhas de pesquisa do programa.
- Qualidade da produção de discentes e egressos.

Por fim o quesito Impacto na Sociedade, busca verificar os impactos gerados pela formação de recursos humanos e a produção de conhecimentos do PPG. Neste quesito avaliam-se a relevância social, internacionalização e inovações, bem como o destino e atuação dos egressos. Este quesito é formado por quatro itens:

- Impacto e caráter inovador da produção intelectual – bibliográfica, técnica e/ou artística, em função da natureza do programa.
- Destino, atuação e avaliação dos egressos do programa em relação à formação recebida.
- Impacto da inserção social e econômica do programa.
- Internacionalização e visibilidade do programa.

Acredita-se que esta pesquisa possa contribuir na elaboração de indicadores para atender os três quesitos.

Uma outra e importante avaliação da CAPES é o Qualis-periódicos.

De acordo com a CAPES (2019) o Qualis-Periódico trata-se de uma ferramenta que tem por objetivo classificar a produção científica dos programas de pós graduação no que se refere aos artigos publicados em periódicos científicos. Essa avaliação é feita baseada nas informações fornecidas por meio do módulo Coleta de Dados da Plataforma Sucupira. Por meio desta coleta, a CAPES disponibiliza uma lista com a classificação dos veículos utilizados pelos Programas de Pós-Graduação para divulgação da sua produção.

Essa estratificação da qualidade da produção dos programas é realizada de forma indireta, o que permite atestar a qualidade dos artigos e de outros tipos de

produção a partir dos periódicos científicos, que tem a classificação realizada pelas áreas de avaliação, ou seja, somente são avaliados aqueles periódicos que receberam produções dos alunos dos programas, utilizando critérios gerais e específicos de cada área de avaliação da CAPES e que são disponibilizados nos respectivos Documento de Área. Essa avaliação estabelece oito estratos – A1, A2, B1, B2, B3, B4, B5 e C – para classificar a produção intelectual publicada em periódicos científicos, sendo A atribuído como nível mais elevado e C como nível mais baixo (CAPES, 2019).

2.5 BIBLIOMETRIA

De acordo com a Plataforma Sucupira (2020) a Ciência da Informação é uma área da Ciências Sociais Aplicadas e possui três subáreas: Biblioteconomia, Arquivologia e Teoria da Informação. A Bibliometria é uma especialidade da Biblioteconomia e de acordo com Faria (2001) é uma técnica de contagem de toda e qualquer informação escrita. A etimologia da bibliometria consiste das palavras gregas para “livro” (biblos) e “Medida” (metron) (SUGIMOTO, 2014).

A maioria das descobertas científicas e resultados de pesquisas são publicados em revistas científicas, anais de congressos científicos ou em sites. Esses trabalhos na maior parte em formato de artigos científicos, geralmente são avaliados por especialistas da mesma área e depois de publicados, são disponibilizados a outros pesquisadores e devem ser citados como referências. O número de citações de um artigo em particular é um reflexo de seu impacto na comunidade científica (GEVENOIS e DURIEUX, 2010).

Segundo os mesmos autores, essas citações fornecem dados que podem ser usados para medir, estatisticamente e matematicamente, a importância relativa de um determinado artigo ou uma revista específica, bem como a conexão entre campos científicos, departamentos de pesquisa ou autores.

Os métodos utilizados nessa medida são coletivamente conhecidos como bibliometria que se baseia na aplicação de métodos matemáticos e estatísticos a livros, artigos e outros meios de comunicação para medir sua quantidade e desempenho, ou seja, a qualidade (GEVENOIS e DURIEUX, 2010).

Já no início do século XX, cientistas começaram a perceber que a quantidade de conhecimento científico gerado, continuamente, ultrapassa a sua capacidade de leitura. O volume do conhecimento gerado através de novos periódicos e publicações, diante da impossibilidade de adquirir todos, levaram os profissionais da informação, a buscarem maneiras para selecionar os mais significativos, como a cientometria, infometria e bibliometria (FARIA, 2001).

A bibliometria, qual também ganha destaque neste estudo, é uma área de estudo da Ciência da informação e tem grande importância na análise da produção científica de um país, já que seus indicadores representam o nível de desenvolvimento de um campo científico. Trata-se de uma técnica de análise baseada na quantificação da informação, que tem como objeto de estudo as publicações científicas em diversos níveis de interesse e é usada para avaliar o resultado da produção científica de pesquisadores e estudiosos (FARIA, 2001).

A bibliometria pode ser definida como uma técnica quantitativa e estatística de medição dos índices de produção e disseminação do conhecimento científico, que surge no início do século XX, a partir da necessidade do estudo e da avaliação das atividades de produção e comunicação científica, tendo o seu desenvolvimento inicial na elaboração de leis empíricas sobre o comportamento da literatura. O método de medição da produtividade de cientistas de Lotka (1926), a lei de dispersão do conhecimento científico de Bradford (1934) e o modelo de distribuição e frequência de palavras num texto de Zipf (1949) são uns dos principais marco do desenvolvimento da bibliometria (ARAÚJO, 2006a).

Com o passar dos anos a bibliometria extrapolou os limites das bibliotecas e passou a, também, considerar revistas acadêmicas e artigos de periódicos, autores e instituições, referências bibliográficas, citações, agradecimentos, patentes e posteriormente, com o surgimento da internet, a bibliometria ganha um reforço para verificar as produções dos acadêmicos, se suas obras têm ou não impacto ao longo do tempo e os efeitos disso (SUGIMOTO, 2014).

Se tratando de impacto das produções acadêmicas é preciso se atentar que impacto, que é o que efetivamente as citações medem, é ambíguo de qualidade, ou seja, a qualidade do artigo não tem uma relação direta apenas com o número de vezes que o um autor é citado e sim com outros fatores como o número de artigos publicados pelo autor, o idioma dos artigos, a cultura de cada comunidade científica entre outros (FARIA, 2001).

Segundo Sugimoto (2014) o interesse pela bibliometria, amplamente interpretado, nunca foi tão grande pelas corporações, universidades e órgãos de financiamento. E esse interesse não é por acaso. Um dos critérios utilizados por governos e órgãos multinacionais na decisão para direcionar seus recursos destinados à pesquisa e ao desenvolvimento é a avaliação do trabalho científico através de indicadores quantitativos.

A produção científica tornou-se um instrumento, que está sendo utilizado, para decidir para onde serão destinados os recursos à pesquisa, recursos que são limitados e compete com os demais setores de investimento público (VELHO, 2011).

A bibliometria ao fornecer novas informações, sobre o estado da Ciência produzida por um país ou por um grupo, pode auxiliar a tomada de decisões e o gerenciamento de pesquisas, mas que sozinha, não justifica uma decisão ou substitui especialistas. O termo bibliometria, tornou-se genérico para toda uma gama de medidas e indicadores específicos com a finalidade de medir a produção de pesquisa científica e tecnológica, não apenas através da literatura científica, mas também das patentes (OKUBO, 1997)

A bibliometria, sempre tem como fonte, um banco de dados. Geralmente esses dados são extraídos, ainda brutos, de uma base de dados estabelecida por empresas ou por instituições públicas ou privadas. Após processamento especial, esses dados, agora em forma de indicadores bibliométricos, passam a ilustrar os resultados da atividade da Ciência e tecnologia e se demonstram como ferramentas práticas, que podem ser usadas em combinação com outros indicadores, com o objetivo de avaliar o estado atual da Ciência (OKUBO, 1997)

2.5.1 Base de dados: Plataforma Lattes

Bases de dados são sistemas que armazenam e disseminam as informações relacionadas aos registros das publicações científicas, resultando em uma enorme variedade de dados, que pode ser específica de uma área do conhecimento, de uma região ou de uma instituição (BASSOLI, 2017).

A popularização desses repositórios de produção científica ocorreu no mundo inteiro e tornaram-se indispensáveis para a preservação e visibilidade da

informação científica, contrapondo o modelo de comunicação científica tradicional, que se estabelece na publicação de artigos e trabalhos científicos por revistas especializadas, onde o acesso ocorre por meio de assinatura paga pelos interessados, dificultando o acesso à informação produzida (MATIAS, 2015).

No Brasil há uma base de dados que concentra dados de pesquisadores de todo o país, trata-se da Plataforma Lattes, do CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) que armazena e disponibiliza os currículos de pesquisadores, docentes e discentes de todo o país, juntamente, desta forma, com os dados da produção científica nacional (MATIAS, 2015).

A partir dos anos 80 houve a preocupação dos dirigentes do CNPq quanto a utilização de um formulário padrão para registro dos currículos dos pesquisadores brasileiros. Além de, através deste formulário, permitir a avaliação curricular do pesquisador, o CNPq tinha como objetivo criar uma base de dados que possibilitasse uma série de levantamento de dados e geração de estatísticas sobre a distribuição da pesquisa científica no Brasil.

Nesta época, quando a internet ainda não se fazia presente no Brasil, este formulário de captação de dados era em papel e posteriormente os dados eram digitados em um sistema informatizado e somente em 1990, que os formulários eletrônicos para a captação de dados curriculares passaram a ser desenvolvidos (CNPq, 2019a).

Já no final da década de 80, as universidades e instituições de pesquisa do país já tinham acesso à base de currículos de pesquisadores brasileiros, que contava com aproximadamente 30.000 currículos. No ano de 2016, somente de pesquisadores, a base de dados da Plataforma Lattes contava com mais de 130.000 registros, como mostra o Quadro 3 e isso vem aumentando consideravelmente (CNPq, 2019a).

Quadro 3: Crescimento do número de Currículos Lattes

Principais Dimensões	1993	1997	2002	2006	2010	2014	2016
Instituição	99	181	268	403	452	492	531
Grupos	4.402	8.632	15.158	21.024	27.523	35.424	37.640
Linhas de Pesquisas	ND	ND	50.473	76.719	106.715	139.141	147.392
Pesquisadores	21.541	33.980	56.861	90.320	128.892	180.262	199.566
Pesquisadores Doutores	10.904	18.724	34.349	57.586	81.726	116.427	130.140

Fonte: Adaptado de (CNPQ, 2019b).

A adoção do Currículo Lattes pelas principais universidades, institutos, centros de pesquisa e fundações de amparo à pesquisa dos estados como instrumento para a avaliação de pesquisadores se tornando um padrão nacional no registro da vida pregressa e atual dos pesquisadores do país contribuiu para o aumento da quantidade de currículos na plataforma. Qualquer pessoa pode pesquisar um Currículo Lattes e ter acesso a informações públicas e individuais de cada usuário cadastrado na plataforma (MOREIRA; DIAS e MOITA, 2016).

O Currículo Lattes é adotado pela maioria das instituições de fomento, universidades e instituto de pesquisa do país. Apresenta ricas, confiáveis e abrangentes informações referentes a vida pregressa e atual dos estudantes e pesquisadores do país. Por esse motivo, atualmente é um elemento obrigatório para a análise de mérito e competência das ações de financiamentos na área da Ciência e tecnologia (CNPq, 2019c).

Diante disto, esta fonte de dados passou a ser o próprio objeto de estudos de muitos pesquisadores, que analisam como esses dados podem contribuir para a construção de indicadores bibliométricos (BASSOLI, 2017).

Um dos problemas para realizar estudos bibliométricos através dos currículos Lattes, está na realização ou sumarização desses dados a partir de um grande grupo de usuários cadastrados de médio ou grande porte, já que essa ação requer um grande esforço mecânico e que geralmente é passível de falhas (ROSSI; DAMACENO; MENA-CHALCO, 2018).

Como pode ser observado Figura 3 atualmente a plataforma Lattes contabiliza mais de 6,5 milhões de currículos.

Figura 3: Currículos indexados na Plataforma Lattes em 2019



Fonte: CNPq (2019)

2.5.2 Análise da produção científica a partir de Indicadores

A Ciência pode ser considerada como um sistema de produção de informação, quando publicada em formatos permanentes e disponível para uso comum. Então, deste modo, a Ciência pode ser vista como uma empresa com insumos e resultados, ou seja, entradas e saídas. A medição dessas duas categorias é a base dos indicadores científicos. Há grandes esforços da Ciência científica concentrados no desenvolvimento de metodologias apropriadas para a construção desses indicadores (SPINAK, 1998).

De acordo com Haddad, Mena-Chalco e Sidone (2018), indicadores bibliométricos abrangem métricas de publicações e citações e têm sido cada vez mais utilizados para quantificar e qualificar a produção científica de pesquisadores. São classificados em:

- Indicadores quantitativos ou de produção: Que se referem a contagens de itens da produção bibliográfica;
- Indicadores qualitativos ou de citações: Medem o desempenho da produção de um pesquisador e buscam refletir diferenças nas estratégias de atuação dos pesquisadores;
- Indicadores estruturais ou de ligação: Podem ser calculados a partir das relações de coautorias endógenas, que medem as conexões entre autores por meio de colaborações científicas e identificam a importância de cada pesquisador na rede endógena, que considera apenas as colaborações de docentes internos de uma instituição.

Os indicadores de produção são desenvolvidos pela análise quantitativa das publicações que podem ser a partir dos tipos de publicação (livros, capítulos, artigos etc.), por instituições vinculadas, por países de origem, por área do conhecimento ou por ano. Esses indicadores podem ser a respeito dos números absolutos das publicações, e de percentuais de crescimento, divisão por autores, por estratificação dos periódicos publicados, entre outros. Os Indicadores de citação são desenvolvidos pela análise de citações recebidas pelas publicações. Eles tentam, principalmente, mensurar o alcance, a influência, o impacto dos trabalhos publicados no campo científico. E os Indicadores de ligação são desenvolvidos pela quantificação de co-ocorrências de alguns dados das publicações científicas –

autores, palavras-chave, citações e outros – para mapear redes de colaboração científica e cooperações internacionais (FARIA et al., 2010).

Fora das universidades e institutos de pesquisas, os Indicadores de Ciência, Tecnologia e Inovação (CT&I) também se destacam como recursos fundamentais para gestores, em especial na tomada de decisões mais racionais e sustentáveis acerca das políticas de Ciência e tecnologia e inovação do país (NIT MATERIAIS, 2019).

Esses indicadores também são utilizados para definir as políticas públicas do Brasil. O Estado de São Paulo foi pioneiro em reconhecer o papel vital da Ciência e da pesquisa tecnológica como política pública, já que considera esses elementos de extrema importância para a sociedade, tornando a pesquisa científica, um dever primordial do estado (LAFER, 2011).

Além do planejamento e execução de políticas, os indicadores de produção científica contribuem para uma melhor compreensão sobre a Ciência pela própria comunidade científica, pela comunidade empresarial e por outros segmentos da sociedade. A Ciência, tecnologia, inovação e competitividade são impulsionadas pela produção científica e os indicadores, resultado da medição desta produção, podem contribuir de diversas maneiras, como na análise da dinâmica das diferentes áreas científicas e se tratando de agências governamentais nacionais e internacionais de fomento à pesquisa, esses indicadores são utilizados na formulação, execução e acompanhamento de políticas públicas de Ciência e tecnologia (FARIA et. al., 2011).

Tanto é a importância dos indicadores de Ciência, tecnologia, inovação, que desde antes de 2001, a FAPESP (2019) vem consolidando o seu programa de Indicadores em São Paulo. O referido programa tem como objetivo apresentar estatísticas e indicadores que reflitam a situação recente e principais tendências da atividade científica e tecnológica no estado comparativamente aos esforços e resultados nacionais e internacionais.

Os resultados do programa da FAPESP citado, são divulgados por meio de publicações multitemáticas e panorâmicas, editadas a cada três ou quatro anos e também por meio de boletins estatísticos periódicos, onde pode-se conferir indicadores atualizados segundo as principais dimensões de interesse como: os dispêndios em pesquisa e desenvolvimento, a formação de recursos humanos para pesquisa e inovação, a produção e o impacto de publicações científicas, a atividade

de patenteamento no Brasil e no exterior entre outros.

2.6 GENEALOGIA ACADÊMICA

Durante a idade média, surgiram os sobrenomes ou nomes de família para a identificação das pessoas de um povo, o que antes disso, era somente utilizado pelos reis e nobres. Em uma época de grande expansão demográfica, com o intuito de buscar uma diferenciação, os homens mais comuns passaram a utilizar como sobrenomes as designações de seus ofícios ou habilidades, de seus lugares, de suas condições socioeconômicas, de plantas ou animais, adotando desta forma uma infinidade de nomeações que pudessem os identificar (INGESC, 2019).

Atualmente, muito além de mera designação, o sobrenome é um patrimônio da família que representa toda uma linhagem se estendendo por várias gerações, identificando características físicas e comportamentos semelhantes. Entretanto, essa descendência vai além do plano genético, se desenvolvendo-se também no campo histórico (INGESC, 2019).

É possível identificar várias formas de herança através do estudo das genealogias e, para isto, utilizam-se o heredograma, também conhecido como árvore genealógica, que é a representação gráfica dos indivíduos relacionados por parentesco (ICB, 2019).

As árvores genealógicas, que ilustram a recomposição das linhagens, se mostram de grande utilidade no processo de esquematização e demonstram figurativamente os descendentes de toda uma família. “A reconstrução histórica da formação familiar conduz, portanto, a interpretações capazes de estabelecer uma ponte entre o passado e o presente, entre os ancestrais e seus descendentes, revelando-se como a maior homenagem que se pode prestar aos antepassados” (INGESC, 2019).

Michel Foucault, inspirado em Friedrich Nietzsche, propôs aos historiadores a abordagem genealógica da histórica, concluindo que a abordagem genealógica serve para qualquer análise histórica. Aqueles que se interessam pela pesquisa de sua árvore genealógica, de modo geral, estão em busca de sua origem (FERREIRA, 2005).

Muitas vezes quem busca pela sua árvore genealógica está apenas interessado na sua linhagem. Preocupado com determinado sobrenome, procura

aquele que carrega em seu próprio nome ou aquele que julga ser o mais nobre. Os genealogistas profissionais sempre evocam a nobreza, como é o caso das empresas especializadas em brasões de família, que sempre enaltecem, qualquer família, como nobre e antiga (FERREIRA, 2005).

A utilização da genealogia para documentar indivíduos e seus relacionamentos têm como objetivo a obtenção de conhecimento sobre a origem, evolução e disseminação de grupos inter-relacionados. Essa tarefa de documentação auxilia o entendimento da formação e tendências futuras de grupos de pessoas (ROSSI; MENA-CHALCO, 2014).

De forma similar a descrita, caracteriza-se a GA, que tem como resultado as árvores genealógicas acadêmicas, que por meio de métricas topológicas, permite categorizar indivíduos através de sua linhagem acadêmica, possibilitando a obtenção de novos conhecimentos importantes para a compreensão do cenário científico de uma área (ROSSI; MENA-CHALCO, 2014).

Segundo Mena-Chalco (2016) GA é o estudo da herança intelectual que se resulta dos relacionamentos entre orientadores e seus alunos a nível de mestrado e doutorado e que através de uma representação gráfica, a árvore de GA (genograma acadêmico) permite compreender, mais facilmente, as redes de parentesco acadêmico, que podem ser caracterizadas pela forma como suas relações são organizadas e pode ser classificada como:

- Genealogia honorífica: Ênfase na descendência de um pesquisador para evidenciar o impacto nas gerações de acadêmicos.
- Genealogia egocentrista: Ênfase na ascendência de um pesquisador para evidenciar os ancestrais mais importantes.
- Genealogia histórica: Ênfase na contextualização histórica (análise histórica) do desenvolvimento de uma área. Geralmente considerados os “pioneiros/fundadores das áreas”.
- Genealogia paradigmática: Ênfase em como o conhecimento e práticas epistemológicas são transmitidas através das relações de orientação.
- Genealogia analítica: Ênfase na aplicação de métodos matemáticos e computacionais para criar e estudar as relações de orientação.

A GA não deve ser confundida com “Genealogia Intelectual” que é um termo usado para descrever qualquer tipo de influência, sem a necessidade de ter um

orientador formal ou institucional ou com a “Genealogia Científica” que é um termo usado há muito tempo para a investigação “histórica” da pessoa sob estudo e que possui relações com outras Ciências como a Biologia e a Estatística, diante das heranças físicas e psíquicas relacionadas a todos os ancestrais de duas famílias e suas combinações (SCHAEFFER, 1960); (SUGIMOTO, 2014).

Como estudo quantitativo do patrimônio intelectual operacionalizado através de cadeias de estudantes e seus orientadores, a GA deixou de ser criticada por ter baixos níveis de generalização e rigor, sendo usada apenas com o objetivo de homenagear estudiosos distintos ou satisfazer a curiosidade de um ilustrado sobre seus antepassados acadêmicos para contribuir com aqueles que estudam Ciência sob as perspectivas históricas, filosóficas, sociológicas e científicas, diante do aumento de fontes e meios para coletar dados em grande escala que possibilitam estudos rigorosos nessa área (SUGIMOTO, 2014).

A reprodução disciplinar é facilitada pela produção de novos estudantes de doutorado. No entanto, a produção deve ser seguida de retenção: os estudantes de doutorado não devem apenas completar seus estudos, mas permanecer na academia após a graduação para continuar o ciclo. O pressuposto implícito nas concepções de reprodução disciplinar é que os alunos se tornarão mestres na mesma disciplina em que foram treinados (SUGIMOTO et al., 2011).

Através das árvores acadêmicas, pode-se facilmente deduzir como são as relações pessoais entre os cientistas, por meio das relações entre orientador e aluno, explicar as alianças com respeito a colaborações ou até mesmo rivalidades e conflitos pessoais. Em termos práticos, pode explicar prontamente o sucesso na obtenção de posições acadêmicas e colocação acadêmica nas instituições (ANDRAOS, 2006).

Segundo o mesmo autor o estudo das árvores acadêmicas permite conhecer a cronologia das ideias de uma questão ou de uma área de pesquisa, associando-as aos principais atores na formação dessas ideias ou pelo menos de seus descendentes científicos diretos, voltando-se para o que, de fato, é importante e merecedor de consideração. Uma grande árvore do conhecimento é composta pelo conhecimento das conexões entre as ideias e o contexto e as circunstâncias em que as descobertas foram feitas.

Chegar nos criadores das descobertas e ler os seus trabalhos originais publicados, ao invés de se basear apenas na interpretação de outros autores é

possível se aprofundar nessas descobertas, entender como elas foram feitas, averiguando suas limitações e assim conhecer as conexões entre as ideias e o contexto do autor. As descobertas desses autores podem preparar o cenário de várias décadas de novas pesquisas. A pesquisa não deve fechar problemas e sim abrir novas áreas de pesquisas. Mesmo que uma ideia possa parecer sem esperanças em termos de compreensão no contexto atual, esta pode ser adequada para longos anos de investigação científica (ANDRAOS, 2006).

Segundo Rossi, Damaceno, Mena-Chalco (2018) as atividades de orientação, supervisão ou mentoria acadêmica promovem um relacionamento qual tem uma relação direta com a evolução do orientado, da instituição, da Ciência e da sociedade em termos gerais. Inúmeras iniciativas por parte das comunidades acadêmicas são realizadas, com o objetivo de documentar, analisar e classificar estruturas que facilitem a GA.

Analisar os relacionamentos de orientação acadêmica sob a forma de uma estrutura genealógica (e.g., grafo ou árvore de genealogia acadêmica), permite um maior entendimento sobre a comunidade científica, a caracterização do acadêmico por meio de seus relacionamentos e a identificação do impacto gerado por estes acadêmicos na formação dessa comunidade. (ROSSI; DAMACENO; MENA-CHALCO, 2018, p. 3).

O caminho para o trabalho de um cientista viver além de seu tempo é que eles preencham o próximo geração de acadêmicos com pessoas que eles orientaram para que suas ideias, contribuições e visões continuem influenciando o pensamento científico. A vida de um erudito é finita, mas sua contribuição é ampliada, reforçada e prolongada através de sucessivas gerações de aprendizes (ANDRAOS, 2006).

Desta forma a genealogia fornece um meio de medir e analisar a propagação de conhecimento científico, a obtenção de insumos quantitativos para mensurar as interações acadêmicas e a migração de área de um acadêmico. A migração de área por um acadêmico tem um impacto direto no cenário do conhecimento, mudando a trajetória esperada da próxima geração de mestres (SUGIMOTO, 2014).

Além disso, segundo David; Hayden (2012) a análise das genealogias acadêmicas fornece informações úteis sobre os ambientes, apontando os pesquisadores mais produtivos em suas carreiras posteriores.

Para Rossi; Mena-Chalco (2014) a análise genealógica acadêmica se mostra importante por apresentar a possibilidade de avaliar o impacto das orientações acadêmicas no desenvolvimento científico de áreas específicas do conhecimento e na identificação dos principais atores, ou grupo de maior relevância na proliferação do conhecimento.

Uma forma eficaz para extrair conhecimento em contextos, muitas vezes, de difícil interpretação é representar os relacionamentos de indivíduos na forma de redes e diferentes estudos foram dedicados à documentação, análise e classificação de árvores genealógicas acadêmica através de relacionamento de orientação (ROSSI; MENA-CHALCO, 2014).

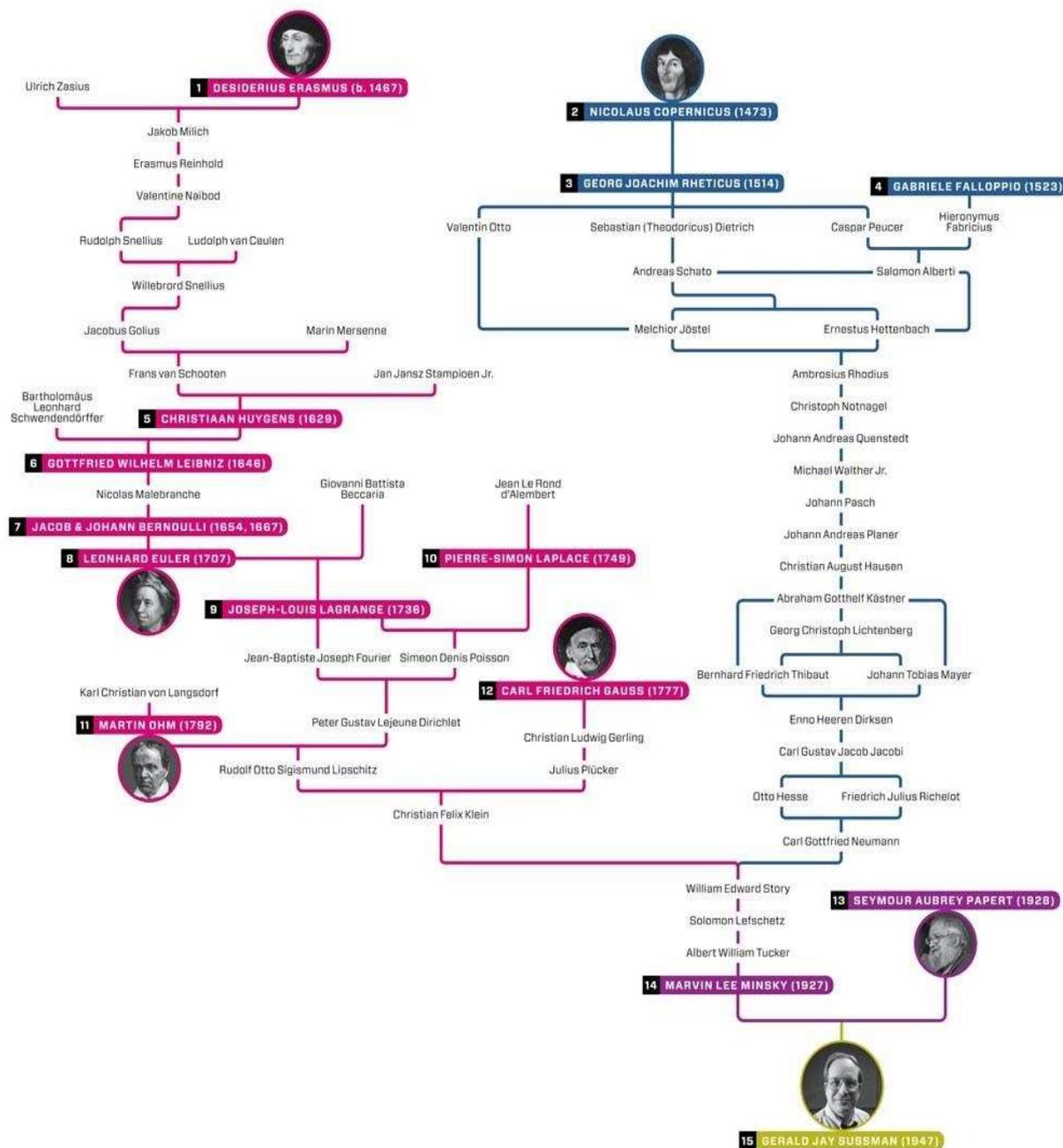
A GA oferece uma grande oportunidade para avaliar pesquisadores e suas áreas de pesquisa sob a perspectiva do treinamento de recursos humanos, de forma que algumas áreas específicas do conhecimento estão tomando diferentes medidas para estruturar suas genealogias acadêmicas, que podem ser analisadas como uma rede social, como o Projeto de Genealogia Matemática (PGM), que tem mais de 240.000 doutores em matemática registrados e também o NeuroTree, que foi projetado para a área na NeuroCiência, especificamente, com cerca de 120.000 acadêmicos registrados (MENA-CHALCO, 2016).

Para o entendimento, toma-se como base o Projeto de Genealogia Matemática que de acordo com Mulcahy (2017) este projeto foi realizado há vinte anos, em 1996, por Harry Coonce, então matemático da Universidade Estadual de Mankato, em Minnesota. Esta ferramenta através de seu banco de dados, tornou-se um padrão de referência para matemáticos em todo mundo, através de contribuições e correções de usuários, tendo como objetivo traçar a história intelectual das Ciências matemáticas, listando todos os indivíduos que receberam um doutorado em matemática. Para cada indivíduo, planeja-se mostrar o seguinte: O nome completo do destinatário do diploma; O nome da universidade que concedeu o diploma; O ano em que o grau foi concedido; O título completo da dissertação; O nome completo do (s) orientador (es).

De acordo com o autor supracitado, o PGM oferece, mediante pagamento, pôsteres personalizados de genealogia matemática, pois identificou que tanto os indivíduos quanto os departamentos com interesse em seu patrimônio gostam de ver os destaques exibidos nas árvores genealógicas. Esses pôsteres são comprados para presentear aniversários, aposentadoria, bem como presentes de

formatura, além de ser um modo de representar visualmente a herança intelectual de um departamento para os visitantes. Um exemplo desses pôsteres, pode ser observado na Figura 4

Figura 4: Pôster personalizado de GA oferecidos pelo PGM



Fonte: (ARBESMAN, 2011)

O portal do PGM possui uma interface simples, como pode ser observado na Figura 5 onde pode-se pesquisar qualquer outro matemático, inserindo apenas o nome.

Figura 5: Página inicial do Portal do PGM

https://www.genealogy.math.ndsu.nodak.edu/index.php

Mathematics Genealogy Project

Home
Search
Extrema
About MGP
Links
FAQs
Posters
Submit Data
Contact
Mirrors
Donate

A service of the [NDSU Department of Mathematics](#), in association with the [American Mathematical Society](#).

Quick Search [Advanced Search](#)

240831 records as of 1 April 2019
View the [growth](#) of the genealogy project

Search About MGP Links FAQs Posters Submit Data Contact

The Mathematics Genealogy Project is in need of funds to help pay for student help and other associated costs. If you would like to contribute, please [donate online](#) using credit card or bank transfer or mail your tax-deductible contribution to:

Mathematics Genealogy Project
Department of Mathematics
North Dakota State University
P. O. Box 6050
 Fargo, North Dakota 58108-6050

Fonte: PGM (2019)

Ao pesquisar por Carl Friedrich Gauss, um importante matemático alemão, reconhecido pelas suas contribuições a essa área como a teoria dos números, geometria, teoria da probabilidade, teoria das funções entre outras (GRAY, 2019) o portal do PGM demonstra a sua fecundidade intelectual, trazendo muitas informações, como a sua instituição de ensino, o ano, o país, o antecessor e os sucessores como pode ser observado na Figura 6. De acordo com o PGM (2019) o último descendente deste matemático é Andreas Bartels no ano de 2014.

Figura 6: Resultados da pesquisa feita no PGM

https://www.genealogy.math.ndsu.nodak.edu/id.php?id=18231

Mathematics Genealogy Project

Carl Friedrich Gauß

[Biography](#)

Ph.D. Universität Helmstedt 1799 

Dissertation: *Demonstratio nova theorematum omnium functionum algebraicarum rationalium integram unius variabilis in factores reales primi vel secundi gradus resolvi posse*

Advisor 1: [Johann Friedrich Pfaff](#)

Students:
Click [here](#) to see the students listed in chronological order.

Name	School	Year	Descendants
Bessel, Friedrich	Georg-August-Universität Göttingen	1810	39429
Dedekind, J. W. Richard	Georg-August-Universität Göttingen	1852	
Encke, Johann			11596
Gerling, Christian	Georg-August-Universität Göttingen	1812	64798
Germain, Sophie	Georg-August-Universität Göttingen	1830	
Gräffe, Carl	Georg-August-Universität Göttingen	1825	
Klinkerfues, Wilhelm	Georg-August-Universität Göttingen	1855	
Listing, Johann	Georg-August-Universität Göttingen	1834	1810
Riemann, Bernhard	Georg-August-Universität Göttingen	1851	2244
Schumacher, Heinrich	Georg-August-Universität Göttingen	1806	91
Spehr, Friedrich Wilhelm	Georg-August-Universität Göttingen	1824	
Stern, Moritz	Georg-August-Universität Göttingen	1829	
Svanberg, Gustaf	Georg-August-Universität Göttingen and Uppsala Universitet	1825	2
von Staudt, Karl	Friedrich-Alexander-Universität Erlangen-Nürnberg	1822	
Wichmann, Moritz	Georg-August-Universität Göttingen	1843	

A service of the [NDSU Department of Mathematics](#), in association with the [American Mathematical Society](#).

Fonte: Elaborada pela autora no portal do PGM

Portanto, qualquer pessoa que atua na área da matemática pode consultar os seus antecessores e verificar se na sua linhagem tem algum desses nomes notáveis.

O PGM também oferece outras contribuições, como auxílio na coleta de informações e referências para artigos, auxiliando a aprendizagem sobre matemáticos, também, para aqueles que não são da área da matemática, sendo futuramente, uma ferramenta inestimável para os historiadores da matemática, estendendo sua utilidade para cientistas sociais e agências governamentais. Se tratando de publicações científicas, o PGM permite identificar, mais facilmente, quem são as pessoas e onde elas se encaixam no espectro matemático, evitando

possíveis conflitos de interesse. Depois, há editores de revistas acadêmicas, os quais muitos concordam quanto à utilidade do PGM, ao proporcionar uma imagem de quem são as pessoas e onde elas se encaixam no espectro matemático (MULCAHY, 2017).

Naturalmente, existem limites para a utilidade do PGM. A inclusão e atualização das informações dependem dos usuários e nem sempre todas as pessoas concordam que seu supervisor oficial era seu verdadeiro pai ou mãe matemáticos, resultando em um relacionamento inexato. Além disso, qualquer pessoa pode enviar dados para o PGM, sendo difícil comprovar as informações, já que às vezes nem as teses em questão nem os registros oficiais correspondentes sobrevivem (MULCAHY, 2017).

No Brasil há um projeto com a mesma finalidade do PGM, a Plataforma Acácia, que possui uma interface simples de ser utilizada, como pode ser observada na Figura 7.

Figura 7: Página Inicial da Plataforma Acácia

A **Plataforma Acácia** foi concebida com o intuito de **documentar** as relações formais de orientação no contexto dos programas de pós graduação brasileiros. Isto é feito utilizando dados disponibilizados pela **Plataforma Lattes**, que atualmente concentra mais de 6,3 milhões de currículos acadêmicos. A documentação é feita sob a forma de grafos de genealogia acadêmica, em que cada **vértice** representa um **pesquisador** e cada **aresta** uma **relação de orientação concluída** entre dois pesquisadores (**orientador e orientado**). Para maiores informações sobre o processo de construção da plataforma, consulte a página **Sobre**.



Fonte: Plataforma Acácia (2019)

O projeto da Plataforma Acácia foi idealizado no ano de 2016, tendo seu lançamento em 2018 e que apesar de ainda estar em desenvolvimento, fornece informações precisas da GA brasileira a partir dos dados dos currículos que estão

indexados na Plataforma Lattes. Desta forma, não é possível inserir ou eliminar relações de orientação a partir desta plataforma (MENA-CHALCO et al., 2019).

Diferente da plataforma PGM, que depende das informações cedidas pelos usuários, a Plataforma Acácia tem suas relações de orientação atualizadas a partir de uma nova coleta de dados registrados nos currículos Lattes, sendo que a última atualização do site ocorreu em novembro de 2017 (MENA-CHALCO et al., 2019).

O acesso é livre e pode-se pesquisar uma única pessoa por vez, sendo que a pesquisa considera somente as relações de mestrado e doutorado. Na Figura 8, pode ser observado o resultado de uma pesquisa realizada na Plataforma Acácia. Para este exemplo, foi pesquisado pelo Prof. Dr. José Angelo Rodrigues Gregolin da Universidade Federal de São Carlos (UFSCar).

Figura 8: Resultados da pesquisa feita na Plataforma Acácia

Plataforma Acácia
Genealogia Acadêmica do Brasil

Pesquise mestres e doutores atuantes no Brasil

Acadêmicos: 1.272.590
Relações de orientação: 1.404.109

Digite o nome do pesquisado:

Jose Angelo Rodrigues Gregolin

Grande Áreaⁱ: Engenharias
Áreaⁱ: Engenharia De Materiais E Metalúrgica
Instituiçãoⁱ: Universidade Federal De São Carlos

Descendência (Ds)ⁱ: 74
Índice Genealógico (IG)ⁱ: 3
Fecundidade (Fc)ⁱ: 20
Fertilidade (Ft)ⁱ: 5
Gerações (G)ⁱ: 3
Relações (R)ⁱ: 74
Primos (Pr)ⁱ: 413

Ascendentes Descendentes

N	Nome	Orientações	Ds	IG	Fc	Ft	G	R	Pr
1	Adriana Regina Martin	P 2001	1	0	1	0	1	1	910
2	Adriana Tahereh Pereira Spinola	M 2006	0	0	0	0	0	0	443
3	Anderson Crovador Massura	M 2005	0	0	0	0	0	0	443
4	Andre Luis Castelo Branco Leite	M 2013	0	0	0	0	0	0	443

Fonte: Elaborada pela autora na Plataforma Acácia.

De forma organizada, a Plataforma Acácia disponibiliza várias informações sobre a pessoa pesquisada. Como pode ser observado do lado esquerdo da Figura 8 é possível identificar a Área, a instituição, o ano da titulação e do lado direito os seus descendentes. No nome de cada descendente está habilitado um link que faz

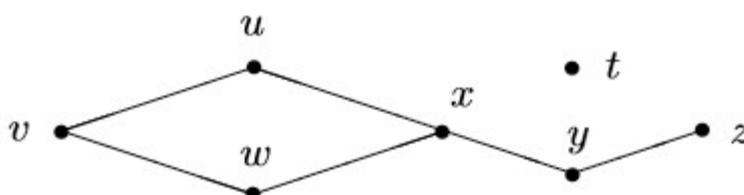
o redirecionamento para uma outra página onde são exibidas as mesmas informações do nome selecionado. Também é possível verificar os ascendentes de cada nome clicando na guia Ascendentes e realizar pesquisas por áreas na guia Análise de Áreas.

Um estudo realizado pelos desenvolvedores da Plataforma Acácia descreve como as árvores de genealogia podem ser caracterizadas por meio de métricas de avaliação de grafos, que caracterizam um indivíduo, onde se baseiam as informações fornecidas pela plataforma.

Em teoria dos grafos, um grafo é um par (V,A) em que V é um conjunto arbitrário de objetos denominados vértices e A é um subconjunto de pares não ordenados de V , denominadas arestas. Os vértices ou nó são as unidades fundamentais que forma os grafos (FEOFILOFF; KOHAYAKAWA; WAKABAYASHI, 2011).

A Figura 9 é um exemplo de um grafo, onde as vértices são t, u, v, w, x, y, z e as arestas são vw, uv, xw, xu, yz e xy .

Figura 9: Exemplo de grafo



Fonte: (FEOFILOFF; KOHAYAKAWA; WAKABAYASHI, 2011).

De acordo com Rossi e Mena-Chalco (2014), através das métricas é possível atribuir um valor numérico para qualificar um indivíduo pela topologia de sua árvore, obtendo informações ou padrões que auxiliem na compreensão a respeito da árvore do indivíduo, capacidade de propagação e diferenciação de outras árvores. Desta forma, utilizando cálculos matemáticos, os autores definem as seguintes métricas:

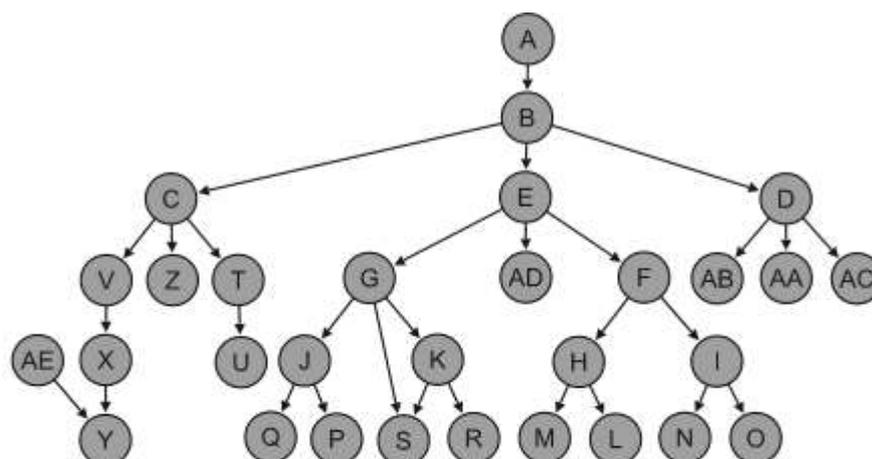
- Fecundidade (f): Esta métrica tem o objetivo de dimensionar a árvore através do número de vértices que ela representa. É importante para a classificação de um vértice raiz (orientador), baseando-se na quantidade de descendentes que este influenciou.

- Fecundidade ponderada (fp): De forma similar ao da fecundidade, esta métrica tem o objetivo de atribuir um peso maior para os vértices que estão mais próximos do vértice raiz, de forma que os relacionamentos diretos têm maior peso no cálculo do valor da métrica.
- Número de folhas (nf): Os números de folhas representam a quantidade de vértices não fecundos na árvore, ou seja, indivíduos que não orientaram nenhum aluno. Quanto maior o número de folhas maior é o potencial de crescimento da árvore.
- Profundidade (p): Fornece o grau de maturidade da árvore genealógica formada a partir de um vértice raiz, medindo a quantidade de arestas existentes entre o vértice raiz e um vértice mais distante que possa ser alcançado.
- Largura (l): Esta métrica mede a quantidade de relacionamentos diretos que um vértice raiz possui, ou seja, os descendentes diretos de um indivíduo, refletindo a produtividade (em termos de orientação) deste.
- Maior Largura (ml): Tem a função de identificar o maior número de relacionamentos em um mesmo nível da árvore genealógica, demonstrando a influência de um vértice na propagação dos relacionamentos em sua árvore.
- Distância Média (dm): Define a proximidade entre um vértice raiz e todos os vértices pertencentes à sua ascendência, ou seja, a média entre um vértice raiz e os vértices pertencentes a sua linhagem. Quanto menor for o valor da distância média, maior é a proximidade existente entre os vértices de uma árvore.
- Média dos menores caminhos (mmc): Tem o objetivo similar ao da Distância Média, diferenciando-se nos caminhos utilizados para o cálculo, ponderando os valores obtidos e representando a distância média entre os indivíduos pertencentes à árvore.
- Índice H (h): Tem por objetivo considerar a quantidade e a qualidade genealógica (no sentido de perpetuidade) dos relacionamentos dos vértices da árvore. O índice H genealógico (h) de um vértice é definido como o maior número h de relações que este vértice possui com

outros vértices que tenham pelo menos, o mesmo número de relacionamentos cada um.

Na Figura 10 é possível observar os resultados dos cálculos das métricas apresentadas acima para os vértices mais representativos de uma árvore. Como foi explicado acima, é possível entender as características de cada nó. Os cálculos realizados que resultam na Fecundidade Ponderada (fp), na Distância Média (dm) e na Média dos Menores Caminhos (mmc) podem ser conferidos no trabalho dos autores, já que neste momento não é do interesse deste trabalho se aprofundar nas fórmulas e sim nos seus conceitos.

Figura 10: Métricas aplicadas em uma árvore genealógica



Vértice	f	fp	nf	p	l	ml	dm	mmc	h
A	28	3,51	15	5	1	9	3,52	0,13	1
B	27	6,53	15	4	3	9	2,64	0,10	3
C	6	3,61	3	3	3	3	1,43	0,24	1
D	3	3,00	3	1	3	3	0,75	0,25	0
E	15	5,03	9	3	3	7	2,13	0,14	2

Fonte: (ROSSI; MENA-CHALCO, 2014)

Assim como os desenvolvedores do Projeto Plataforma Acácia, há outros pesquisadores que têm como objeto de estudos a GA brasileira a partir dos dados dos currículos Lattes, como o Professor Alberto Henrique Frade Laender, do Departamento de Ciência da Computação da Universidade Federal de Minas

Gerais, que tem como projeto o *Scence Tree*, qual sua plataforma ainda está em desenvolvimento.

2.6.1 Estudos correlatos de Genealogia Acadêmica

Diferentemente do que já foi apresentado, os estudos a seguir têm como objetivo estudar a GA de um departamento, de uma instituição ou de uma área específica ao invés de um único pesquisador. Todos eles tiveram como fonte de dados os currículos da Plataforma Lattes.

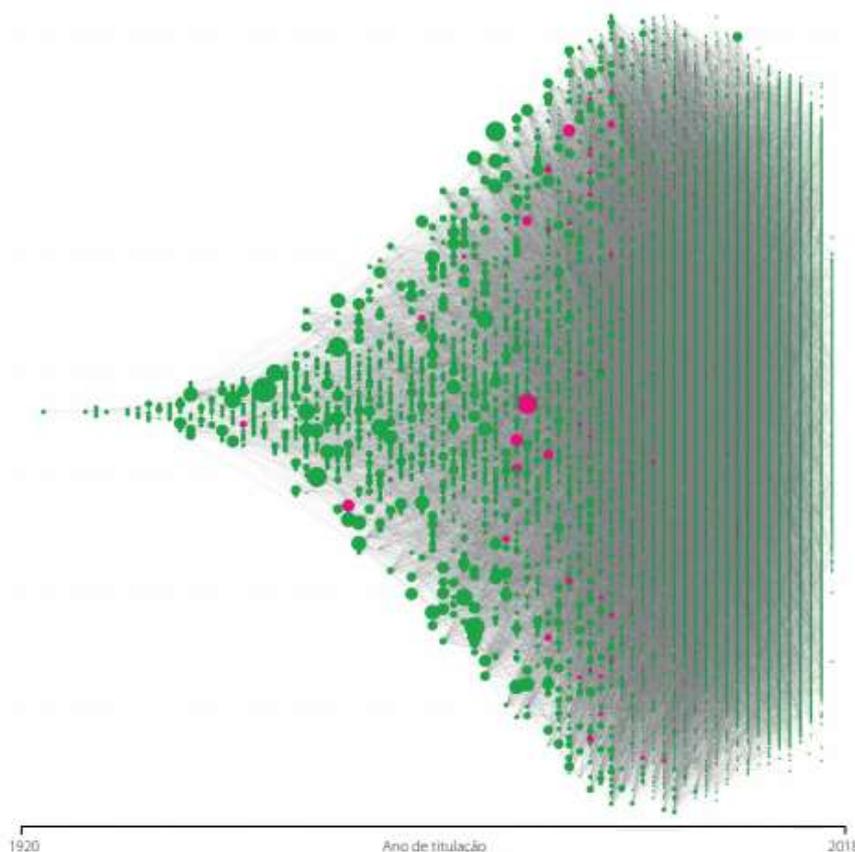
2.6.1.1 Genealogia Acadêmica da ENSP

O estudo sobre a GA da Escola Nacional de Saúde Pública Sergio Arouca, da Fundação Oswaldo Cruz (ENSP) é o mais próximo do proposto nesta pesquisa. De acordo com Viana et. al. (2019) o estudo citado tem como objetivo construir e analisar a GA da ENSP, de modo a identificar suas origens e a relação existente entre a área da Saúde Coletiva e as demais áreas do conhecimento.

O que difere o estudo da ENSP deste trabalho de pesquisa, envolvendo a GA do PPGCEM/UFSCar, é que o mesmo mede a relevância da instituição do ponto de vista do número de orientações diretas de doutorado concluídas; a captação de pesquisadores se dá de forma manual através de nomes contido no acervo da instituição e depois na busca de seus currículos na Plataforma Lattes.

Além disso, se difere na metodologia para excluir pesquisadores e relações que não fazem parte do contexto da Instituição.

Na Figura 11 pode ser observado o resultado gráfico do estudo da GA da ENSP, o que de acordo com o autor supracitado demonstrou o potencial do conceito de grafos e de GA para a análise da comunidade científica institucional.

Figura 11: Genealogia Acadêmica da ESPN

Fonte: (VIANA et al., 2019)

2.6.1.2 Genealogia Acadêmica da FEAUSP

Em 2018, foi lançada a GA do departamento de Economia da Faculdade de Economia, Administração e Contabilidade da USP (FEAUSP) com o objetivo de documentar os parentes acadêmicos de seus pesquisadores e dimensionar a influência de um pesquisador na academia. O estudo partiu de 43 docentes do departamento Economia da FEAUSP (denominados de Grupo Inicial) resultando em 819 sucessores (HADDAD et al., 2018).

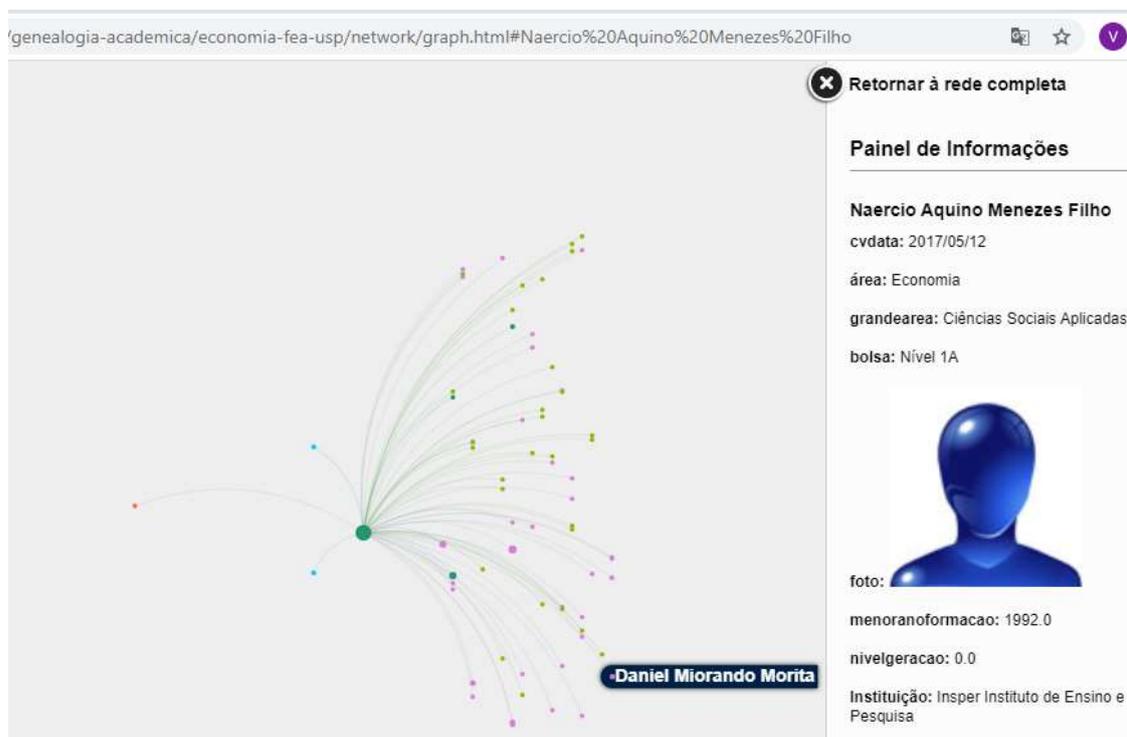
Este estudo além do grafo, tem como resultado, uma tabela pela qual é possível se verificar o pesquisador, os pais acadêmicos e todos os seus filhos acadêmicos informando nome completo, um link que direciona para o currículo lattes, a grande área, a área, a instituição, o ano de formação e o tipo de orientação.

De forma que um pesquisador pode atuar em mais de um departamento ou universidade durante sua carreira e diferentemente do que está sendo proposto na metodologia desta pesquisa, a genealogia da FEAUSP não considera apenas os descendentes da linhagem do departamento.

Exemplo disso, pode ser citado o pesquisador Naercio Aquino Menezes Filho, que tem sua árvore genealógica representada na Figura 11. De acordo com FEA/USP (2018) ele é da área da Economia e possui 67 filhos acadêmicos.

No entanto, de acordo com o seu Currículo Lattes, ele também atua em outras instituições, no caso o Insper - Instituto de Ensino e Pesquisa (LATTES, 2020a).

Figura 12: Grafo da GA de Naercio Aquino Menezes Filho



Fonte: (FEA/USP, 2018)

Porém, ao analisar os resultados, pôde-se observar que nem todos seus descendentes se formaram na FEAUSP, desta forma, estes não deveriam aparecer na árvore deste departamento, como é o caso de um de seus descendentes, o acadêmico Daniel Miorando Morita que, de acordo com o seu Currículo Lattes, conclui o seu mestrado em Economia no Instituto de Ensino e Pesquisa (LATTES, 2020b).

Portanto, o acadêmico citado é um descendente somente do pesquisador e não da FEAUSP.

Observam-se na literatura outros estudos correlatos em GA que não partem de um único pesquisador, como a GA da área da protozoologia que está disponível no site da Sociedade Brasileira de Protozoologia e que foi desenvolvida com o objetivo de identificar e caracterizar os grafos de GA da área (ELIAS; FLOETER-WINTER; MENA-CHALCO, 2016).

Ganha destaque também, a GA da área de Engenharia de *Software* no Brasil, lançado em 2016, com o objetivo de identificar e caracterizar os grafos de GA da área (MENA-CHALCO; MALDONADO; MURTA, 2016).

Ambas não foram amplamente discutidas pois, apesar de apresentarem interesses em comum em relação a GA, diferem em relação ao objetivo desta pesquisa. Enfatiza-se que o diferencial desta pesquisa está em sua metodologia a qual permitiu alcançar, de uma forma simples, todos os descendentes exclusivos da linhagem do PPGCEM/UFSCar, ou seja, todas as pessoas que têm sua origem acadêmica no programa. Para tanto, após o levantamento dos descendentes diretos por meio dos resultados do *software* AGLattes, estes foram usados como sementes em uma segunda execução do *software*. A metodologia detalhada pode ser conferida na Seção 3.5 deste trabalho.

2.6.1.3 Limitações da GA aplicada ao estudo de Programas de pós-graduação

Para Matias (2015) *softwares* como o AGLattes é uma forma alternativa de extração de dados da Plataforma Lattes, que podem sobrecarregar os servidores da Plataforma Lattes. Para evitar esse tipo de problema, a Plataforma Lattes dispões de um convênio no qual fornece às instituições de Nível Superior um acesso direto por meio de uma *web service*, para consulta de currículos em formato XML.

Porém para ter acesso ao Extrator Lattes não é tão simples. De acordo com a Plataforma Lattes (2016) para ter acesso a esses dados é necessário seguir um protocolo, preenchendo formulários e Termo de Responsabilidade, além de ser necessário seguir orientações específicas e submeter essas informações digitalmente ao Serviço da Presidência e de Apoio aos Órgãos Colegiados (SEPRE)

do CNPq.

Corrêa et. al (2017) explicam que a extração em lotes de informação na Plataforma Lattes, realizada por *softwares* com propósito semelhantes ao do AGLattes, pode a qualquer momento ter o seu funcionamento impedido pelo sistema de segurança da Plataforma Lattes, o que seria um enorme retrocesso e prejudicial para diversas pesquisas e ações que se utilizam dessas ferramentas.

Os autores ressaltam que ainda não há métodos eficazes de consultas aos dados e que a demanda por esses dados está em crescimento para desenvolvimento de diversas pesquisas, principalmente pelo fato dos dados estatísticos oferecidos pelo CNPq na Plataforma Lattes são de pouca utilidade.

Além disso há outras limitações tanto para ambos os meios de extração de dados da Plataforma Lattes.

No caso do AGLattes, apesar do *software* realizar o tratamento nos dados, existem circunstâncias que dificultam a obtenção de alguns dados, como por exemplo, a inexistência de alguns currículos, divergência no nome de orientadores e orientados ou abreviações, currículos com dados incompletos e desatualizados e falta de associação do ID Lattes dos pesquisadores (MENA-CHALCO; MALDONADO; MURTA, 2016).

Devido a esses fatores, alguns dados podem estar incoerentes com a realidade, principalmente por se tratar de um programa que teve sua origem antes de 1999, ano que de acordo com o CNPQ(2019), foi lançado e padronizado o Currículo Lattes como sendo o formulário de currículos a ser utilizados no âmbito do Ministério da Ciência e Tecnologia e CNPq. Muitos dos egressos mais antigos dos PPG que já existem há vários anos não tem currículos lattes e, portanto, informações sobre eles não são consideradas nos estudos de GA feitas a partir de dados da Plataforma Lattes.

Acrescente-se também que muitos egressos dos PPG acabam seguindo carreira fora do mundo acadêmico, e, portanto, não atualizam seus Currículos Lattes.

Isso explica a crescente a mobilização de estudiosos, como Maciel (2018) que tem a esperança de contribuição do seu trabalho para a sensibilização dos pesquisadores para a importância na atualização constante dos dados dos currículos, tanto para a ampliação da visibilidade sobre a atividade científica e tecnológica nacional, quanto para o uso dos dados dos currículos em estudos

bibliométricos.

Este trabalho adere-se a essa mobilização, na esperança de que todo pesquisador se conscientize da importância de se manter o currículo atualizado em prol a benefícios próprios e da Ciência brasileira.

Uma última limitação importante da genealogia baseada nos dados da Plataforma Lattes é que os docentes indicados como sementes no estudo de um PPG podem ser orientadores em mais de um PPG. O estudo de um PPG baseado em sementes que orientam em mais de um PPG pode ter seus resultados em termos de formação de mestres e doutores contaminados pela contagem dos descendentes das sementes provenientes de outros PPG.

Os estudos de GA baseados na Plataforma Lattes têm sido adotados e considerados úteis, mesmo com a imprecisão dos resultados decorrentes dessa contaminação. É admissível considerar-se a formação de mestres e doutores pelos docentes sementes de um PPG – o que se obtém com o uso do AGLattes no estudo de PPG's atualmente - como uma aproximação válida da formação de mestres e doutores de um PPG

No entanto, a separação de descendentes ligados diretamente ao PPG em estudo dos descendentes ligados a outros PPG aumenta a precisão dos resultados e representa mais fielmente a contribuição de um PPG na formação de pós-graduandos.

2.7 A ENGENHARIA E A IMPORTÂNCIA DO PPGCEM/UFSCAR

Não é possível imaginar o mundo sem a figura dos engenheiros. Essa tão nobre profissão se destaca desde a antiguidade, mesmo ainda sem um nome formal para a profissão das pessoas que realizavam tarefas e obras de engenharias, porém de forma empírica e regras práticas. O primeiro estabelecimento de ensino de engenharia, a *École Nationale des Ponts et Chaussés*, data-se de 1747 e foi a primeira a diplomar profissionais com esse título. Leonardo Da Vinci e Galileu, dos Séculos XV e XVII, podem ser considerados os precursores da engenharia científica, regida pelas leis da física e da matemática (TELLES, 1984).

Leonardo da Vinci, não teve seu estudo publicado, onde fez a primeira

tentativa de aplicar a estática para a determinação das forças atuando em uma estrutura simples. Galileu, por sua vez, publicou em 1638 o seu famoso livro *As Duas Novas Ciências*, que trata, entre outros assuntos, da resistência de vigas e de colunas, sendo assim o primeiro livro, em todo mundo, no campo da resistência dos materiais (TELLES, 1984).

E desde então a área da engenharia foi se estruturando à medida também que se desenvolviam as Ciências matemáticas. Porém, só foi possível chegar a um conjunto sistemático e ordenado de doutrinas, que constituíram a primeira base teórica da engenharia no Século XVIII (TELLES, 1984).

A história da Engenharia no Brasil começa em 1549 com a fundação do Governo Geral e da Cidade do Salvador por Thomé de Souza, quem trouxe consigo de Portugal, um grupo de profissionais construtores e a ordem do Rei D. João III para que fizessem uma “fortaleza de pedra e cal e uma cidade grande e forte...como melhor pode ser”. Foram grandes os desafios, até a profissão ser regulamentada no Brasil por D. Pedro I, com a lei de 28 de agosto de 1828. Até então, muita gente exercia a profissão sem o ser (TELLES, 1984).

Muitos desafios também tiveram que ser superados para que fosse criado o primeiro curso de graduação em Engenharia de Materiais da América Latina, na recém fundada Universidade Federal de São Carlos. O principal percussor dessa grande iniciativa foi o físico e químico Sergio Mascarenhas de Oliveira, que participou da criação da UFSCar e que tinha a visão de que o desenvolvimento tecnológico brasileiro dependia do melhor aproveitamento dos materiais, gerando riqueza para o país, através da criação de emprego e da exportação (SAVIGNANO, 2012).

Dentre esses desafios, está a oposição e resistência principalmente da Engenharia Metalúrgica e da Química, que acreditavam não ser necessária e precipitada a criação do curso. Com isso, vieram as reações e denúncias junto ao Conselho Federal de Educação com instauração de inquérito, além de toda polêmica e resistências surgidas em torno da estruturação do curso. Mas isso não foi o suficiente para impedir que fosse realizado o primeiro vestibular para “Engenharia de Ciência de Materiais”, como foi inicialmente chamado o curso e, em 1974, trinta e quatro estudantes se formavam nessa carreira (SAVIGNANO, 2012).

Além de ser pioneiro no Brasil, o curso de Engenharia de Materiais é até hoje o único que oferece cursos com ênfases específica em Materiais Metálicos,

Materiais Poliméricos e Materiais Cerâmicos (“Revista Edição Comemorativa DEMa 40 Anos”, 2012).

Em 1972 é instituído o Departamento de Engenharia de Materiais (DEMa), somente após dois anos do início das atividades acadêmicas da UFSCar. Não demorou muito para que o DEMa se consolidasse como um órgão acadêmico de geração e desenvolvimento de atividades interdisciplinares de ensino (graduação, pós-graduação, *stricto sensu* e cursos de extensão) de pesquisa científica, desenvolvimento tecnológico e prestação de serviços, formando mais de 1700 engenheiros (“Revista Edição Comemorativa DEMa 40 Anos”, 2012).

Além de Sérgio Mascarenhas, outros grandes e importantes nomes foram responsáveis pelo projeto audacioso e de sucesso, como José Roberto da Silva, Augusto Marcondes Agnelli, Rosário Elida Suman Bretas, Sebastião Elias Kuri, Levi de Oliveira Bueno, José Alexandrino de Souza e Vanderlei Belmiro Sverzut. Para Mascarenhas a Engenharia de Materiais é a “maior medalha do peito” da UFSCar (“Revista Edição Comemorativa DEMa 40 Anos”, 2012).

O projeto bem-sucedido teve sua ampliação em 1979, quando foi criado o primeiro curso de mestrado em Engenharia de Materiais com três áreas de concentração: cerâmica, metalurgia e polímeros. Em 1987 surgiu então o doutorado em Ciência e Engenharia de Materiais com as mesmas áreas de concentração, segmentando desta forma o Programa de Pós-Graduação em Ciência e Engenharia de Materiais (PPGCEM/UFSCar) que tem como principal objetivo contribuir para o desenvolvimento científico e tecnológico do Brasil através da formação de mestres e doutores (“Revista Edição Comemorativa DEMa 40 Anos”, 2012).

A implantação do curso de Engenharia de Materiais impactou na economia regional e principalmente na cidade São Carlos, que contribuiu para que a cidade obtivesse o título oficial de capital nacional da tecnologia. A área de materiais com o apoio e liderança do DEMa-UFSCar, foi uma grande geradora de empresas de base tecnológica através da pesquisa e formação de profissionais qualificados. Muitas empresas de base tecnológica na cidade São Carlos, surgiram a partir de laboratórios das universidades, muitas com influência do DEMa. Fatos como esses, elevou a cidade de São Carlos como a cidade com maior densidade de doutores do Brasil (SAVIGNANO, 2012).

O PPGCEM/UFSCar, atualmente, conta com 35 docentes credenciados, os quais, em sua maioria, pertencentes ao DEMa, e ao Departamento de Química da

UFSCar e EMBRAPA de São Carlos, muitos deles são reconhecidos internacionalmente (PPGCEM/UFSCar, 2019).

Desde a sua fundação, a UFSCar teve um grande crescimento e conta hoje com quatro Campis, que estão localizados nas cidades de São Carlos, Araras, Sorocaba e Lagoa do Sino. Possui quarenta e três Programas de Pós-Graduação, dentre os quais, quatro deles possuem nota sete de acordo com a avaliação da CAPES, tanto no mestrado (ME) como no doutorado (DO) conforme pode ser observado no Quadro 4, onde também podem ser conferidas as notas dos demais programas (PLATAFORMA SUCUPIRA, 2020).

Quadro 4: Relação dos PPG da UFSCar e suas respectivas notas.

Programa	Nota	Programa	Nota
Ciência e Engenharia dos Materiais	7	Enfermagem	4
Engenharia Química	7	Engenharia Civil	4
Fisioterapia	7	Engenharia Urbana	4
Química	7	Estatística - UFSCar/USP	4
Educação Especial (educação do indivíduo especial)	6	Biotecnologia e Monitoramento Ambiental	4
Psicologia	6	Física	4
Sociologia	6	Genética Evolutiva e Biologia Molecular	4
Antropologia Social	5	Linguística	4
Ciência Política	5	Matemática	4
Educação	5	Planejamento e Uso de Recursos Renováveis	4
Engenharia de Produção	5	Terapia Ocupacional	4
Filosofia	5	Agricultura e Ambiente	3
Estudos de Literatura	4	Agroecologia e Desenvolvimento Rural	3
Biotecnologia	4	Ciência da Computação	3
Ciência da Computação	4	Ciência da informação	3
Ciência dos Materiais	4	Educação em Ciências e Matemática	3
Ciência, Tecnologia e Sociedade	4	Engenharia de produção	3
Ciências Ambientais	4	Geografia	3
Ciências Fisiológicas	4	Gerontologia	3
Ecologia e Recursos Naturais	4	Imagem e som	3
Economia	4	Produção Vegetal e Bioprocessos associados	3
Educação	4		

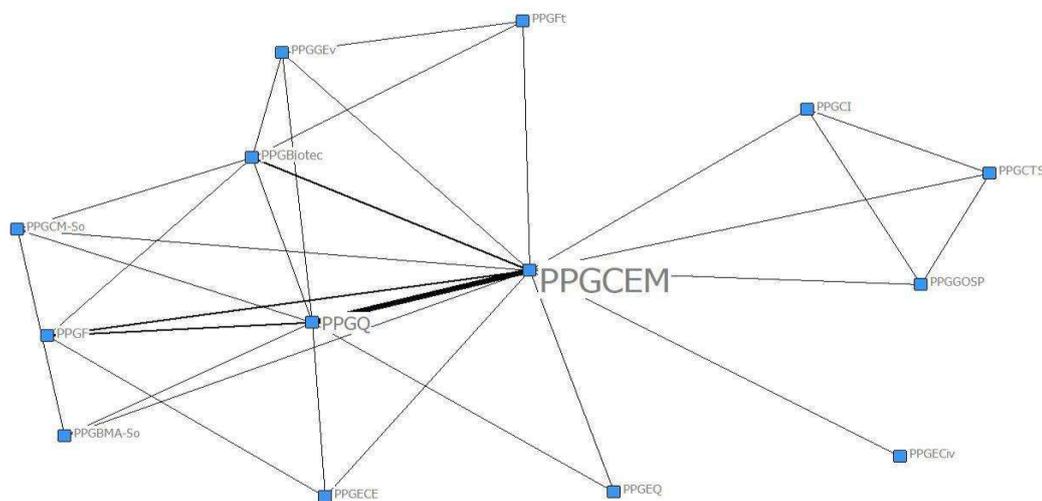
Fonte: (PLATAFORMA SUCUPIRA, 2020)

Pelo fato dos docentes do PPGCEM/UFSCar estarem atrelados a outros Programas de Pós-Graduação da UFSCar, identifica-se uma expressiva produção de artigos, o que contribui para que este programa, há 12 anos, receba nota máxima (nota 7) na avaliação da CAPES, garantindo o grau de excelência padrão internacional (MACIEL, 2018).

A relevância do estudo do PPGCEM/UFSCar pode ser, também, constatada nos resultados da avaliação dos Programas de Pós-Graduação realizado pela CAPES, onde dos 32 (trinta e dois) cursos avaliados e reconhecidos na área de Engenharia de Materiais e Metalúrgica, o curso do PPGCEM/UFSCar é um dos três que obtiveram a nota máxima em sua avaliação (Sucupira, 2019), comprovando que o programa é uma referência nacional e internacional na sua área e importante para a evolução da Ciência e Engenharia de Materiais no Brasil.

O PPGCEM/UFSCar apresenta significativa colaboração, como pode ser observada na Figura 13, relativa à produção de artigos, no quadriênio 2013-2016, identificada tanto por meio da coautoria entre docentes de PPG's distintos quanto pelo fato de um docente do PPGCEM/UFSCar também estar credenciado em outro(s) PPG, como o Programa de Pós-Graduação em Química (PPGQ), Programa de Pós-Graduação em Física (PPGF), o Programa de Pós-Graduação em Ensino de Ciências Exatas (PPGECE), o programa de Comunicação e Informação (PPGCI), Interdisciplinar (PPGCTS), Biotecnologia (PPGBiotec) e Administração Pública e de Empresas, Ciências Contábeis e Turismo (PPGGOSP) (MACIEL, 2018).

Figura 13: Rede de colaboração do PPGCEM/UFSCar



Fonte: (MACIEL, 2018)

As Engenharias como um todo apresentam caráter multidisciplinar e interdisciplinar, no que se refere as pesquisas científicas, formação de recursos humanos, e novas tecnologias. Isso ocorre, basicamente, porque há o envolvimento com outras áreas correlatadas quando se trata de desenvolvimentos científicos e tecnológicos(PLATAFORMA SUCUPIRA, 2019).

3 MÉTODO E DESENVOLVIMENTO

O grande privilégio em se fazer pesquisa é poder conhecer mais de perto uma realidade (ECHER, 2001). Para se chegar neste conhecimento, foram adotados procedimentos para a realização desta pesquisa, quais estão apresentados neste capítulo.

3.1 ABORDAGEM E TIPOLOGIA DA PESQUISA

A pesquisa quantitativa é voltada para tratar os aspectos da realidade que podem ser quantificados e tem como foco representar a população que está sendo estudada através da representatividade numérica, se centrando na objetividade, tendo suas raízes no pensamento positivista lógico e tende a enfatizar o raciocínio dedutivo, as regras da lógica e os atributos mensuráveis da experiência humana. Já quanto a natureza, a Pesquisa Aplicada, objetiva gerar conhecimentos para aplicação prática, dirigidos à solução de problemas específicos e que envolve verdades e interesses locais. De acordo com os objetivos da pesquisa, ela pode ser exploratória, qual está deve proporcionar maior familiaridade com o problema a fim de torná-lo mais explícito ou a construir hipóteses através da análise de exemplos que estimulem a compreensão. (SILVEIRA; CORDOVA, 2009).

Desta forma, o presente trabalho é de abordagem quantitativa. Quanto à natureza, trata-se de uma pesquisa aplicada, uma vez que ela vai gerar conhecimentos para aplicações práticas na solução de problemas específicos. Quanto aos objetivos é uma pesquisa exploratória, de forma que ela busca familiaridade com problemas poucos conhecidos, exigindo a aquisição de novos conhecimentos em ferramentas como o *software* AGLattes-v2, Sistema Operacional Linux, Oracle-VirtualBox, VantagePoint, Lattes Machine e linguagem de programação Python.

Foi feito levantamento bibliográfico, principalmente a partir do Portal de Periódicos da CAPES, para identificação artigos científicos, livros e outras publicações contendo o estado da arte em relação aos temas pesquisados e que permitisse a discussão dos resultados.

Enquadra-se também, como estudo de caso, tendo em vista que foi feito um estudo aprofundado sobre o PPGCEM/UFSCar a fim de revelar um conhecimento mais abrangente sobre ele e que permitisse a generalização dos resultados para estudo de outros PPG.

3.2 FERRAMENTAS

No Quadro 5 é possível visualizar, de forma sintetizada, as ferramentas que foram utilizadas para esta pesquisa.

Quadro 5: Síntese das ferramentas utilizadas

Processo	Ferramenta
Baixar currículos da Plataforma Lattes	<i>software</i> AGLattes
Filtragem e análise dos dados resultantes do AGLattes	VantagePoint
Levantamento da Produção Científica	Lattes Machine a partir de dados fornecidos pelo AGLattes.
Seleção de dados do AGLattes e Construção de indicadores de formação e produção científica	Microsoft Excel

Fonte: Elaborado pela autora

Como visto, desde a sua fundação, entre os atuais docentes e os que passaram pelo PPGCEM/UFSCar, somam-se 58 docentes. Apesar da possibilidade de localizar o Currículo Lattes de cada um manualmente, direto na Plataforma Lattes, o mapeamento de seus descendentes e o levantamento de dados sobre eles, resultaria em um trabalho custoso e inviável, devido ao tempo que isso levaria. Diante disto, iniciou uma investigação sobre as ferramentas já existentes, que pudessem ser utilizadas para atingir o objetivo desta pesquisa.

Logo, os estudos foram concentrados na Plataforma Acácia, um web site, onde é possível consultar ascendentes e descendentes de um acadêmico brasileiro, em nível de mestrado e doutorado conforme descrito na Seção 2.4.1 deste trabalho. Porém, um problema encontrado foi que esta plataforma permite apenas consultar um acadêmico por vez e o resultado se dá através de uma lista, não sendo possível consultar de uma única vez todos os acadêmicos de um programa inteiro, devido ao funcionamento dos filtros existentes no site, o que seria inviável para a realização desta pesquisa, já que para tanto, seria necessário pesquisar de uma única vez todos os docentes do PPG.

Contudo, a partir da Plataforma Acácia, foi possível chegar até o *software* AGLattes, que é o *software* que obtém a GA de pesquisadores a partir de seus currículos indexados na Plataforma Lattes, a nível de mestres e doutores.

O *software* AGLattes foi desenvolvido e, cedido para estudos a versão Aglattes-v2, pelos pesquisadores Rafael Jeferson Pezzuto Damaceno, Prof. Dr. Jesus P. Mena-Chalco e Luciano Rossi, com o objetivo principal de documentar as relações formais de orientação no contexto dos PPG's brasileiros sob a forma de grafos de GA, em que cada vértice representa um pesquisador e cada aresta uma relação de orientação concluída entre dois pesquisadores (orientador e orientado) (DAMACENO; ROSSI; MENA-CHALCO, 2019).

De acordo com Damasceno (2019) é a partir dos dados obtidos por este *software* que a Plataforma Acácia executa suas ações.

Desta forma, esta foi a ferramenta utilizada para baixar os currículos da plataforma Lattes para a execução desta pesquisa e para tanto, foi necessário estudar o seu funcionamento e suas principais características estão descrita no Apêndice A deste trabalho.

O *software* AGLattes disponibiliza seus resultados em dados brutos divididos em arquivos em formato XML, XLSX e um grafo no formato GDF. São dados brutos pois, não estão organizados em sua totalidade. Portanto, foi necessário utilizar outras ferramentas para transformar esses dados em informação.

E obter informação desses dados foi um dos desafios encontrados para a execução deste trabalho, exigindo pesquisas em ferramentas computacionais que atendessem a demanda de forma menos complexa, tendo como princípio de que, de acordo com Ribeiro (2014) a Ciência da Informação é o campo de estudos apropriado para desenvolver novos estudos acerca de dados e informação. Assim,

esperava-se usar uma ferramenta menos complexa, de forma que qualquer cientista da informação pudesse replicar esta pesquisa ou adotá-la em outros estudos.

Após um levantamento sobre as necessidades de acordo com o tempo disponível para a execução da pesquisa, optou-se por focar as análises dos dados nos formatos XML e XLSX.

Diante do grande volume de dados que um PPG pode apresentar, foi feito um levantamento sobre as possíveis ferramentas que pudessem analisá-los de uma única vez. Três ferramentas foram testadas para a realização da análise:

- **VantagePoint:** A primeira ferramenta que foi testada foi o VantagePoint para analisar os resultados dos dados no formato XML. O VantagePoint é uma poderosa ferramenta para mineração de textos que tem como principal vantagem a facilidade de análise dos dados através de vários recursos que permitem a realização de filtros, grupos, thesaurus, listas e posteriormente exportar os conjuntos de dados limpos diretamente para o Excel para uso em outras ferramentas (VANTAGEPOINT, 2019). É uma ferramenta muito utilizada para se realizar bibliometria. As desvantagens para o uso desta ferramenta é que se trata de um *software* proprietário, ou seja, para utilizá-lo é necessário obter licenças de uso junto ao produtor e a utilização do *software* requer treinamento (FARIA, 2015).
- **Banco de dados Microsoft SQL Server:** Outra ferramenta a ser testada para a análise desses arquivos foi a Microsoft SQL Server com o gerenciador de banco de dados Microsoft SQL Server Management Studio. Nesta ferramenta foram explorados os arquivos do tipo XLSX. Esta ferramenta apresentou grande facilidade para o manuseio dos dados sem a necessidade de se realizar ações filtragem de forma manual. Porém, este método exige conhecimento em Banco de Dados, Gerenciador de Banco de Dados e linguagem SQL. Por este motivo, esse método de análise foi considerado complexo para cumprir o objetivo desta pesquisa.
- **Microsoft Excel:** A última ferramenta a ser testada foi o editor de planilhas da Microsoft Excel, que apesar de ser um *software* proprietário, está mais acessível às pessoas, comparado ao *software* VantagePoint e permite avaliar todos os descendentes até mesmo

quais o *software* não conseguiu identificar o *idlattes*. Quando isso ocorre, o *software* atribui um número aleatório ao *idlattes* o que permite a análise dos dados. Essa ferramenta foi testada tanto para filtrar os dados como também para a sua análise. A desvantagem da utilização desse *software* para filtragem dos dados é a necessidade de ações manuais o que requer tempo e atenção, além do conhecimento a nível intermediário da ferramenta. Através desta ferramenta, foi constatado que é possível realizar a triagem de todas as gerações do programa, mas para tanto, é necessário utilizar recursos como macros e filtros avançados, quais podem sofrer alterações constantes de acordo com a necessidade de uso pelo PPG, além de exigir um conhecimento intermediário a avançado destes recursos.

Diante do contexto apresentado optou-se por utilizar a ferramenta Microsoft Excel para selecionar os descendentes diretos, bem como a construção dos indicadores.

Utilizou-se a ferramenta VantagePoint para filtrar e analisar os dados em arquivos em formato *XML*.

Já para o levantamento das publicações científicas utilizou-se a ferramenta Lattes Machine, que posteriormente, tiveram seus indicadores construídos também na ferramenta Microsoft Excel. Enfatiza-se que só foi possível realizar a bibliometria das publicações científicas a partir dos dados de entrada adquiridos por meio do *software* AGLattes.

De acordo com Matias (2015) a ferramenta Lattes Machine funciona através de um conjunto de scripts que realiza a extração, tratamento e sincronização de metadados do Currículo Lattes com o DSpace, uma ferramenta de código aberto que permite a construção de repositórios digitais provendo o armazenamento, acesso e preservação a longo prazo e que foi desenvolvida pelo Instituto Tecnológico de Massachusetts – MIT e a Hewlett Packard Corporation – HP iniciado em 2002.

3.3 PROCEDIMENTO DE COLETA DOS DADOS E DELIMITAÇÃO DA AMOSTRA

Para o desenvolvimento desta pesquisa, foi delimitado o estudo no PPGCEM/UFSCar da Universidade Federal de São Carlos. Teve como escolha o referido programa considerando os fatores que já foram relatados neste trabalho, como o fato deste programa ser considerado um marco no desenvolvimento científico e tecnológico da área no Brasil, já contar 40 anos na formação de pós-graduandos e possuir uma número expressivo de mestres e doutores formados de tal forma que, no caso dos procedimentos adotados, mostrarem-se viáveis para o estudo do PPGCEM/UFSCar, seriam viáveis também no estudo de outros PPG.

De acordo com o site do programa, até dezembro de 2019, foram formados 879 Mestres e 412 Doutores, formando um total de 1291 descendentes. Apesar de tentativas, não foi possível ter acesso a esses dados para identificar quantos deles obtiveram ambas as titulações no programa para saber com exatidão a quantidade de egressos, o que foi representado na Figura 14.

Figura 14: Formação de mestres e doutores pelo PPGCEM/UFSCar



Fonte: Adaptado de PPGCEM/UFSCar (2019)

Como o PPGCEM/UFSCar iniciou suas atividades em 1979, foi necessário fazer uma pesquisa junto à Secretaria Geral de Planejamento e Desenvolvimento Institucionais (SPDI) da UFSCar, para o levantamento de todos os docentes que

atuaram e ainda atuam no programa e seus respectivos *id/attes*, totalizando cinquenta e oito docentes entre os anos de 1979 a 2019, os quais podem ser conferidos no Quadro 6.

Quadro 6: Sementes do PPGCEM/UFSCar 1978 - 2019

Adhemar Colla Ruvolo Filho	João Baptista Baldo	Oscar Balancin
Alberto M. Jorge Junior	José Angelo Gregolin	Pedro A.de P. Nascente
Alessandra de A. Lucas	Jose Antonio Eiras	Pedro Iris Paulin Filho
Ana Candida M. Rodrigues	José Arana Varela	Piter Gargarella
Anselmo Ortega Boschi	José Augusto M. Agnelli	Rinaldo Gregório Filho
Carlos Alberto Della Rovere	José de A. Rodrigues	Roberto Tomasi
Carlos Henrique Scuracchio	José Donato Ambrósio	Rodrigo Bresciani Canto
Carlos Ventura D'Alkaine	José Eduardo Spinelli	Rosário E. S. Bretas
Claudemiro Bolfarini	José Manoel Marconcini	Ruth H. G. A. Kiminami
Claudio Shyinti Kiminami	Leonardo B. Canto	Sati Manrich
Conrado Ramos M. Afonso	Levi de Oliveira Bueno	Sebastião Elias Kuri
Daniel Rodrigo Leiva	Lidiane Cristina Costa	Sebastião V. C. Junior
Dulcina M. P. F. de Souza	Luiz Antonio Pessan	Sílvia Helena P. Bettini
Edgar Dutra Zanotto	Luiz H. C. Mattoso	Tomaz Toshimi Ishikawa
Edson Roberto Leite	Marcello R. B. Andreetta	Victor Carlos Pandolfelli
Elias Hage Junior	Márcio R. Morelli	Vitor Luiz Sordi
Elson Longo da Silva	Marco Aurelio L. Cordeiro	Walter José Botta Filho
Guilherme Zepon	Maria Zanin	Walter Libardi
Hans Jürgen Kestenbach	Maurizio Ferrante	
José A. de Souza	Nelson G.de Alcântara	

Fonte: (UFSCAR, 2019)

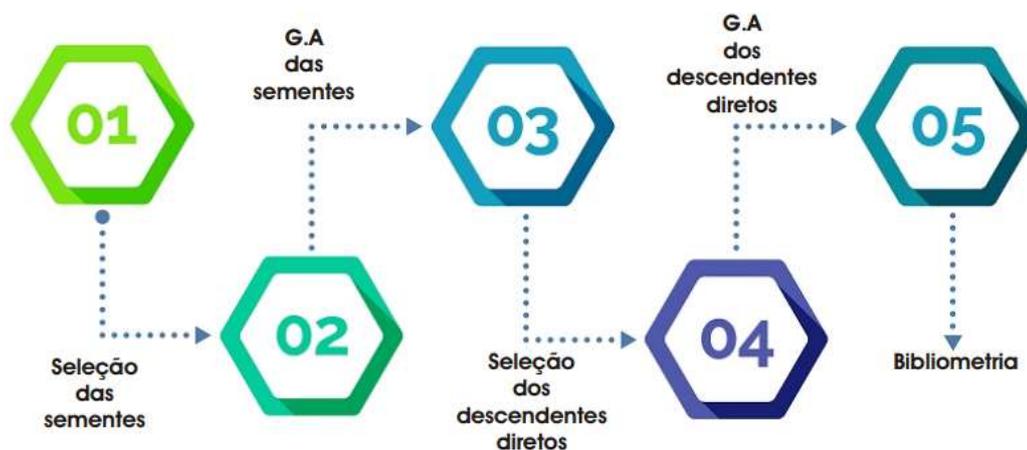
Esses docentes são a origem do programa, os quais a partir de agora, serão denominados como “sementes”. Diante destas informações, tinha-se a certeza da grande quantidade e riqueza de dados que poderiam ser coletados através da GA do PPGCEM/UFSCar.

3.4 NOVO PROCEDIMENTO DE GENEALOGIA ACADÊMICA E BIBLIOMETRIA

Nesta Seção está descrito o novo procedimento adotado para o estudo da contribuição de um PPG para a formação de mestres e doutores, a partir da

separação dos descendentes diretamente ligados ao PPG. O novo procedimento envolveu uma dupla aplicação da GA. A primeira aplicação adotou os docentes do PPG como sementes, o que é tradicionalmente feito nos estudos atuais. A partir das planilhas geradas na primeira aplicação, foi possível separar a primeira geração de descendentes diretamente ligados ao PPG estudado – aqueles que foram orientados pelos docentes-sementes no PPG estudado e não em outros PPG - para serem usados como sementes da segunda aplicação da Genealogia. O resultado da segunda aplicação é que todos os descendentes agora identificados, ou são diretamente ligados ao PPG estudado ou foram orientados por docentes formados no PPG estudado. Esse novo procedimento e sua aplicação ao PPGCEM/UFSCar, está representado na Figura 15 e é descrito a seguir.

Figura 15: Fluxograma do procedimento da pesquisa



Fonte: Elaborado pela autora

Após selecionar as sementes e executar o *software* AGLattes, com o objetivo de levantar dados apenas do PPGCEM/UFSCar e dos seus descendentes, sem considerar os orientados de suas sementes em outros programas ou em outras universidades, adotou-se os alguns critérios para a triagem dos descendentes diretos e indiretos, para posteriormente baixar os currículos da Plataforma Lattes por meio do *software* AGLattes que estão descritos a seguir.

3.4.1 Triagem dos descendentes diretos

O levantamento dos descendentes diretos foi realizado a partir do arquivo *7_arestas*, o qual contém dados referentes a relação orientador e orientado de todos os acadêmicos, tanto dos descendentes como dos ascendentes que teve alguma relação com as sementes do PPGCEM/UFSCar e seus descendentes. No caso, são dados de 5 gerações em uma única planilha, somado aos dados de todos os seus ascendentes, totalizando 3122 registros.

Para identificar somente os descendentes diretos do programa, filtrou-se pelos campos, na seguinte ordem:

- 1) Instituição: Na filtragem deste campo foram considerados os seguintes termos atribuídos no preenchimento do Currículo Lattes:
 - “universidade federal de são carlos”;
 - “programa de pósgraduação em Ciência e engenharia de materiais da ufscar”;
 - “programa de pósgraduação em Ciência e engenharia de materiais”
 - “programa de pós graduação em Ciência e engenharia de materiais da ufscar”;
 - “departamento de engenharia de materiais”,
 - “programa de pósgraduação em Ciência e engenharia de materiais; ufscar”,

Após a filtragem restaram 1389 registros.

- 2) Curso: Na filtragem deste campo foram considerados todos os nomes que tinham em sua descrição a palavra “materiais” e “doutorado”. A palavra “doutorado” foi considerada pois havia apenas dois registros que indicavam essa descrição para o campo Curso. Desta forma, foi considerado viável conferir o dado diretamente na Plataforma Lattes, atestando que os dois registros pertencem a acadêmicos que se formaram no PPGCEM/UFSCar. Após a filtragem restaram 1002 registros.
- 3) Origem_Nome: Na filtragem deste campo foram desconsiderados todos os nomes que não pertencem ao PPGCEM/UFSCar, de acordo com a lista de docentes adquirida junto à Secretaria Geral de Planejamento e

Desenvolvimento Institucionais (SPDI) da UFSCar, como já foi mencionado na seção 3.2 deste trabalho. Foram desconsiderados os nomes de Aparecido Junior Menezes, Fábio de Lima Leite e Flávio Salgado Politi.

Após a filtragem restaram 993 registros. No entanto, dentre esses registros constam os dados dos descendentes que fizeram mestrado e doutorado no PPGCEM/UFSCar. Notou-se que o mesmo acadêmico pode ser descendente de orientadores diferentes, desta forma, foi adotado o critério baseado no processo seguinte, que é a bibliometria das publicações científicas desses acadêmicos.

Baseada no critério de que as publicações deveriam ser contadas a partir da primeira titulação do acadêmico, foi considerada a titulação de mestrado, excluindo a formação de doutorado desses acadêmicos em específico. Também foram excluídos os descendentes que não continham *idlattes*, pois esta é variável indispensável para a execução da ferramenta Lattes Machine. Assim, foram removidos 367 registros sem *idlattes*. Após a exclusão dos registros especificados, restaram 586 registros, que foram consideradas as novas sementes do PPGCEM/UFSCar.

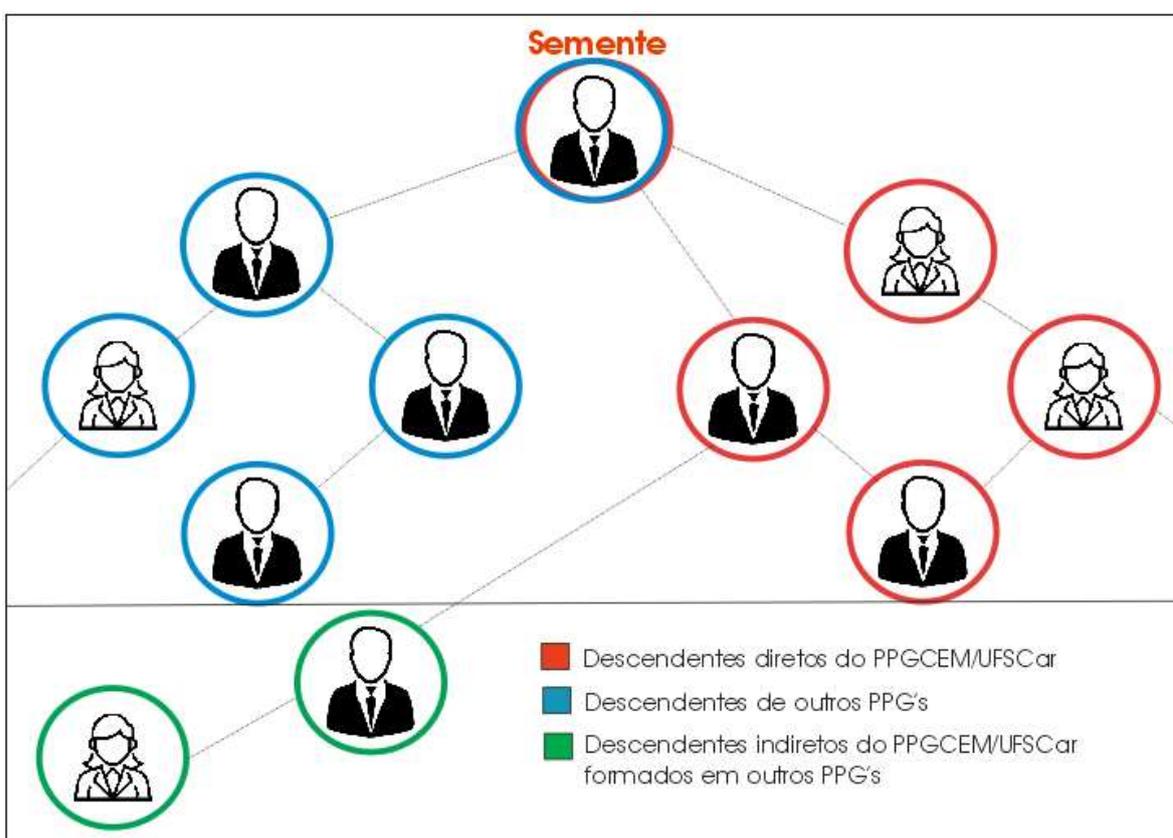
3.4.2 Triagem dos demais descendentes do PPGCEM/UFSCar

Uma solução, embora simples, mas eficiente para encontrar todos os descendentes da mesma linhagem do PPGCEM/UFSCar foi executar o *software* pela segunda vez, agora, com todos os descendentes diretos como sementes, porém configurado apenas como [descendência], ou seja, desconsiderando a ascendência.

Na Figura 16 é possível visualizar como a classificação dos descendentes diretos e indiretos foi realizada para a execução desta pesquisa. No centro da figura está a semente do programa PPGCEM/UFSCar que pode atuar em programas distintos ao mesmo tempo, estando representada pelas cores vermelha e azul. Do lado direito da figura os descendentes diretos estão representados pela cor vermelha. Estes, são todos os acadêmicos que fizeram mestrado ou doutorado no programa. Do lado esquerdo da figura, estão os descendentes da semente, porém

que fizeram mestrado ou doutorado em outros PPG's. Embora estes acadêmicos sejam descendentes da mesma semente, eles não podem ser considerados como descendentes do programa, bem como suas produções científicas não devem ser contabilizadas para o programa. Já na parte inferior da figura, estão os descendentes indiretos do PPGCEM/UFSCar, representados pela cor verde. Estes acadêmicos desfrutaram de todo o empenho realizado pelo PPGCEM/UFSCar durante toda a sua história, na área de Engenharia de Materiais e Metalúrgica, através do conhecimento que chegou até eles por meio de seus ascendentes.

Figura 16: Método de classificação dos descendentes



Fonte: Elaborado pela autora

Portanto, através desta metodologia, foi possível chegar nos acadêmicos que foram formados pelo PPGCEM/UFSCar, que são os descendentes diretos do programa e nos seus descendentes indiretos, que mesmo formados em outros departamentos ou instituição, de alguma forma, os conhecimentos produzidos no

currículo, que por algum motivo, como a sua exclusão, não constava mais na base de dados da Plataforma Lattes. Para solucionar o problema, o número lattes 2462579245129970 em nome de Marcos Vinícius Gerotto, foi excluído do arquivo sementes. Após esse processo, o *software* executou normalmente, resultando em 1503 currículos baixados em arquivos no formato XML, incluindo os descendentes diretos e indiretos.

3.4.3 Análise bibliométrica da segunda aplicação da Genealogia Acadêmica

A segunda aplicação da GA resultou na coleta de 1208 currículos de descendentes indiretos do PPGCEM. Esses currículos foram recuperados em arquivo XML e importados para o *software* VantagePoint utilizando-se o filtro XML Lattes desenvolvido pelo NIT/Materiais. Os currículos foram contados segundo informações neles presentes, permitindo a construção de indicadores referentes aos descendentes diretos e indiretos do PPGCEM, tais como número de mestres e doutores formados, ano de formação dos mestres e doutores, tipo de atuação profissional, instituição e departamento de atuação, país, estado e cidade de atuação, dentre outros. De forma alternativa, foi testada a construção dos mesmos indicadores a partir das planilhas resultantes da segunda aplicação da GA e utilizando-se o *software* Excel e recursos de programação. Os resultados alcançados são similares aos obtidos via XML e VantagePoint.

4 RESULTADOS E DISCUSSÃO

O *software* AGLattes foi desenvolvido, especificamente, para atender os estudos referentes a GA. Desta forma, foi realizada uma análise por meio de ferramentas e procedimentos, verificando o seu potencial na extração de informações relevantes para este estudo, conforme descritos na seção Métodos e Desenvolvimento.

A contribuição desse estudo refere-se, principalmente, na oportunidade do uso de uma ferramenta que possa apoiar na avaliação ou no acompanhamento dos Programas de Pós Graduação, contribuindo para que estes tenham conhecimento sobre seus egressos, permitindo levantar informações como por exemplo, quem são, onde estão, o que estão fazendo, entre outras informações, o que vem ao encontro com que a CAPES espera atualmente desses programas .

Outra contribuição está na resolução de uma limitação da GA no caso de estudos em PPG's, que se refere a propagação da análise com os descendentes que não pertencem ao programa, reduzindo a contaminação dos dados.

Em relação a associação da GA com a Bibliometria, neste estudo foi feita de duas maneiras: Primeiro, aplicou-se a Bibliometria nos resultados da GA, para contabilizar dados e levantar informações sobre a atuação do programa por meio de seus descendentes diretos e indiretos, como formação de recursos humanos, a localização e atuação entre outras. Posteriormente a Bibliometria foi aplicada para se conhecer a produção científica deles e isso só foi possível com a associação das duas técnicas.

Fazer a coleta de dados dos Currículos Lattes sem o uso da GA demandaria um trabalho enorme porque a priori precisaria ter os *idlattes* ou nomes completos de todos os alunos do programa e posteriormente fazer um mapeamento manual de cada um de seus descendentes. Já com a GA fazendo a bibliometria a partir de seus resultados, só é preciso os *idlattes* dos docentes (sementes).

Ao trabalhar com os Currículos Lattes, sabe-se das limitações, pois nem todas as pessoas têm o currículo cadastrado na plataforma e há atrasos em suas atualizações. No entanto estudos como esse, que exploram e confirmam a potencialidade dos dados da Plataforma Lattes, incentivam para que tanto a plataforma como *softwares* como o AGLattes continuem melhorando,

proporcionando ganhos significativos para todos os envolvidos.

O novo procedimento empregado para a GA aplicada ao estudo de um PPG, considerado inovador - executando o *software* a primeiro momento com os *idlattes* dos docentes e posteriormente com o *idlattes* dos descendentes diretos - permitiu chegar a todos os descendentes exclusivos do programa, ou seja, em todos os acadêmicos que têm a sua origem no PPGCEM/UFSCar.

Esta execução do *software* em duas etapas, apesar de exigir um considerado esforço no primeiro momento, sendo necessário ajustes no *software*, apresentou-se eficaz para atingir os resultados esperados, não trazendo ruídos de outros acadêmicos que não pertencem a linhagem do programa.

Ademais o estudo confirmou a importância do PPGCEM/UFSCar para a propagação da Engenharia de Materiais no Brasil.

Neste capítulo estão descritos os indicadores desenvolvidos a partir da GA do PPGCEM/UFSCar, os quais se referem apenas a este programa e que foram desenvolvidos de acordo com os dados coletados pelo AGLattes da Plataforma Lattes.

É preciso enfatizar que este trabalho não tem por objetivo avaliar a atuação de seus docentes no programa e sim verificar, como é possível construir indicadores bibliométricos através dos dados da GA do programa, e o quanto eles podem contribuir na autoavaliação de qualquer PPG.

Desta forma, para a explanação dos indicadores e sua importância foi necessário evidenciar alguns números, o que não significa mérito ou demérito de qualquer docente. Como já visto, todos foram e são importantes para o programa, para a área de Engenharia de Materiais e Metalúrgica e para a Ciência como um todo.

4.1 INDICADORES DE FORMAÇÃO DE RECURSOS HUMANOS

Os indicadores descritos nesta seção foram desenvolvidos baseados no objetivo da proposta metodológica de autoavaliação da CAPES, que de acordo com a CAPES (2019) é esperado que no exercício de autoavaliação, os PPG's visem o monitoramento da qualidade do programa e seu processo formativo e mantenham foco na formação discente pós-graduada na perspectiva da inserção social e/ou científica e/ou tecnológica e/ou profissional, presencial e/ou a distância do

programa. No entanto, perante a delimitação do tema e o cronograma da pesquisa, não houve estudos aprofundados no que tange os requisitos para atender a proposta metodológica de autoavaliação e nem o processo de avaliação da CAPES.

Desta forma, os indicadores também foram construídos baseados em situações hipotéticas das necessidades de um PPG, o que pode variar de acordo com os objetivos de cada programa.

4.1.1 Contribuição na formação de mestres e doutores

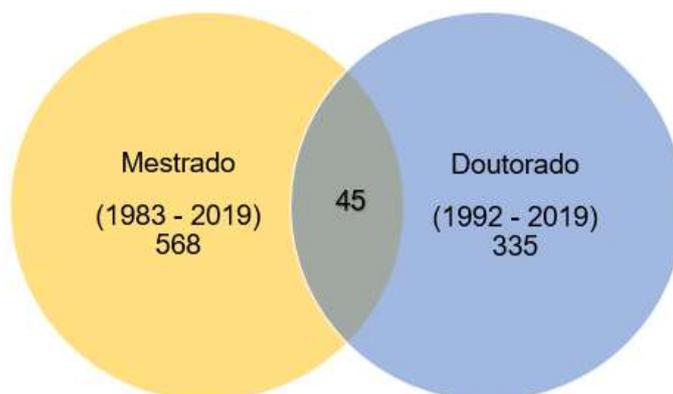
O primeiro Indicador extraído a partir dos dados processados foi o de contribuição do PPGCEM/UFSCar na formação de mestres e doutores. Este indicador foi desenvolvido a partir dos dados dos descendentes diretos do programa. Após a triagem dos descendentes como fora mencionado na seção 3.4.1 deste trabalho, foram identificados 993 (novecentos e noventa e três) descendentes diretos, onde aplicou-se filtros no campo Titulação. Primeiro pesquisou pelas titulações de doutorado e posteriormente de mestrado, contando a seguir somente os descendentes que tinham duas titulações pelo programa.

De acordo com Ribeiro (2007) um dos critérios de avaliação dos PPG's pela CAPES é a capacidade do programa em formar mestres e doutores. Desta forma, a Figura 18 indica a atuação do programa na formação de recursos humanos de acordo com os dados do Currículo Lattes, permitindo um rápido acesso a esses números e possibilitando, de maneira global, avaliar os resultados do PPGCEM/UFSCar. Entre os anos de 1983 a 2019 foram apuradas as formações de 568 mestres e entre os anos de 1992 a 2019, de 335 doutores. Ou seja, o *software* conseguiu mapear 76,50% dos descendentes diretos do PPGCEM/UFSCar, conforme números indicados pelo programa, mencionados na Figura 14. Sendo que deste total foi possível apurar que 45 deles fizeram o mestrado e o doutorado no PPGCEM/UFSCar.

Essa amostra apurada pelo *software* AGLattes é considerada satisfatória e suficiente para se realizar as análises referentes ao programa e essa diferença de 298 descendentes diretos, pode ser explicada pela inexistência de Currículos Lattes de alguns pesquisadores, visto que o programa é mais antigo que a própria Plataforma Lattes, e também por erros no preenchimento do currículo, como foi

mencionado nas limitações do uso da GA.

Figura 18: Formação de recursos humanos pelo PPGCEM/UFSCar



Fonte: Elaborado pela autora

4.1.2 Formação por ano e por titulação

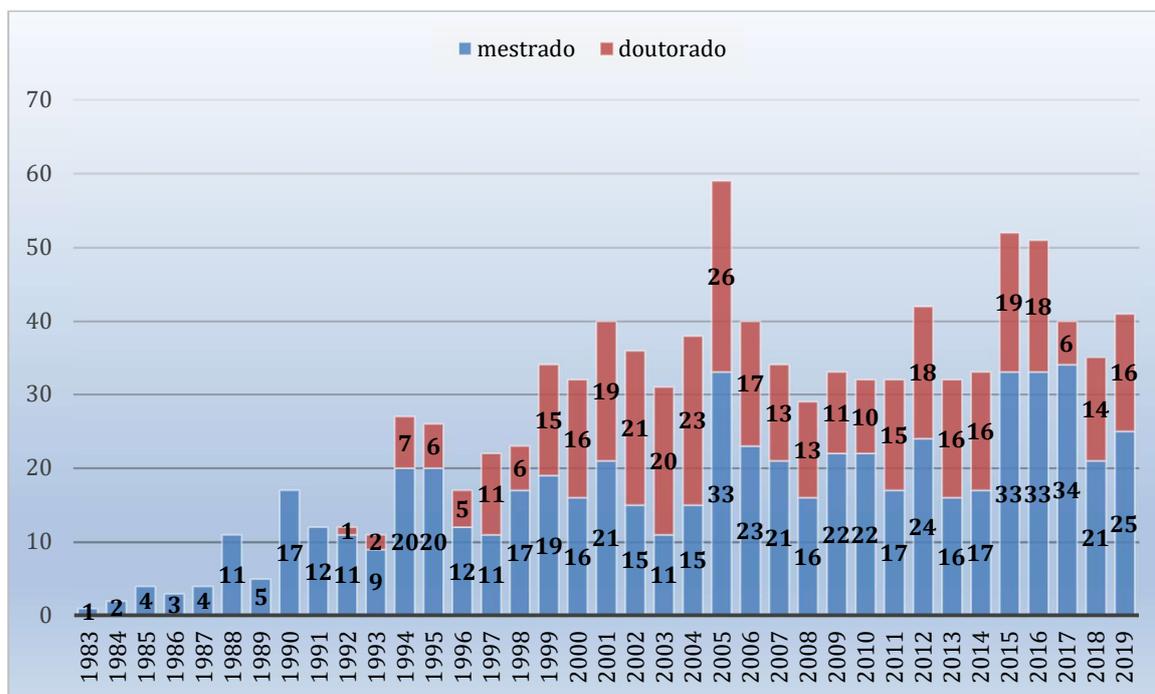
A partir dos 993 registros dos descendentes diretos buscou-se conhecer a atuação do programa na formação de recursos humanos ao longo do tempo e, portanto, foi construído um indicador que representasse a distribuição desta formação durante o período estudado.

A Figura 19 representa este indicador, e através de sua análise pode ser observado que nos anos de 2005, 2015 e 2016 o programa atingiu os melhores resultados, comparados com a série histórica. A partir de 1999 o programa manteve uma média aproximada de 38 formados por ano, ficando abaixo desta média entre os anos de 2007 até 2011, 2013 e 2014. Outra informação que pode ser visualizada é que nos anos de 2002, 2003 e 2004 os números de doutores formados no programa, superaram o número de mestres, porém em 2017 o número de doutores formados teve uma queda, voltando a crescer em 2018 e 2019.

Essa informação pode ser útil para que o programa avalie o seu desempenho diante dos fatores que ocasionaram essas variações dos resultados tanto positivos, quanto negativos. Como o objetivo deste indicador foi verificar a potencialidade dos dados para se obter números sobre a atuação do programa na formação de

recursos humanos ao longo do tempo, esta pesquisa não se aprofundou nos fatores que causaram essa variação, já que podem ter sido motivadas por vários eventos, e não haveria tempo hábil para constatação das mesmas.

Figura 19: Descendentes diretos por ano e por titulação



Fonte: Elaborado pela autora

4.1.3 Indicador de Atuação Profissional

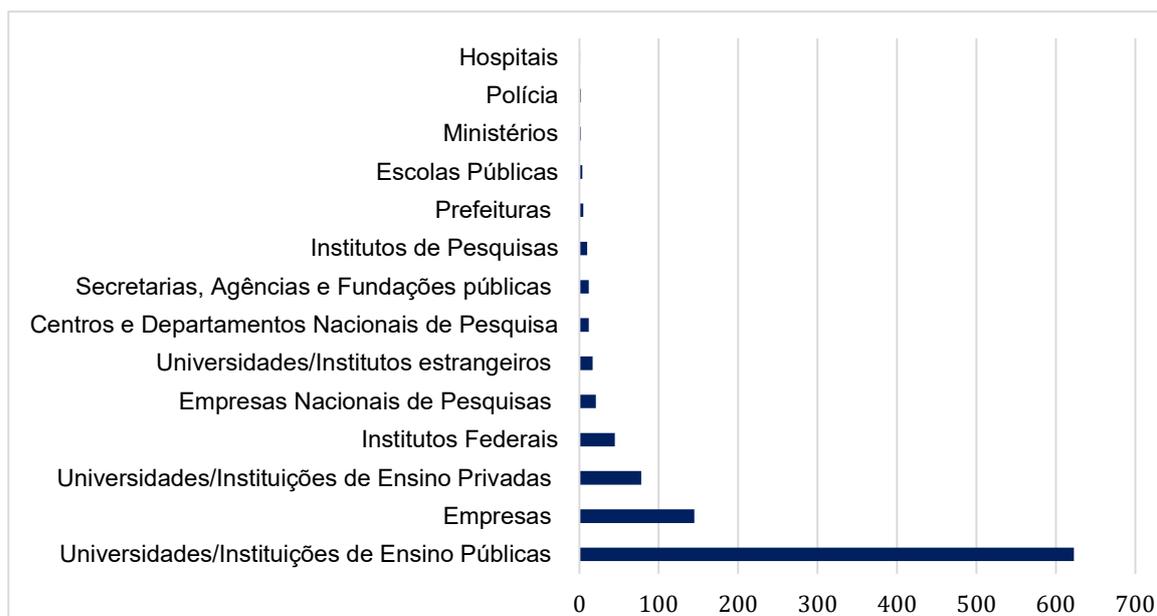
Por meio do endereço profissional contido no Currículo Lattes foi possível identificar em quais tipos de instituições estão atuando os descendentes do PPGCEM/UFSCar. A filtragem de dados foi realizada no VantagePoint e foi criado um thesaurus com os principais tipos de instituições. Para a construção deste indicador foram considerados todos os descendentes diretos e indiretos, totalizando 1503 registros. Deste total, 526 currículos não tinham o endereço profissional indicado, representando aproximadamente 35%. A análise foi realizada sobre 977 currículos, que representou aproximadamente 65% do total e, portanto, uma amostra considerada significativa.

Como pode se observar na Figura 20 a maioria, um total de 623

descendentes, 63,77%, atuam em universidades e instituições de ensino públicas como as universidades federais, estaduais e centros tecnológicos como a Fatec (Faculdade de Tecnologia). Essa atuação no meio acadêmico é o que se espera da formação de mestres e doutores, portanto este número é considerado satisfatório

Também é possível constatar que 145 deles, 14,84%, estão atuando em empresas, se destacando as empresas onde atuam mais de três descendentes: ArcelorMittal Tubarão, Brasken S.A e Petrobrás. Já nas Universidades e Instituições privadas há uma concentração de 7,98% e Institutos Federais, 4,61%. Nas Empresas nacionais de pesquisas, como a Embrapa, se concentram 2,15% dos descendentes. Com pequena representatividade, há descendentes que indicaram em seus currículos endereços profissionais como hospital e Polícia Militar.

Figura 20: Atuação profissional dos descendentes



Fonte: Elaborado pela autora

4.1.4 Formação por orientador

Para a construção deste indicador foram considerados os 993 descendentes diretos. No Quadro 7 pode ser observado o ranking dos docentes que tiveram um maior número de orientados para os que tiveram menor número, desde a fundação do PPGCEM/UFSCar. Se for considerada a quantidade de formação, em destaque

está o docente Victor Carlos Pandolfelli, que contribuiu para a formação de 61 descendentes diretos, constatando a sua importante atuação.

No entanto, se a quantidade de formação for relacionada com o tempo de contribuição para o programa, o melhor índice é da docente Alessandra de Almeida Lucas, que em quatro anos contribuiu para a formação de 11 descendentes diretos, obtendo o maior índice de 2,75. Este indicador pode contribuir na avaliação do desempenho individual de cada docente e ajustes nas diretrizes do programa, conforme necessidade.

Quadro 7: Formação por orientador e tempo de atuação

Docente	Qtd.	Anos	Índice	Docente	Qtd.	Anos	Índice
Victor Carlos Pandolfelli	61	29	2,10	Leonardo Bresciani Canto	11	7	1,57
Claudemiro Bolfarini	57	23	2,48	Alberto Moreira J. Junior	10	22	0,45
Rosário Elida Suman Bretas	54	36	1,50	Ana Candida M. Rodrigues	10	29	0,34
Elias Hage Junior	52	31	1,68	Edson Roberto Leite	10	25	0,40
Luiz Antonio Pessan	39	25	1,56	Hansjürgen Kestenbach	10	26	0,38
Nelson Guedes De Alcântara	39	37	1,05	Maria Zanin	10	14	0,71
Anselmo Ortega Boschi	36	32	1,13	Pedro Iris Paulin Filho	10	26	0,38
Edgar Dutra Zanotto	34	37	0,92	Silvia Helena Prado Bettini	10	7	1,43
Sebastiao Vicente C. Junior	34	33	1,03	Conrado Ramos M. Afonso	9	7	1,29
Claudio Shyinti Kiminami	33	29	1,14	Roberto Tomasi	9	22	0,41
José Augusto M. Agnelli	33	28	1,18	Rodrigo Bresciani Canto	9	3	3,00
Márcio Raymundo Morelli	31	19	1,63	Adhemar Colla R. Filho	8	24	0,33
Oscar Balancin	28	34	0,82	José Eduardo Spinelli	8	7	1,14
Dulcina Maria P. F. De Souza	27	31	0,87	Sati Manrich	8	4	2,00
Ruth Herta G. A. Kiminami	26	29	0,90	Rinaldo Gregório Filho	6	26	0,23
Jose De Anchieta Rodrigues	25	36	0,69	Vitor Luiz Sordi	6	3	2,00
Jose Alexandrino De Sousa	24	38	0,63	Jose Arana Varela	5	28	0,18
Luiz Henrique C. Mattoso	24	26	0,92	Jose Antonio Eiras	4	20	0,20
Walter Jose Botta Filho	24	32	0,75	Daniel Rodrigo Leiva	3	3	1,00
Sebastião Elias Kuri	19	32	0,59	Jose Donato Ambrósio	3	5	0,60
Joao Baptista Baldo	17	29	0,59	Carlos Alberto D. Rovere	2	2	1,00
Jose Angelo R. Gregolin	15	25	0,60	Lidiane Cristina Costa	2	1	2,00
Tomaz Toshimi Ishikawa	15	33	0,45	Piter Gargarella	2	2	1,00
Pedro A. De Paula Nascente	14	27	0,52	Carlos H. Scuracchio	1	2	0,50
Elson Longo Da Silva	13	27	0,48	Guilherme Zepon	1	1	1,00
Levi De Oliveira Bueno	13	32	0,41	Jose Manoel Marconcini	1	3	0,33
Maurizio Ferrante	13	35	0,37	Marcello R. B. Andreeta	1	3	0,33
Alessandra de A. Lucas	11	4	2,75	Marco A. L. Cordeiro	1	0	1,00
Carlos Ventura Dalkaine	11	28	0,39	Walter Libardi	1	18	0,06

Fonte: Elaborado pela autora

4.1.5 Aderência ao programa após a formação

Da mesma forma foi possível identificar que dos 993 descendentes diretos formados pelo PPGCEM/UFSCar, 18 deles foram absorvidos pelo próprio programa após a formação, resultando no indicador de aderência, representando 31% do total das 58 sementes, como pode ser observado no Quadro 8, onde também é possível identificar quem são essas pessoas, quem foram os seus orientadores, o ano e a titulação que obtiveram no programa. Esses números refletem positivamente para o programa, o que significa que seus descendentes estão dando continuidade nas pesquisas e ideias de seus antecessores, impulsionando também a área.

Ao realizar a análise do Quadro 8 e do Quadro 7 em conjunto, pode-se observar também, que o docente Elias Hage Junior orientou 52 alunos, sendo que 4 deles foram absorvidos pelo programa após a formação: José Donato Ambrósio, Lidiane Cristina Costa, Luiz Antonio Pessan e Sati Manrich. E destes resultaram em outros mais 52 descendentes diretos, constatando a importante atuação desses descendentes para o programa.

Quadro 8: Acadêmicos que aderiram ao programa após formação

ANO FORMAÇÃO	TITULAÇÃO	DESCENDENTE DIRETO	ORIENTADOR
1997	DOUTORADO	ALBERTO M. JORGE JR	OSCAR BALANCIN
1998	MESTRADO	ALESSANDRA LUCAS	ROSARIO ELIDA S. BRETAS
2011	DOUTORADO	CARLOS A. D. ROVERE	SEBASTIÃO ELIAS KURI
2003	DOUTORADO	CARLOS SCURACCHIO	ROSARIO ELIDA S. BRETAS
2004	DOUTORADO	CONRADO R. M. AFONSO	CLAUDIO SHYINTI KIMINAMI
2009	DOUTORADO	DANIEL RODRIGO LEIVA	WALTER JOSE BOTTA FILHO
1990	MESTRADO	EDSON ROBERTO LEITE	ELSON LONGO DA SILVA
2016	DOUTORADO	GUILHERME ZEPON	CLAUDEMIRO BOLFARINI
2007	DOUTORADO	JOSÉ DONATO AMBRÓSIO	ELIAS HAGE JUNIOR
2004	DOUTORADO	LEONARDO B. CANTO	LUIZ ANTONIO PESSAN
2012	DOUTORADO	LIDIANE CRISTINA COSTA	ELIAS HAGE JUNIOR
1987	MESTRADO	LUIZ ANTONIO PESSAN	ELIAS HAGE JUNIOR
1988	MESTRADO	LUIZ H. C. MATTOSO	SEBASTIAO V. CANEVAROLO JR
2013	DOUTORADO	MARCO A. L. CORDEIRO	EDSON ROBERTO LEITE
2009	MESTRADO	PITER GARGARELLA	CLAUDIO SHYINTI KIMINAMI
1993	DOUTORADO	SATI MANRICH	ELIAS HAGE JUNIOR
1997	DOUTORADO	SILVIA HELENA P. BETTINI	JOSÉ AUGUSTO M. AGNELLI
1994	DOUTORADO	VITOR LUIZ SORDI	LEVI DE OLIVEIRA BUENO

Fonte: Elaborado pela autora

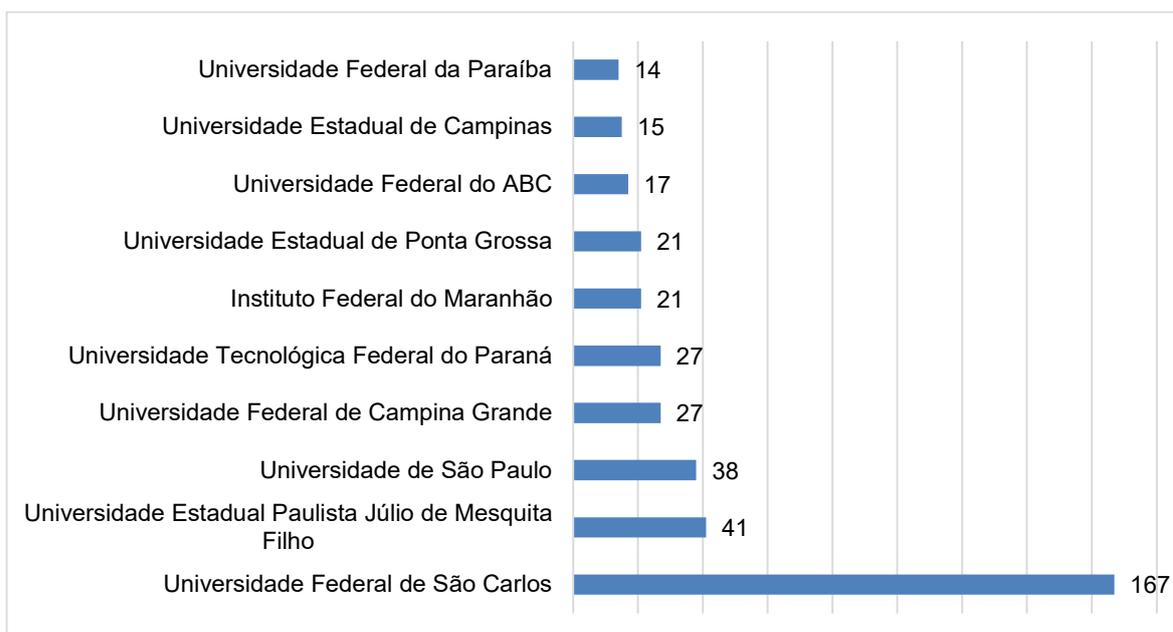
4.1.6 Atuação por instituição

Este indicador foi desenvolvido a partir dos dados de endereço profissional contidos no Currículo Lattes dos descendentes diretos e indiretos do PPGCEM/UFSCar, nos quais foram informadas as instituições onde eles trabalham. Para tanto foram selecionadas as dez instituições onde mais se concentram os descendentes do programa.

Pode se observar na Figura 21 que a instituição onde mais se concentra os descendentes do PPGCEM/UFSCar é a própria instituição do programa com 167 descendentes. Ainda no Estado de São Paulo, a Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP) e Universidade de São Paulo (USP) concentram-se 41 e 38 descendentes, respectivamente e depois a Universidade Federal do ABC com 17 descendentes e Universidade Estadual de Campinas (UNICAMP) com 15 descendentes. As demais instituições nos outros estados brasileiros, pode-se constatar que duas universidades de Paraíba estão entre as dez instituições que mais empregam os descendentes do PPGCEM/UFSCar.

Este indicador pode contribuir para identificar o destino dos descendentes do programa e o impacto social.

Figura 21:Atuação por Instituição

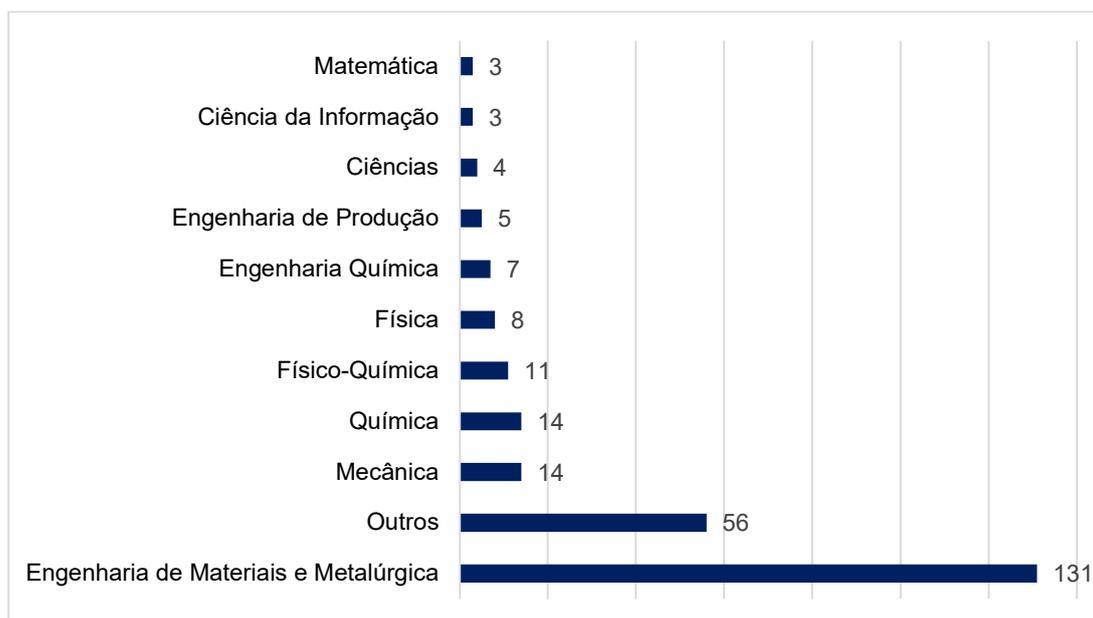


Fonte: Elaborado pela autora

4.1.7 Indicador de atuação por departamento

No indicador representado pela Figura 22, pode ser constatado o departamento de atuação dos descendentes diretos e indiretos do PPGCEM/UFSCar. A classificação foi feita através de um agrupamento de acordo com os departamentos mais indicados, de forma que departamentos com duas ou menos indicações foram classificados em “outros”. Observa-se que 131 deles indicaram em seu Currículo Lattes que atuam no departamento de Engenharia de Materiais e Metalúrgica. Com este indicador pode-se verificar as áreas em que eles estão atuando e auxiliar na compreensão da continuidade da área de Engenharia de Materiais e Metalúrgica pelos descendentes após a formação

Figura 22: Atuação por Departamento



Fonte: Elaborado pela autora

4.1.8 Localização dos descendentes por país

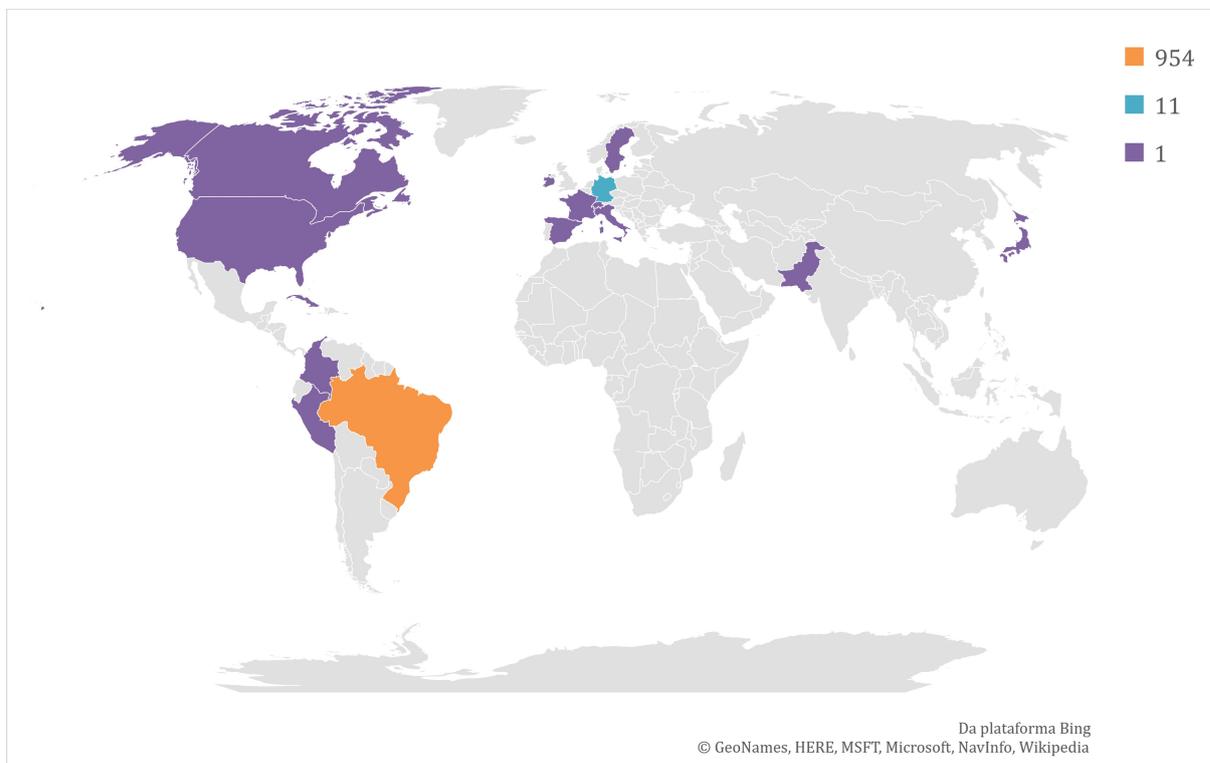
Para a construção deste indicador foram considerados todos os descendentes do programa, tanto os diretos como os indiretos, totalizando 1503

registros. Através do país indicado no endereço profissional do Currículo Lattes foi realizada uma contagem, permitindo identificar a concentração de descendentes por país, tendo por objetivo demonstrar o impacto internacional do programa.

Do total, 537 currículos dos descendentes não consta o país no endereço profissional. Observou-se na Plataforma Lattes que há um controle de preenchimento dos campos referentes ao endereço profissional, sendo alguns campos obrigatórios. No entanto não se sabe a partir de quando esse controle passou a existir., o que pode explicar a ausência do país nos currículos.

Conforme pode ser observado na Figura 23, o PPGCEM/UFSCar tem descendentes em vários países, onde a maioria, um total de 954, se encontra no Brasil, representados pela cor laranja. Em destaque está também a Alemanha, onde podem ser encontrados 11 descendentes, representados pela cor azul. Também há um descendente em cada um desses países: Canadá, Colômbia, Cuba, Espanha, Estados Unidos, França, Irlanda, Itália, Japão, Paquistão, Peru, Suíça e Suécia representados pela cor roxa.

Figura 23: Localização dos descendentes por país



Fonte: Elaborado pela autora

4.1.9 Localização dos descendentes por regiões brasileiras

Visto que, a maioria dos descendentes diretos do PPGCEM/UFSCar atuam no Brasil, considerou-se importante criar um indicador também com base no Endereço Profissional extraído dos currículos, identificando como eles estão distribuídos no país.

A Figura 24 ilustra o resultado deste indicador, onde pode-se observar que seus descendentes estão distribuídos em quase todo território brasileiro, presentes em 23 estados e no Distrito Federal. A maior concentração desses está no estado de São Paulo, seguido pelo estado do Paraná e Minas Gerais.

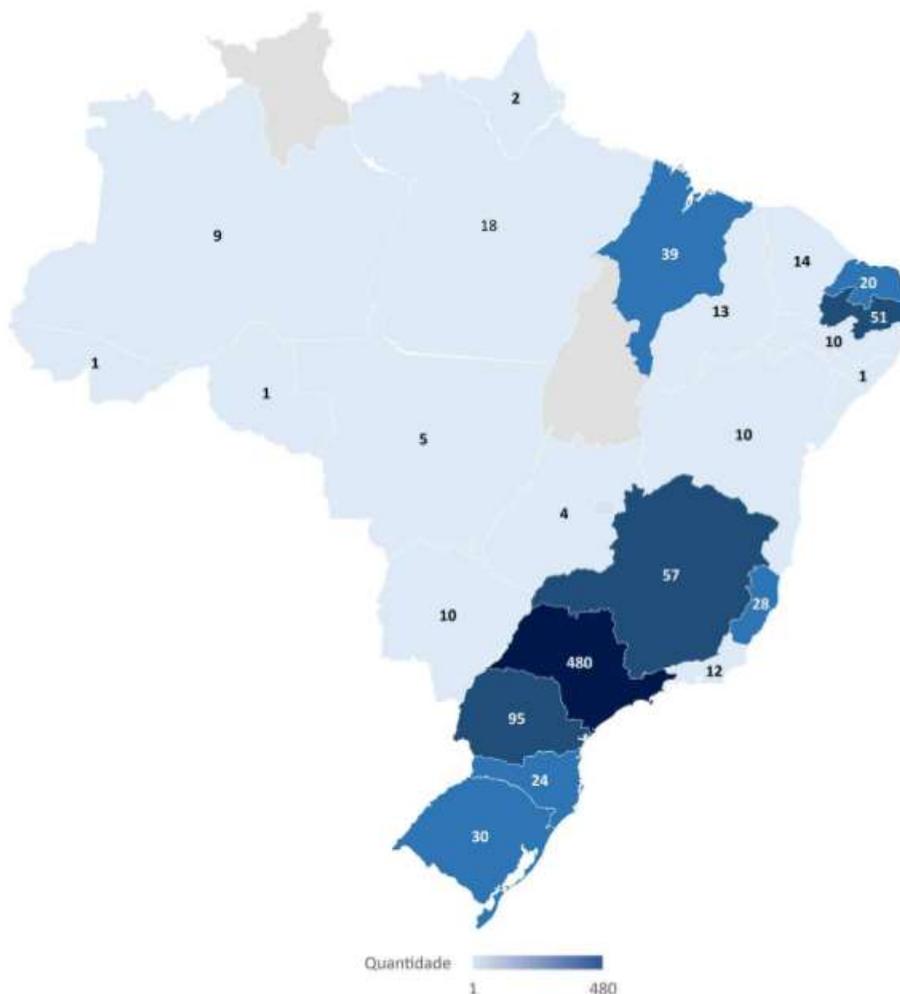
Observa-se, que a região nordeste possui 170 descendentes, 17,82% do total, apesar da grande distância da sede do programa. Nesta região, se destaca o estado de Paraíba, com 51 descendentes, 30%, do total da região.

Em João Pessoa encontra-se a Universidade Federal de Paraíba e em Campina Grande a Universidade Federal de Campina Grande, renomadas instituições na área de Engenharia de Materiais e Metalúrgica, o que pode explicar essa propagação do conhecimento.

Outro fator que pode explicar a concentração de descendentes nessa região é que de acordo com a CAPES (2019) a UFSCar é uma instituição promotora, através de seu PPGCEM, do Programa de Doutorado Interinstitucional – Dinter, que tem por objetivo viabilizar a formação a nível de doutorado os docentes do quadro permanente de instituições distantes dos grandes centros de ensino à pesquisa, de modo a diminuir as assimetrias existentes, fomentar a produção acadêmica e fortalecer as linhas de pesquisas que respondam às demandas relacionadas ao desenvolvimento local e regional, contribuindo para o desenvolvimento da área de Engenharia de Materiais e Metalúrgica nessa região.

Desta forma, por meio deste indicador pode ser constatado o impacto do PPGCEM/UFSCar a nível nacional, além de poder contribuir para estudos mais aprofundados no desenvolvimento regional e diminuição das diferenças científico-tecnológicas entre as regiões brasileiras.

Figura 24: Localização dos descendentes por regiões brasileiras



Fonte: Elaborado pela autora

4.1.10 Cidades com maior concentração de descendentes

Como foi constatado pelo indicador das regiões brasileiras, a maioria dos descendentes do programa está localizada no Estado de São Paulo. Logo, entendeu-se a necessidade da criação deste indicador, detalhando a análise por cidade, para verificar se havia uma cidade específica com maior concentração de descendentes do programa e entender o fluxo do conhecimento produzido.

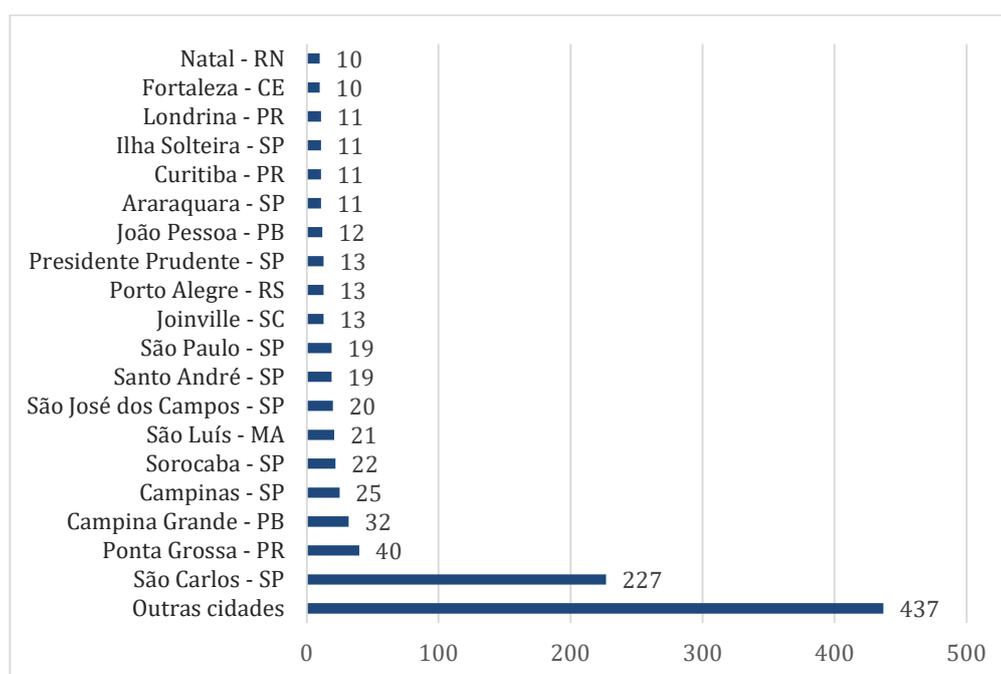
Como ilustrado na Figura 25 a grande maioria encontra-se na cidade de São Carlos – SP, cidade sede do PPGCEM/UFSCar onde se encontram 227 descendentes. As justificativas, hipotéticas, deste número são: A UFSCar tem sua sede em São Carlos; a cidade é considerada um polo tecnológico, o que propicia a absorção destes descendentes pelo mercado, em São Carlos estão localizadas

outras instituições de ensino e por fim, a cidade possui instituições de Ciência e Tecnologia renomadas a nível internacional: UFSCar, USP e Embrapa, o que favorece a fixação dos descendentes do programa na cidade. Constata-se, também, uma concentração de descendentes na cidade de Ponta Grossa – PR com 40 e Campina Grande – PB, com 32 até chegar em Fortaleza – CE e Natal – RN com dez cada, confirmando a importância do PPGCEM/UFSCar para a propagação da área e do desenvolvimento da pós graduação na região dessas cidades.

Na cidade com maior concentração, São Carlos, foi constatado que 68,28% dos descendentes atuam na Universidade Federal de São Carlos, 12,33% atuam na Universidade de São Paulo, 4% atuam na Embrapa, o que confirma o que foi afirmado no indicador anterior e mostra a visibilidade e importância do PPGCEM/UFSCar na formação de pesquisadores para a própria Universidade.

Na sua sequência, está a cidade de Ponta Grossa onde foi possível constatar que 52,5% dos descendentes atuam na Universidade Estadual de Ponta Grossa e 17,5% atuam na Universidade Tecnológica Federal de Ponta Grossa, constatando a importância do PPGCEM/UFSCar para o desenvolvimento da região destas cidades.

Figura 25: Cidades com maior concentração de descendentes



Fonte: Elaborado pela autora

4.1.11 Áreas de formação de mestres e doutores

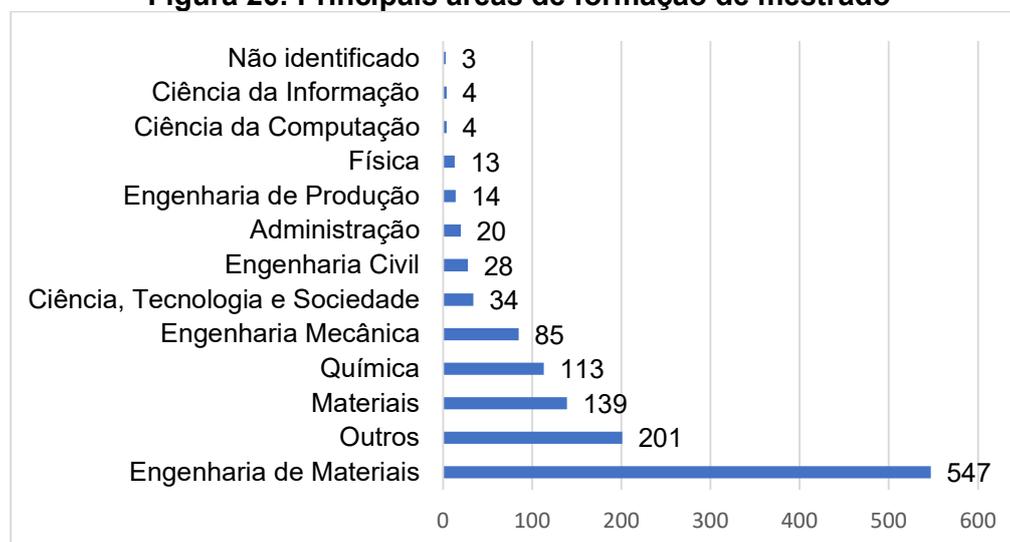
Como visto, a continuidade dos trabalhos de pesquisas desenvolvidos por um PPG depende da continuidade das pesquisas por seus formados. Desta forma, a formação de mestres e doutores na área precisa ser contínua. Com base nisto, foi criado o indicador de área de formação somente dos descendentes indiretos, já que os descendentes diretos pertencem, por formação, a área de Engenharia de Materiais e Metalúrgica.

Este indicador foi desenvolvido com o objetivo de avaliar se a continuação de formação na área está sendo mantida. A tendência é que um aluno que se torne mestre ou doutor na área formem outros mestres e doutores também na área.

A partir do curso indicado pelos descendentes indiretos na Formação Acadêmica/Titulação em seus Currículos Lattes, foi feita uma classificação para mensurar as áreas mais indicadas por eles, as quais podem ser conferidas na Figura 26, onde 547 relacionaram a sua formação de mestrado ou doutorado com o termo “Engenharia de Materiais”, ou seja, 45%.

Há aqueles que relacionaram suas áreas com o termo “materiais” totalizando 139 acadêmicos. Este indicador possibilita observar a contribuição do PPGCEM/UFSCar na formação de recursos humanos na área de Engenharia de Materiais e Metalúrgica, o que pode contribuir na continuidade e no direcionamento das estratégias do programa.

Figura 26: Principais áreas de formação de mestrado

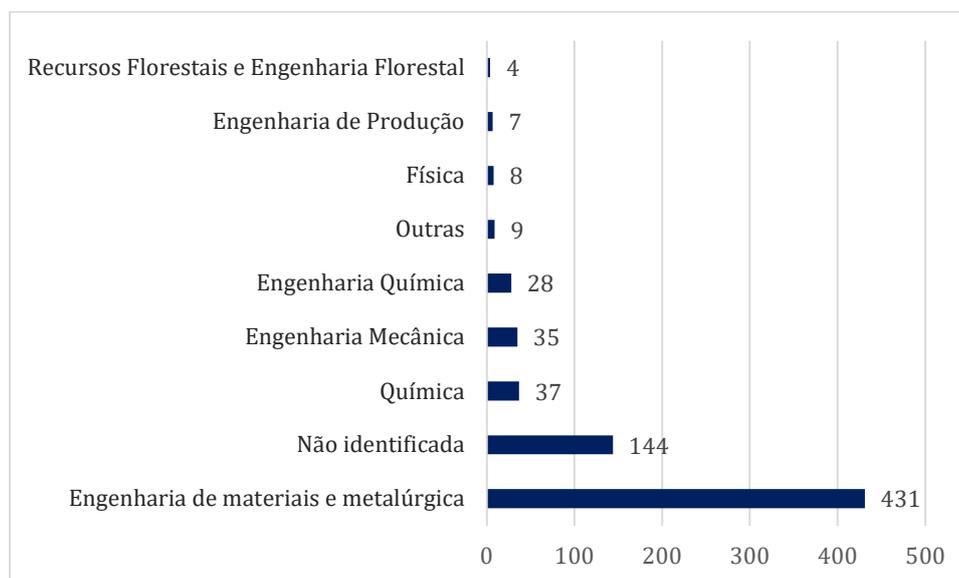


Fonte: Elaborado pela autora

Da mesma forma que o mestrado, também foi levantado as áreas de formação dos descendentes que concluíram o doutorado, desta vez considerando todos os descendentes. Como pode ser observado na Figura 27 um total de 729 titulações de doutorado, 431 ou 59% das titulações foram na área de Engenharia de Materiais e Metalúrgica.

Não tiveram suas áreas identificadas apresentaram um total de 20%. Química e Engenharia Mecânica tem uma representação de 5%, seguidas pela área de Física e Engenharia de Produção com uma representatividade de 1%. As outras áreas Administração, Engenharia Civil; Engenharia Agrícola, Engenharia Nuclear, Matemática, Medicina, Odontologia, Ciência da computação, Ciência e Tecnologia de Alimentos, Engenharia Aeroespacial, Engenharia Elétrica, Engenharia Sanitária, Genética, GeoCiências, Geografia, História e Letras somam 5% do total.

Figura 27: Principais áreas de formação de doutorado



Fonte: Elaborado pela autora

4.2 INDICADORES DE PUBLICAÇÕES CIENTÍFICAS

Durante o desenvolvimento deste indicador, percebeu-se o equívoco por partes dos descendentes quanto a classificação de áreas, onde muitas classificações das subáreas não tinham nenhuma relação com a área. Ao buscar

por respostas para esta questão na Plataforma Lattes, foi constatado que há um há um controle pelo sistema para que o usuário desta plataforma cadastre e classifique corretamente a área para as suas publicações, ficando livremente a classificação de especialidades. Desta forma, acredita-se que as publicações com divergências na classificação de áreas tenham ocorrido antes deste controle.

Portanto, para a criação do subconjunto de dados da área de Engenharia de Materiais e Metalúrgica, foram consideradas as publicações que foram classificadas na área ou subárea ou especialidade de acordo com a Árvore do Conhecimento - Áreas do conhecimento – Engenharias do CNPq. Todas as demais, que não se enquadraram na área citada, foram classificadas como Outras Áreas.

4.2.1 Tipos de publicações

Do total de 1507 áreas apuradas, foi feita uma classificação que considerou a área de Engenharia de Materiais e Metalúrgica, as suas subáreas e especialidades, de acordo com áreas de conhecimento da Plataforma Sucupira (2020), separando-se desta forma a área do programa das demais áreas.

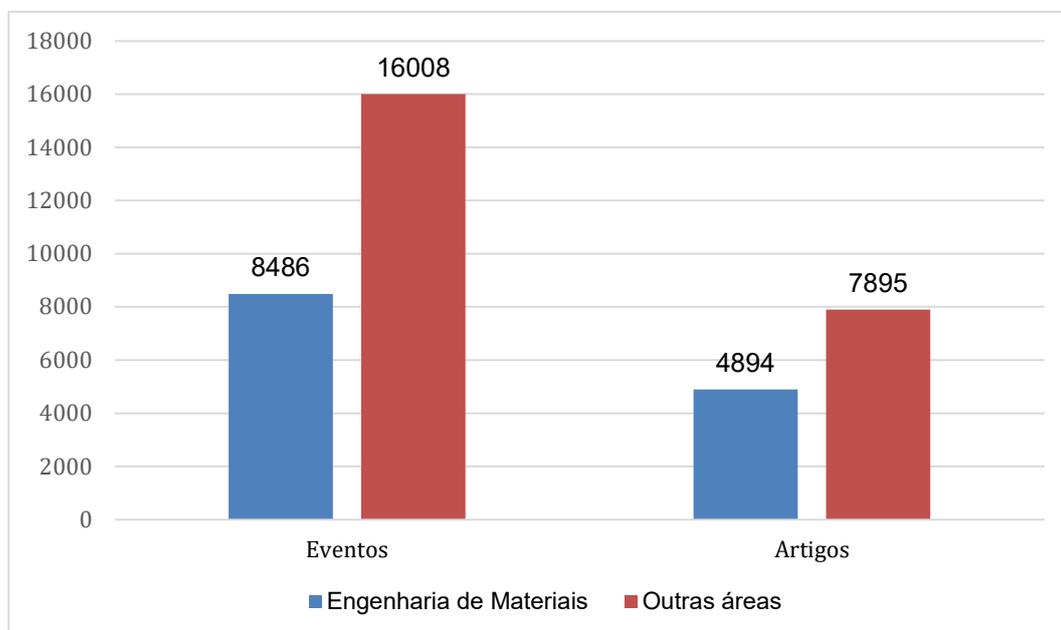
O primeiro indicador a partir destes dados, permite realizar a análise dos registros bibliográficos relativos aos artigos publicados em periódicos e eventos científicos, fazendo um comparativo com as publicações realizadas na área de Engenharia de Materiais e Metalúrgica e nas demais áreas, entre os anos de 1979 até a data de extração de dados.

Após a extração de dados foram apuradas 24.494 publicações em eventos científicos. Através da Figura 28 é possível observar que a área de Engenharia de Materiais e Metalúrgica, que é a área de concentração do PPGCEM/UFSCar, teve 8.486 publicações, que representa 34,6% do total. Já nas outras áreas em conjunto, apurou-se 16.008 publicações, 65,34%. Quanto as publicações de artigos em periódicos, esses somam um total de 12.789 e a área do programa representa 38,26% e nas demais áreas 61,74%.

Considerando o grande número de áreas indicadas, o número de publicações na área do referido programa se torna expressivo. A concentração de publicações em outras áreas pode estar relacionada com o fato de os descendentes do

programa serem absorvidos por outros programas de pós-graduação de áreas correlatas, como a Química.

Figura 28: Tipos de publicações



Fonte: Elaborado pela autora

4.2.2 Publicações de todos os tipos por ano e áreas

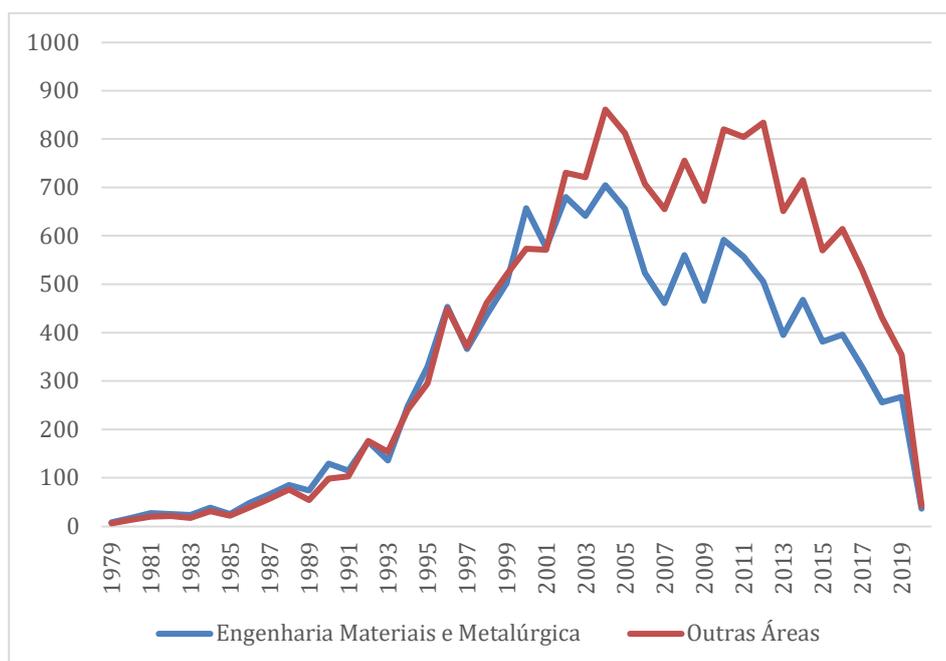
Este indicador foi desenvolvido com o objetivo de apurar a quantidade de publicações de todos os descendentes, por tipo, ano e de acordo com a área em que foram classificadas. Na Figura 29, pode ser observado o resultado da comparação das publicações na área de Engenharia de Materiais e Metalúrgica com as publicações realizadas em outras áreas.

Percebe-se que até o ano de 2002 as pesquisas na área do programa e em outras áreas permaneceram niveladas e depois passa a ter diferença entre os números, quando as pesquisas em outras áreas foram maiores que na área do programa. Isto pode ser o resultado pela procura do programa por pesquisadores de outras áreas. Somado ao fato de que esses pesquisadores publicam com seus orientadores e com o tempo tendem a voltar para suas áreas. Outro fato, é de que esses descendentes sejam absorvidos por outros programas e suas publicações se

voltem para a área dele.

Este indicador pode ser útil para medir os esforços de pesquisas na área motivando a criação de novos indicadores para conhecer as principais áreas de interesse desses descendentes.

Figura 29: Quantidade de publicações por ano e por áreas



Fonte: Elaborado pela autora

4.2.3 Publicações da área do PPGCEM/UFSCar

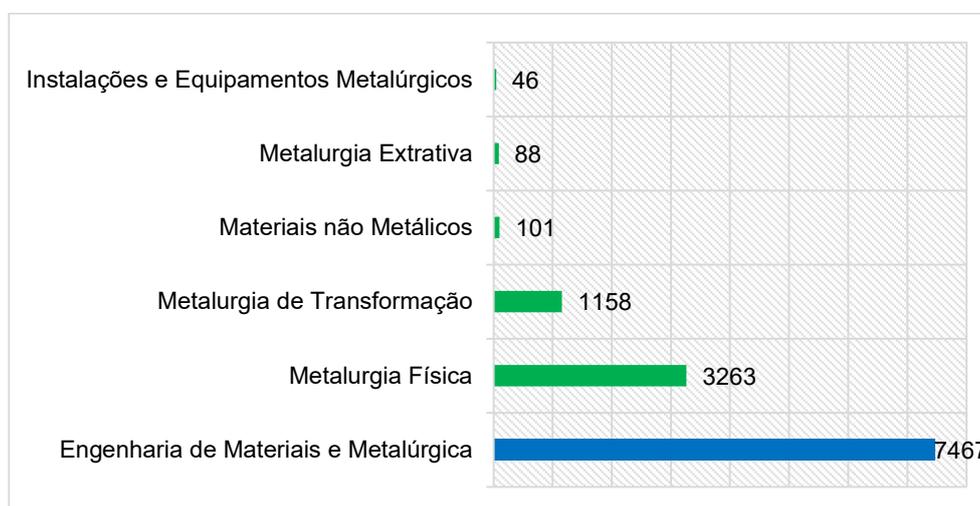
Este indicador revela a quantidade de publicações dos descendentes que indicaram a área de Engenharia de Materiais e Metalúrgica e das respectivas subáreas em que suas pesquisas foram realizadas.

Este indicador foi desenvolvido realizando uma filtragem das áreas cadastradas pelos descendentes do programa no Currículo Lattes, foi feita uma contagem das indicações da área da Engenharia de Materiais e Metalúrgica e das suas subáreas de acordo com a árvore do conhecimento do CNPq. De acordo com o CNPq (2020) as subáreas que compõe a área de Engenharia de Materiais e Metalúrgica são: Metalurgia Física, Metalurgia de Transformação, Materiais não

Metálicos, Metalurgia Extrativa, Instalações e Equipamentos Metalúrgicos.

A quantidade de vezes que a área e subáreas foram indicadas podem ser conferidas na Figura 30, onde pode ser observado que a área de Engenharia de Materiais e Metalúrgicas foram indicadas em 7.467 publicações. No entanto a subárea Metalurgia Física foi indicada em 3.263 publicações, seguida da Metalurgia de Transformação em 1.158 publicações, Materiais não Metálicos em 101 publicações, Metalurgia Extrativa em 88 publicações e por último a subárea Instalações e Equipamentos Metalúrgicos que foi indicada em 46 publicações.

Figura 30:Quantidade de publicações por área e subáreas



Fonte: Elaborado pela autora

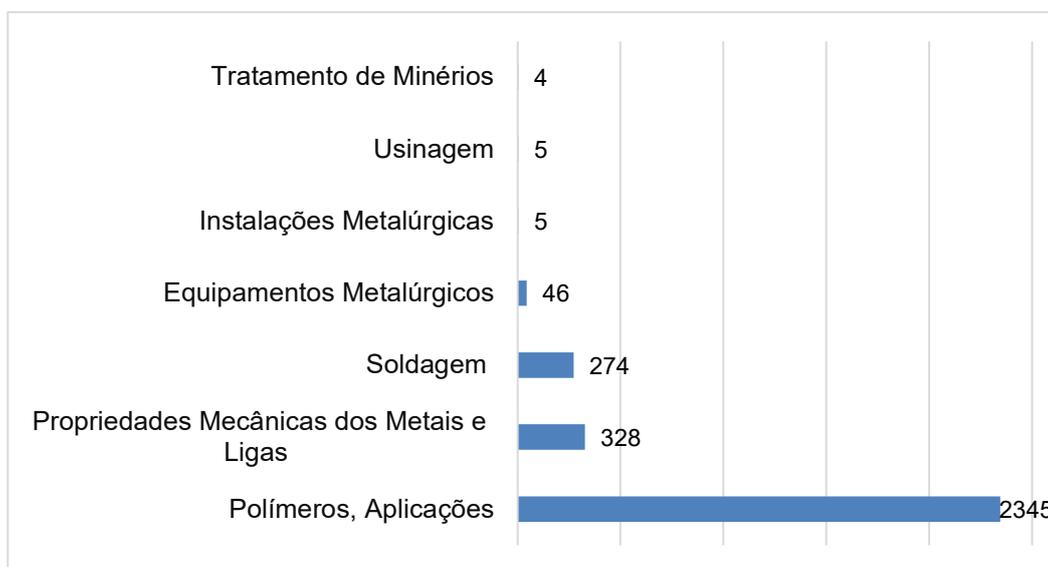
4.2.4 Publicações por especialidade

Com a mesma motivação que o indicador anterior, buscou-se conhecer como os descendentes do programa classificam suas pesquisas de acordo com as especialidades, com o objetivo de conhecer a especialidade que se concentra um maior esforço de pesquisa. Para a construção deste indicador foram consideradas as publicações que continham em suas indicações o nome exato da especialidade e que tenham alguma relação com a área de Engenharia de Materiais e Metalúrgica ou suas subáreas.

Como pode ser observado na Figura 31 a maior especialidade indicada pelos descendentes no cadastro de suas publicações é a Polímeros, Aplicações com

2.345 indicações, seguidas pelas especialidades Propriedades Mecânicas dos Metais e Ligas com 328 indicações, Equipamentos Metalúrgicos com 46 indicações; Instalações Metalúrgicas com 5 indicações, Usinagem com 5 indicações e Tratamento de Minérios com 4 indicações. As demais especialidades da área de Engenharia de Materiais e Metalúrgica não apresentaram indicações.

Figura 31: Quantidade de publicações por especialidade



Fonte: Elaborado pela autora

4.2.5 Periódicos com maiores publicações

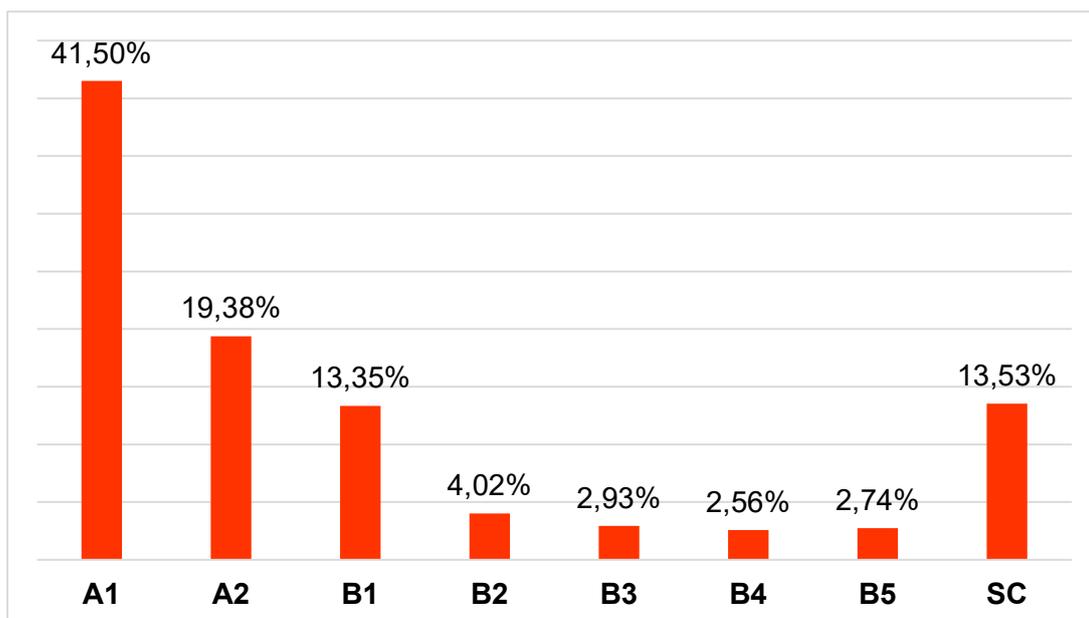
Para analisar a qualidade das produções intelectuais mais recentes dos descendentes do programa, foram utilizadas como amostras as publicações do período 2017 – 2020. Por meio da ferramenta VantagePoint essas produções foram agrupadas de acordo com os veículos onde foram divulgadas. A partir do ISSN desses veículos foi feita uma pesquisa na Plataforma Sucupira para verificar suas classificações de acordo com os instrumentos avaliativos da CAPES.

Segundo a Plataforma Sucupira (2020) Qualis afere a qualidade dos artigos e de outros tipos de produção, a partir da análise de qualidade dos veículos de divulgação. No entanto, o novo modelo do Qualis Referência, se encontra em fase de discussão e aprimoramento, desta forma, tomou-se como referência a

classificação dos veículos no Quadriênio de 2013-2016, onde foram estabelecidos oito estratos de qualidade A1, A2, B1, B2, B3, B4, B5 e C, sendo A1 atribuído como nível mais elevado e C como nível mais baixo.

Na área de Engenharia de Materiais e Metalúrgica somaram-se 547 publicações em revistas científicas no Quadriênio de 2017 – 2020. Essas publicações foram agrupadas acordo com a revista científica e quantidade, que apresentavam duas ou mais publicações. Com isso, obteve-se o indicador representado pela Figura 32, onde pode-se constatar que 60,88% das publicações foram realizadas em veículos classificados como A1 e A2 que correspondem a maior pontuação nos indicadores de produtividade intelectual da CAPES.

Figura 32: Publicações por Qualis Periódico



Fonte: Elaborado pela autora

No Quadro 9 podem ser conferidas as revistas científicas que mais receberam publicações na área de Engenharia de Materiais e Metalúrgica no Quadriênio de 2017 – 2020.

Pode-se observar que o periódico científico que se destaca com 52 publicações no Quadriênio 2017 – 2020 é o *Ceramics International*, o qual foi classificado como A1 na última avaliação da CAPES. Este indicador tem a finalidade de atestar a importância da produção intelectual dos PPG's.

Quadro 9: Principais Revistas Científicas

ISSN	Publicações	Qualis	Revista	Qualis
0272-8842	52	A1	Ceramics International	A1
1516-1439	20	A2	Materials Research (São Carlos. Impresso)	A2
0925-8388	11	A1	Journal of Alloys and Compounds	A1
1980-5373	11	A2	Materials Research	A2
0957-4522	10	A2	Journal of Materials Science. Materials in Electronics	A2
1678-4553	10	B1	Cerâmica	B2
1517-7076	10	B2	Matéria (Ufrj)	B1
1868-2405	9	B4	Refractories Worldforum	B4
0022-3093	8	A1	Journal of Non-Crystalline Solids	A1
0254-0584	8	A1	Materials Chemistry and Physics	A1
0360-3199	8	A1	International Journal of Hydrogen Energy	A1
1932-7447	8	A1	Journal of Physical Chemistry. C	A1
0021-8995	7	A1	Journal of Applied Polymer Science (Print)	A1
0955-2219	7	A1	Journal of The European Ceramic Society	A1
2053-1591	7	A2	Materials Research Express	B5
1809-8797	7	B5	Revista Eletrônica de Materiais E Processos (Ufcg)	A2

Fonte: Elaborado pela autora

4.2.6 Autores com maior publicação por áreas

Para a construção deste indicador foram comparados os dez autores com maiores publicações na área de Engenharia de Materiais e Metalúrgica com os dez autores com maiores publicações em outras áreas. Após a classificação das publicações por área, foi feita uma lista simples no VantagePoint, possibilitando conhecer os principais autores com maior índice de produções científicas da área e o quanto eles contribuem para as demais áreas.

Pode-se observar na Tabela 1, que o autor Elson Longo da Silva apresenta o maior número de publicações tanto na área de Engenharia de Materiais e Metalúrgica como em outras áreas. Todos os autores da tabela são sementes do programa, exceto dois, a Ana Cristina Figueiredo de Melo Costa e Tomás Jeferson Alves de Mélo. Ambos são descendentes direto do PPGCEM/UFSCar e atualmente atuam na Universidade Federal de Campina Grande, Centro de Ciências e Tecnologia, Departamento de Engenharia de Materiais.

Tabela 1: Autores com maior quantidade de publicações por áreas

Engenharia de Materiais e M.	Total	Outras áreas	Total
Elson Longo da Silva	1237	Elson Longo da Silva	1050
Jose Arana Varela	887	Ana Cristina F. de M. Costa	759
Victor Carlos Pandolfelli	820	Jose Arana Varela	754
Ana Cristina F. de M. Costa	720	Ruth Herta G. A. Kiminami	698
Ruth Herta G.A. Kiminami	678	Walter Jose Botta Filho	522
Walter Jose Botta Filho	625	Claudio Shyinti Kiminami	511
Claudio Shyinti Kiminami	484	Márcio Raymundo Morelli	453
Edson Roberto Leite	466	Edson Roberto Leite	453
Anselmo Ortega Boschi	425	Tomás Jeferson A. de Melo	434
Luiz Henrique C. Mattoso	398	Pedro A. de Paula Nascente	422
Total	6740	Total	6056

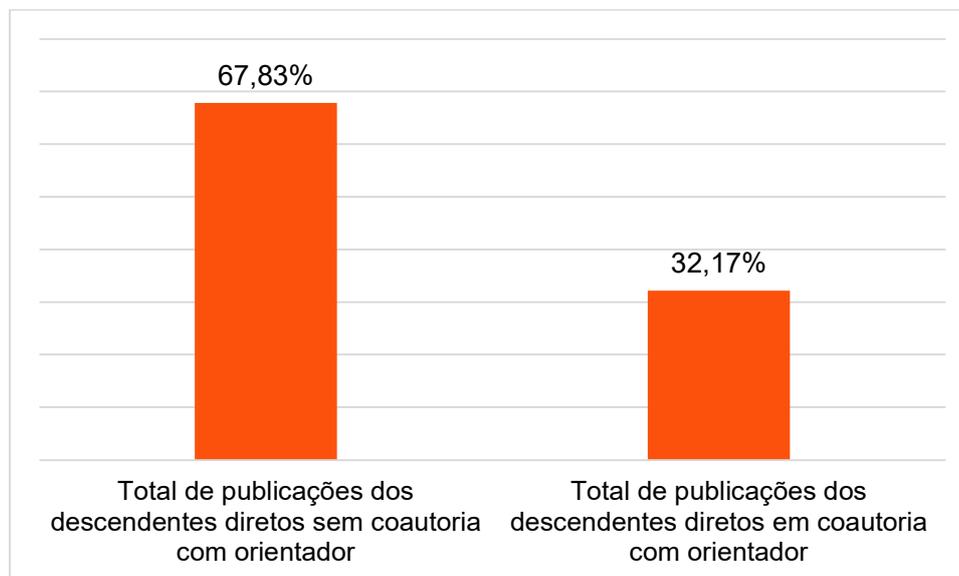
Fonte: Elaborado pela autora

4.2.7 Publicações em coautoria com o orientador

Com o objetivo de analisar por quanto tempo que os descendentes do programa publicaram em coautoria com seus orientadores do próprio programa, foi desenvolvido o indicador de publicações em coautoria. A análise foi feita sobre os descendentes diretos do programa, considerando todas as publicações.

Para chegar no resultado, foi feita uma matriz de publicações no VantagePoint, relacionando as publicações por meio do *idlattes* dos autores. Os números foram transferidos para uma planilha eletrônica, onde foram selecionados apenas os descendentes diretos nas linhas da planilha e os respectivos orientadores nas colunas. No caso de dupla titulação no programa, foi considerada a formação de mestrado.

Dos 586 descendentes diretos com *idlattes* identificados, 481 deles apresentam publicações científicas indicadas nos Currículos Lattes, com um total de 20.478 publicações. Desse total 32,17% foram em coautoria com o orientador, totalizando 6588 publicações, como pode ser observado na **Figura 33**.

Figura 33: Publicações em coautoria com orientador

Fonte: Elaborado pela autora

4.2.8 Análise da relação por meio das publicações científicas

A partir da mesma matriz mencionada no indicador 4.2.7, foi feita uma conferência manual no VantagePoint para se conhecer em quais anos as produções científicas foram realizadas em coautoria com seu orientador. Com isso foi possível avaliar por quanto tempo os descendentes diretos publicaram em coautoria com seus orientadores após a formação e desta forma, mensurar por quanto tempo as ideias desses pesquisadores foram influenciadas diretamente pelos seus orientadores.

Com os números obtidos, foram feitas duas análises. Uma análise contabilizou-se apenas os anos em que houve alguma publicação em coautoria dos descendentes diretos e seus orientadores. A outra análise considerou a longevidade da relação por meio das publicações, ou seja, por quanto tempo o descendente direto do programa mantém um relacionamento com o seu orientador, mesmo que este fique algum período sem realizar publicações científicas, ponderando a diferença entre o último ano em que descendente direto e orientador têm publicações em coautoria e o ano de formação deste descendente.

Como pode ser observado na Figura 34 os descendentes diretos do

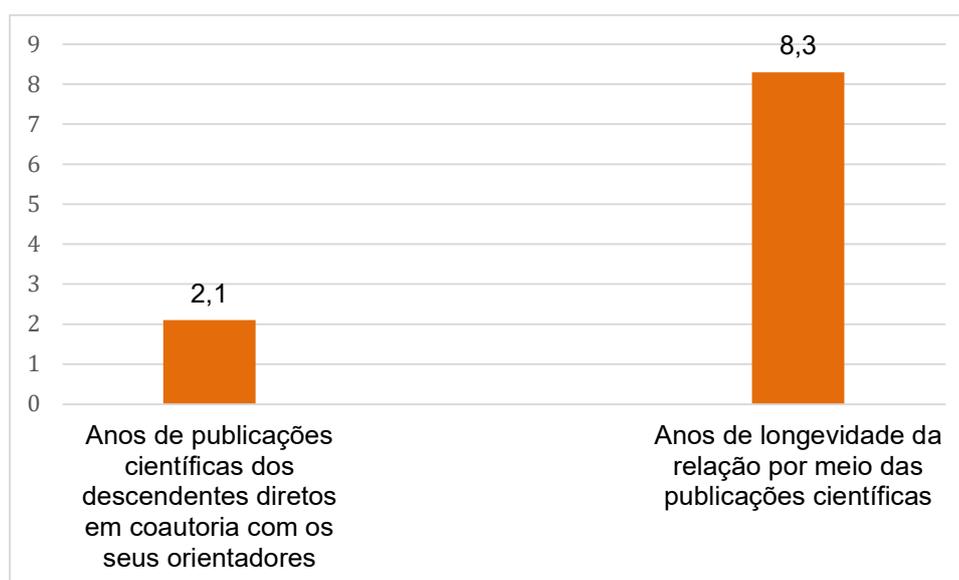
PPGCEM/UFSCar publicam em coautoria com seus orientadores, em média, por 2,1 anos após sua formação.

Porém, isso não quer dizer que o relacionamento termina a partir do momento que eles deixam de produzir juntos por um período. Muitos deles apresentam longos anos de parceria e colaboração por meio de suas publicações, como é o caso do relacionamento entre o descendente direto Luiz Antonio Pessan e seu orientador Elias Hage Junior, sendo o mais longo entre todos os descendentes diretos do programa, uma parceria por 23 anos. O descendente se formou no programa em 1987 e publicou em coautoria com seu orientador pela primeira vez em 1988 e a última vez em 2010.

A média da longevidade do relacionamento após a formação dos descendentes diretos, por meio de suas publicações científicas é, em média, de 8,3 anos.

Portanto, por meio deste indicador, mais uma vez foi possível constatar que GA em conjunto à análise bibliométrica propicia a alcançar informações raras referentes ao PPGCEM/UFSCar, além de certificar a consolidação e a importância do seu trabalho, por meio dos longos anos de parceria entre orientador e orientado nas publicações científicas.

Figura 34: Análise da relação por meio das publicações científicas



Fonte: Elaborado pela autora

5 CONSIDERAÇÕES FINAIS

A internet e o avanço tecnológico possibilitaram melhorias na disseminação e recuperação da informação, principalmente quando se trata da disseminação da informação científica, beneficiando praticamente todas as áreas e aproximando as pessoas do mundo científico, ou seja, o conhecimento que antes era privilégio de poucos, se tornou acessível. Mas, além do meio que possibilitasse esse acesso, também necessitou haver uma quebra de paradigma cultural dos cientistas quanto ao compartilhamento de resultados de suas pesquisas e que muitos ainda resistem.

Entender a dinâmica do conhecimento produzido pelos Programas de Pós-Graduação e os caminhos pelos quais esse conhecimento percorre, através de todos os seus descendentes, se torna uma tarefa muito mais complexa e praticamente impossível sem a utilização de ferramentas computacionais.

Muito se fala nas avaliações do programa realizada pelas instituições de fomento, e claro, isso deve ser uma preocupação constante de cada PPG para garantir a sua continuidade, mas por outro lado, isso nada existiria se não houvesse pessoas que buscam através desses programas, contribuir para a Ciência, bem como um avanço na vida profissional e pessoal.

Essas pessoas saem de um PPG com sementes nas mãos e assim como as sementes das plantas, elas devem cair em solo fértil para que a espécie continue. Diante disso, saber os caminhos que essas pessoas percorrem, como e onde elas estão, o que estão fazendo, se fecundaram novas mentes com seus conhecimentos e o quanto isto está contribuindo para a área e para a Ciência, é uma forma de saber se o que está sendo feito está gerando resultados; funciona como um *feedback* das ações dos Programas de Pós-Graduação.

E isso se confirma, por meio das mudanças nos instrumentos da avaliação dos Programas de Pós-Graduação, que passaram a dar maior ênfase à formação e avaliação de resultados do que nos processos.

A informação, como sempre foi, é algo valioso e apesar de estar disponível e de hoje haver diversos meios para buscá-la, nem sempre se encontra o caminho facilmente, porque é fato que a informação é garantia de competitividade. E a competitividade, de forma imutável, está no DNA do ser humano, e opera em todos os lugares, até mesmo onde não há a necessidade de existir. Para muitos, esquivar-

se dessa realidade é estar ciente da sua descontinuidade.

Porém, muito do que a sociedade se dispõe hoje, são resultados das ações movidas pela competitividade que, em algum lugar do tempo, transpôs o regalo da sociedade; e a internet é um desses exemplos. Apesar da sua origem não ter sido por uma causa nobre, desde o seu surgimento, ela incide diretamente na evolução do ser humano, principalmente no quesito informação. Com tantos dados à disposição, seria um despautério não os utilizar.

E muitos cientistas têm como objeto de pesquisa essa busca por informação por motivos variados. Mas para uma universidade o motivo é pela sua continuidade, de seus programas e a evolução da pesquisa, que garantem muitos benefícios para toda a sociedade, seja pelos desenvolvimentos de pesquisas e seus resultados, bem como desenvolvimento econômico regional e nacional, como geração de empregos, capitalização de empresas, formação de profissionais qualificados, entre outros. E neste trabalho pôde ser constatada a importância de achar meios de se obter informação para, literalmente, se manter vivo em um ambiente altamente competitivo, onde números são de extrema importância e os recursos são, incrivelmente, escassos.

Observou-se, então, como a informação pode ser obtida utilizando uma ferramenta que foi desenvolvida para outros fins. Uma ferramenta que é o resultado de outras pesquisas, tendo a oportunidade de explorá-la, adentrando nas ideias e pensamentos de grandes pesquisadores. Mas antes de conhecê-la, foi necessário buscar o conhecimento sobre o que se pretendia, sobre a área da Ciência da Informação, sobre uma universidade que um dia foi sonho e se transformou em realidade para muitas pessoas e um de seus departamentos, a qual existência, grandeza e importância, foram constatadas nesse estudo. Agora se sabe por que o PPGCEM/UFSCar causa tanto orgulho para todos que tiveram o privilégio de fazer parte da sua história, da história do DEMa e da UFSCar.

A presente pesquisa revelou também a necessidade do CNPq em aumentar os esforços para melhorar a qualidade das informações preenchidas no Currículo Lattes, adotando medidas como por exemplo, envio de notificações sobre as atualizações, melhoria dos textos autoexplicativos, diminuição de textos digitados e design mais intuitivo.

Por outro lado, muitos problemas podem ser evitados se os acadêmicos, durante o período que estiverem no programa, forem orientados quanto a importância

do preenchimento correto de seus Currículo Lattes, bem como sobre a importância de mantê-los atualizados, já que o uso da Plataforma Lattes é uma tendência, principalmente no meio acadêmico e que vem, cada vez mais, ganhando a atenção para estudos bibliométricos.

Portanto, a associação da GA com a Bibliometria demonstrou-se de alto potencial a ser utilizada por Programa de Pós-Graduação, permitindo conhecer toda a sua linhagem por meio das informações sobre seus descendentes, de uma maneira rápida, confiável, que podem complementar as análises bibliométricas e apoiar na avaliação do programa ou no acompanhamento de suas ações, com vistas ao fortalecimento e desenvolvimento do programa, bem como avaliar a atuação de cada professor membro, propiciando estudos mais aprofundados das gerações e de seus resultados, de acordo com o objetivo de cada programa.

Para trabalhos futuros sugere-se realizar a análise dos dados de outros PPG's utilizando a mesma metodologia e assim como a Plataforma Acácia, desenvolver uma ferramenta, que automatize o processo para consultar a descendência de um Programa de Pós-Graduação ou de uma área específica, com o objetivo de tornar esse processo menos complexo. Também sugere-se realizar análise de orientação considerando os coo

REFERÊNCIAS

AMOROS, R. et al. **Find Out How Much Your Country Spends on Research & Development**. Disponível em: <<https://howmuch.net/articles/research-development-spending-by-country>>. Acesso em: 5 set. 2019.

ANDRAOS, J. Scientific genealogies of physical and mechanistic organic chemists. **Canadian Journal of Chemistry**, v. 83, n. 9, p. 1400–1414, 2006.

ARAÚJO, C. A. Bibliometria: evolução histórica e questões atuais. **Em Questão**, Porto Alegre. Vol.12, n. 1, p. 11–32, 2006a.

ARAÚJO, C. A. À. A Ciência como forma de conhecimento. **Ciências & Cognição**, Belo Horizonte. Vol. 8:127-142, 2006b.

ARBESMAN, S. **Math Masters Trace Their Intellectual Lineage**. Disponível em: <<https://www.wired.com/2011/05/st-mathancestry/amp>>. Acesso em: 5 set. 2019.

BACELAR, J. Apontamentos sobre a história e desenvolvimento da impressão. **Universidade da Beira Interior**, Covilhã - Portugal p. 1–6, 1999. Disponível em: <http://www.bocc.ubi.pt/pag/bacelar_apontamentos.pdf>. Acesso em: 20 set. 2019.

BARRETO, A. DE A. Uma história da Ciência da informação. In: **Para Entender a Ciência da Informação**. Salvador: EDUPBA, 2007. p. 242 p.

BARROS, R. F. Comunicação científica: reflexões preliminares para o GT “relevância dos livros acadêmicos na comunicação da pesquisa”. **SciELO**, n. 2003, 2018.

BASSOLI, Marcela. **Avaliação do currículo lattes como fonte de informação para construção de indicadores: o caso da UFSCar**. 2017. 139 f. Dissertação (Mestrado em Ciência, Tecnologia e Sociedade) –Universidade Federal de São Carlos (UFSCar), São Carlos, 2017. Disponível em: Acesso em: 08 set. 2019.

BIBLIOTECA DA FEUP. **Guias Temáticos de Apoio: PUBLICAÇÃO CIENTÍFICA: Contextos e evolução**. Disponível em: <<https://feup.libguides.com/publicacao-cientifica/contextos>>. Acesso em: 17 set. 2019.

CAPES. **Relatório Técnico DAV - Egressos da Pós-Graduação: Áreas**

Estratégicas. p. 38, 2017.

CAPES. **História e missão**. Disponível em: <<https://capes.gov.br/historia-e-missao>>. Acesso em: 24 abr. 2019.

CAPES. Ficha de Avaliação - Grupo de trabalho: Proposta de revisão da Ficha utilizada para a Avaliação dos Programas de Pós-Graduação que é conduzida pela CAPES. p. 22, 2018.

CAPES. **Qualis Periódicos e classificação de produção intelectual**. Disponível em: <<http://www.capes.gov.br/avaliacao/instrumentos-de-apoio/qualis-periodicos-e-classificacao-de-producao-intelectual>>. Acesso em: 17 mar. 2020.

CAPES. **Avaliação da CAPES aponta crescimento da pós-graduação brasileira**. Disponível em: <<https://www.capes.gov.br/sala-de-imprensa/noticias/8558-avaliacao-da-capes-aponta-crescimento-da-pos-graduacao-brasileira>>. Acesso em: 24 abr. 2019.

CAPES. **Programa De Doutorado Interinstitucional - Dinter - Projetos em Andamento**. Disponível em: <http://www.capes.gov.br/images/novo_portal/documentos/Dinter/16122019_Projetos_DINTER_ativos_-_2019.pdf>. Acesso em: 17 fev. 2020.

CAPES. **Autoavaliação de Programas de Pós-Graduação**. Disponível em: <https://www.capes.gov.br/images/novo_portal/documentos/DAV/avaliacao/10062019_Autoavaliacao-de-Programas-de-Pós-Graduacao.pdf>. Acesso em: 27 jan. 2020.

CAPES. **Mudanças na ficha de avaliação valorizam qualidade dos programas**, 2019. Disponível em: <<https://www.capes.gov.br/36-noticias/9370-mudancas-na-ficha-de-avaliacao-valorizam-qualidade-dos-programas>>. Acesso em: 20 mar. 2020.

CHIBENI, S. S. **O que é Ciência?** Departamento de Filosofia da Unicamp – IFCH-Campinas, 2019. Disponível em: <<https://www.unicamp.br/~chibeni/textosdidaticos/ciencia.pdf>>. Acesso em: 03 out. 2019.

CHIBENI, S. S. **Introdução à filosofia da Ciência Aula 1. A s origens da Ciência moderna**. p. 1–5, 2019. Departamento de Filosofia da Unicamp – IFCH-Campinas, 2019. Disponível em:

< <https://www.unicamp.br/~chibeni/textosdidaticos/ciencia.pdf> >. Acesso em: 03 out. 2019.

CNPQ. **Árvore do Conhecimento - Áreas do conhecimento - Engenharias**. Disponível em: <<http://lattes.cnpq.br/web/dgp/engenharias>>. Acesso em: 7 mar. 2020.

CNPQ. **Histórico - Plataforma Lattes - CNPq**. Disponível em: <<http://lattes.cnpq.br/web/plataforma-lattes/historico/>>. Acesso em: 2 abr. 2019a.

CNPQ. **Dados e estatísticas - Plataforma Lattes - CNPq**. Disponível em: <<http://lattes.cnpq.br/web/plataforma-lattes/dados-e-estatisticas/>>. Acesso em: 2 abr. 2019b.

CNPQ. **O que é a Plataforma Lattes**. Disponível em: <<http://lattes.cnpq.br/web/plataforma-lattes/o-que-e/>>. Acesso em: 2 abr. 2019c.

CORRÊA, T. S. et al. O fim do scriptLattes? Uma análise de suas funcionalidades, alternativas para o presente e perspectivas para o futuro. **Revista do EDICC - ISSN 2317-3815**, v. 3, n. 3, p. 138–148, 2017.

CORTELLI, S. C. **A importância da pesquisa para a sociedade**. Disponível em: <<http://web.unitau.br/noticias/2018/08/01/a-importancia-da-pesquisa-para-a-sociedade/>>. Acesso em: 4 set. 2019.

CRUZ, C. H. B. Vannevar Bush uma apresentação. **Revista Latino-americana de Psicopatologia Fundamental**, p. 11–13, 2011.

DAMACENO, R. J. P.; ROSSI, L.; MENA-CHALCO, J. P. A genealogia acadêmica brasileira: evidências de relações de conselheiro-assessor por meio de análise quantitativa. **Scientometrics (2019) 119: 303**, 2019.

DAMACENO, R. J. P. Alterações no software AGLattes [mensagem pessoal de voz via SKYPE]. Mensagem recebida por <rafael.jpd > em 22 jan. 2020.

DAVID, S. V; HAYDEN, B. Y. Journals Plos. **Neurotree: A Collaborative, Graphical Database of the Academic Genealogy of Neuroscience**, 2012.

DENNIS, M. A. **Encyclopedia Britannica**. Disponível em:

<<https://www.britannica.com/biography/Vannevar-Bush>>. Acesso em: 19 set. 2019.

DESJARDINS, J. Innovators wanted: these countries spend the most on R&D. **World Economic Forum**, 2018.

UNIÃO, D. O. DA. No 158, quinta-feira, 16 de agosto de 2018 - **Fundação Coordenação de Aperfeiçoamento de Pessoal De Nível Superior**. p. 2018, 2018.

DROESCHER, F. D.; SILVA, E. L. O pesquisador e a produção científica. **Perspectivas em Ciencia da Informacao**, v. 19, n. 1, p. 170–189, 2014.

DUDZIAK, E. A. Quem financia a pesquisa brasileira? Um estudo Incites sobre o Brasil e a USP. **SIBiUSP**, 2018. Disponível em: <<https://www.aguia.usp.br/noticias/quem-financia-a-pesquisa-brasileira-um-estudo-incites-sobre-o-brasil-e-a-usp/>> Acesso em: 8 set. 2019

ECHER, I. C. A REVISÃO DE LITERATURA NA CONSTRUÇÃO DO TRABALHO CIENTÍFICO Literature review in a scientific work. *Revista gaúcha de enfermagem*, v. 22, p. 5–20, 2001.

FAPESP - Fundação de Amparo à Pesquisa do Estado de S. Paulo. Disponível em: <http://www.fapesp.br/publicacoes/indct/apres/apres.htm/>>. Acesso em: 5 jun. 2019.

FARIA, L. I. L. DE et al. Análise da produção científica a partir de publicações em periódicos especializados. In: **Indicadores de Ciência, Tecnologia e Inovação em São Paulo 2010**. São Paulo: [s.n.]. p. 4.1-4.71.

FARIA, Leandro Innocentini Lopes de. Prospecção tecnológica em materiais: aumento da eficiência do tratamento bibliométrico. Aplicação na análise de tratamentos de superfície resistentes ao desgaste. 2001. 213 f. São Carlos. Disponível em <https://repositorio.ufscar.br/handle/ufscar/644>. Acesso em: 8 set. 2019.

FEA/USP. **Genealogia Acadêmica - Economia - FEA/USP**. Disponível em: <<http://professor.ufabc.edu.br/~jesus.mena/genealogia-academica/economia-fea-usp/network/graph.html>>. Acesso em: 20 dez. 2019.

FERNANDES, C. **Invenção da Imprensa**. Disponível em:

<<https://www.historiadomundo.com.br/idade-moderna/invencao-imprensa.htm>>. Acesso em: 9 set. 2019.

FERREIRA, S. L. Transmissão de sobrenomes entre luso-brasileiros: uma questão de classe. **Boletim de História Demográfica**, p. 302, 2005.

FINEP. **Os segredos do país mais inovador do mundo**. Disponível em: <<http://www.finep.gov.br/noticias/todas-noticias/5989-os-segredos-do-pais-mais-inovador-do-mundo>>. Acesso em: 5 set. 2019.

FEOFILOFF, P.; KOHAYAKAWA, Y.; WAKABAYASHI, Y. Uma Introdução Sucinta à Teoria dos Grafos. **IME USP**, p. 61, 2011.

GEVENOIS, P. A.; DURIEUX, V. Bibliometric Indicators: Quality Measurements of Scientific Publication. **Radiology**, v. 255, n. 2, p. 342–351, 2010.

GLOBAL RESEARCH COUNCIL. 2019 GRC Statement of Principles Addressing Expectations of Societal and Economic Impact. n. 6, p. 6–7, 2019.

GRAY, J. J. **Carl Friedrich Gauss | Biography, Discoveries, & Facts | Britannica.com**. Disponível em: <<https://www.britannica.com/biography/Carl-Friedrich-Gauss>>. Acesso em: 2 abr. 2019.

GROSSMANN, L. O. **PONTES: CORTE DE 42% NO ORÇAMENTO DEIXA MCTIC 'COM A CORDA NO PESCOÇO'**. Disponível em: <<http://www.abc.org.br/2019/04/11/pontes-corte-de-42-no-orcamento-deixa-mctic-com-a-corda-no-pescoco/>>. Acesso em: 5 set. 2019.

HADDAD, E. A.; MENA-CHALCO, J. P.; SIDONE, O. Produção científica e redes de colaboração dos docentes vinculados aos programas de pós-graduação em Economia no Brasil. **Estudos Econômicos (São Paulo)**, v. 47, n. 4, p. 617–679, 2018.

ICB. Instituto de Ciências Biológicas: **Introdução à Análise de Genealogias**. Disponível em: <<https://www2.icb.ufmg.br/grad/genetica/heredogramas.htm>>. Acesso em: 2 set. 2019.

INGESC. Instituto de Genealogia de Santa Catarina: **Estude o passado, viva o presente e entenda o futuro**. Disponível em: <<http://www.ingesc.org.br/>>. Acesso em: 2 set. 2019.

LAFER, C. Indicadores de Ciência, Tecnologia e Inovação em São Paulo 2010. In: **Indicadores de Ciência, Tecnologia e Inovação em São Paulo 2010**. Volume I ed. São Paulo: FAPESP, 2011. p. XIX.

LATTES, P. **EXTRAÇÃO DE DADOS DA PLATAFORMA LATTES-PoSIC**. n. Anexo I, p. 1–6, 2016. Disponível em: <<http://memoria.cnpq.br/documents/313759/dad0813b-6eb1-45b5-9084-fdbd3f6c78cc>>. Acesso em: 25 mar. 2020.

LATTES, P. **Currículo Lattes de Jose Angelo Rodrigues Gregolin**. Disponível em: <<http://buscatextual.cnpq.br/buscatextual/visualizacv.do?id=K4783420D4>>. Acesso em: 3 dez. 2019.

LATTES, P. **CNPq - Plataforma Lattes - Currículo Lattes de Naercio Aquino Menezes Filho**. Disponível em: <<http://lattes.cnpq.br/7403858236965011>>. Acesso em: 5 jan. 2020.

LE COADIC, Y.-F. **A Ciência da informação / Yves - François Le Coadic: tradução de Maria Yêda F. S. de Filgueiras Gomes**. Brasília, DF: Briquet de Lemos/Livros, 1996.

LIMA, W. T. A. **Avaliação de egressos da pós-graduação estrito senso: Política de monitoramento de egressos internacionais**. Revista do Colegio Brasileiro de Cirurgioes, v. 42, n. Suplemento 1, p. 26–27, 2015.

LUCHESI, C. et al. **Fazer universidade - uma proposta metodológica**. 17° ed. São Paulo: Cortez, 2012. Cap.4, pg. 178-188.

MACIEL, R. S. **A Plataforma Lattes como recurso estratégico para a gestão dos programas de Pós - Graduação: Uma análise baseada na produção de artigos científicos**. Dissertação (Mestrado em Ciência da Informação) – Curso de Pós-Graduação em Ciência da Informação, Universidade Federal de São Carlos, São Carlos, 2018.

MARQUES, F. **Ciclo interrompido**. Disponível em: <<https://revistapesquisa.fapesp.br/2019/01/10/ciclo-interrompido/>>. Acesso em: 5 set. 2019.

MATIAS, Mesailde Souza de Oliveira. **Base referencial para o povoamento de repositórios institucionais: coleta automatizada de metadados da Plataforma Lattes**. 2015. 94 f. Dissertação (Mestrado em Gestão de Organizações e Sistemas

Públicos) –Universidade Federal de São Carlos (UFSCar), São Carlos, 2015. Disponível em: <<https://repositorio.ufscar.br/handle/ufscar/6932>>. Acesso em 01 dez. 2019.

MEDEIROS, J. B. **Redação Científica: a prática de fichamentos, resumos, resenhas**. 11. ed. São Paulo: EDITORA ATLAS, 2013.

MENA-CHALCO, J. P. Genealogia acadêmica: uma ferramenta para investigar a origem, evolução e disseminação de áreas do conhecimento. **5º Encontro Brasileiro de Bibliometria e Cientometria**, n. Jul., 2016.

MENA-CHALCO, J. P. et al. **Plataforma Acácia - Genealogia Acadêmica do Brasil**. Disponível em: <<http://plataforma-acacia.org/>>. Acesso em: 2 abr. 2019.

MOREIRA, T. H. J.; DIAS, T. M. R.; MOITA, G. F. Uma estratégia baseada em árvores genealógicas científicas para visualização da relação orientador-orientado In: **Encontro Brasileiro de Bibliometria e Cientometria**, 5., 2016, São Paulo. Anais... São Paulo: USP, 2016. p. A91
1, 2, 3 Centro Federal de Educação Tecnológica de Minas Gerais Powered

MULCAHY, C. The Mathematics Genealogy Project Comes of Age at 21. **Notices of the American Mathematical Society**, v. 64, n. 05, p. 466–470, 2017.

MENA-CHALCO, J. P.; MALDONADO, J. C.; MURTA, L. G. P. **Genealogia acadêmica da área de Engenharia de Software no Brasil**. Disponível em: <<http://professor.ufabc.edu.br/~jesus.mena/genealogia-academica-eng-software/>>. Acesso em: 19 fev. 2020.

NHACUONGUE, J. A.; FERNEDA, E. O campo da Ciência da informação: Contribuições, desafios e perspectivas. **Perspectivas em Ciência da Informacao**, v. 20, n. 2, p. 3–18, 2015.

NIT MATERIAIS. **Indicadores C&T**. Disponível em: <<http://www.nit.ufscar.br/servicos/indicadores-cet>>. Acesso em: 31 mar. 2020.

NORONHA, D. P. et al. Egressos dos Programas de Pós-Graduação em Ciência da Informação: por onde andam os doutores? **Perspectivas em Ciência da Informação**, V.14, n.2, p. 94 – 107, 2009. Disponível em: <http://www.scielo.br/pdf/pci/v14n2/v14n2a07.pdf>. Acesso em: 3 dez. 2019.

OKUBO, Y. General Distribution OCDE/GD(97)41 STI WORKING PAPERS 1997/1 Bibliometric Indicators and Analysis of Research Systems: Methods And Examples. n. 97, 1997.

PGM. **Projeto de Genealogia Matemática**. Disponível em: <<https://www.genealogy.math.ndsu.nodak.edu/index.php>>. Acesso em: 1 abr. 2019.

PIZZI, M. DE L. B.; SANTOS, M. A. G. DOS. Alguns indicadores sobre os egressos do Programa de Pós-Graduação em Educação da UFSCar nos anos de 2008 a 2016. **Revista Eletrônica de Educação**, v. 11, n. 2, p. 407–428, 2017.

PLATAFORMA ACÁCIA - **Genealogia Acadêmica do Brasil**. Disponível em: <<http://plataforma-acacia.org/>>. Acesso em: 2 abr. 2019.

PLATAFORMA SUCUPIRA. **Qualis Periódicos**. Disponível em <https://sucupira.capes.gov.br/sucupira/public/index.jsf>. Acesso em: 1 fev. 2020.

PPGCEM/UFSCAR. **Programa de Pós-Graduação em Ciência e Engenharia de Materiais**. Disponível em: <<http://www.ppgcem.ufscar.br/>>. Acesso em: 5 set. 2019.

Revista Edição Comemorativa DEMa 40 Anos. . São Carlos: [s.n.]. Disponível em: <http://www.revistakappa.com.br/edicoes/DEMa/files/revista_kappa.pdf>. Acesso em: 5 set. 2019.

RODRIGUES, M.; MATTOS, A. DE M. Ignaz Semmelweis e a febre puerperal: algumas razões para a não aceitação de sua hipótese. **Filosofia e História da Biologia**, v. 10, n. 1, p. 85–98, 2015.

RIBEIRO, C. J. S. Big Data: os novos desafios para o profissional da informação. **Big Data: os novos desafios para o profissional da informação**, v. 1, n. 1, p. 96–105, 2014.

RIBEIRO, R. J. Os critérios da avaliação. Brasília, DF: [s.n.]. 2007. Disponível em: <https://www.capes.gov.br/images/stories/download/artigos/Artigo_10_08_07.pdf>. Acesso em: 24 abr. 2019.

ROSSI, L.; DAMACENO, R. J. P.; MENA-CHALCO, J. P. **Genealogia acadêmica: Um novo olhar sobre impacto acadêmico de pesquisadores**. Disponível em:

<https://www.researchgate.net/publication/326960663_Genealogia_academica_Um_novo_olhar_sobre_impacto_academico_de_pesquisadores>. Acesso em: 28 set. 2019.

ROSSI, L.; MENA-CHALCO, J. P. Caracterização de árvores de genealogia acadêmica por meio de métricas em grafos. **III Brazilian Workshop on Social Network Analysis and Mining (BraSNAM 2014)**, p. 12, 2014.

SANTOS, P. Paul Otlet: Um pioneiro da organização das redes mundiais de tratamento e difusão da informação registrada. **Ciencia da Informacao**, v. 36, n. 2, p. 54–63, 2007.

SARACEVIC, T. Ciência da informação: origem, evolução e relações¹. **BRAPCI, Perspec. Ci. Inf.**, Belo Horizonte, v. 1, n. 1, p. 41-62, jan./jun. 1996

SAVIGNANO, V. Um aniversário da Engenharia de Materiais no Brasil: 40 anos do DEMa da UFSCar. **Brazilian Materials Reserch Society**, 2012.

SCHAEFFER, E. Noções de genealogia científica. Questões Pedagógicas. **Revista de História – USP**, São Paulo, 1960. V. 21 N. 44 /. Disponível em <<<http://www.revistas.usp.br/revhistoria/article/view/120155>>>. Acesso em 30 out. 2019.

SILVA, J. L. C.; FREIRE, G. H. DE A. Um olhar sobre a origem da Ciência da informação: indícios embrionários para sua caracterização identitária. **Encontros Bibli: revista eletrônica de biblioteconomia e Ciência da informação**, v. 17, n. 33, p. 1–29, 2012.

SILVA, J. L.; TUNDISI, J. G. Financiamento à pesquisa. In: **Projeto de Ciência para o Brasil**. Rio de Janeiro: Academia Brasileira de Ciências, 2018. p. 30–32.
SILVEIRA, D. T.; CÓRDOVA, F. P. **Métodos de Pesquisa - Unidade 2 – A Pesquisa Científica**. 1. ed. Porto Alegre: Universidade Federal do Rio Grande do Sul, 2009.

SOUZA, L. G. DE. Relação entre Ciência & pesquisa, en Contribuciones a las Ciencias Sociales, **eumed**, octubre 2011. Disponível em <<http://www.eumed.net/rev/cccss/14/lgs.html>> Acesso em: 13 set. 2019.

SPINAK, E. Indicadores cientímetricos. **IBICIT**, Ci. Inf., Brasília, v. 27, n. 2, p. 141-148, maio/ago. 1998. Disponível em:
<http://revista.ibict.br/ciinf/article/view/795> Acesso em: 18 set. 2019.

SPINAK, E.; PACKER, A. L. **350 anos de publicação científica: desde o “Journal des Sçavans” e “Philosophical Transactions” até o SciELO.** Disponível em: <<https://blog.scielo.org/blog/2015/03/05/350-anos-de-publicacao-cientifica-desde-o-journal-des-sçavans-e-philosophical-transactions-ate-o-scielo/#.XYE2ei5Kjcs>>. Acesso em: 13 set. 2019.

SUCUPIRA, Plataforma. **Qualis Periódicos.** Disponível em: <<https://sucupira.capes.gov.br/sucupira/public/consultas/coleta/programa/quantitativos/quantitativos.jsf?areaAvaliacao=12&areaConhecimento=30300002>>. Acesso em: 1 fev. 2020.

SUCUPIRA, P. **Cursos Avaliados e Reconhecidos.** Disponível em: <<https://sucupira.capes.gov.br/sucupira/public/consultas/coleta/programa/quantitativos/quantitativoBuscaAvancada.jsf>>. Acesso em: 10 mar. 2020.

SUGIMOTO, C. R. et al. Academic Genealogy as an Indicator of Interdisciplinarity: An Examination of Dissertation Networks in Library and Information Science. **Journal of the American Society for Information Science**, 2011.

SUGIMOTO, C. R. Beyond bibliometrics. In: **Political Geography**. London: Massachusetts Institute of Technology, 2014. v. 68p. 365–382.

TELLES, P. C. DA S. História da engenharia no Brasil. **História da Engenharia no Brasil**, v. 2, p. 3–6, 1984.

TORRES, M. DE F. N. DE M.; [HTTP://LATTES.CNPQ.BR/4586218146401718](http://lattes.cnpq.br/4586218146401718). Contribuição à identificação da competência científica de um Instituto Público de Pesquisa por meio de indicadores bibliométricos. p. 1–221, 2018.

TOZONI-REIS, M. F. DE C. A Pesquisa e a Produção de Conhecimentos. **Universidade Estadual Paulista. Prograd. Caderno de Formação: formação de professores didática de conteúdo.**, p. 1–38, 2010.

TSALLIS, C. **A importância da pesquisa para o desenvolvimento do país.** Disponível em: <<https://portal.fei.edu.br/noticia/84/a-importancia-da-pesquisa-para-o-desenvolvimento-do-pais>>. Acesso em: 6 set. 2019.

UFSCAR. **Universidade Federal de São Carlos.** Disponível em: <<https://www2.ufscar.br/>>. Acesso em: 13 set. 2019.

UNESCO, C. N. DA. **Ciência para a sociedade**. Disponível em: <<https://www.unescoportugal.mne.pt/pt/temas/ciencia-para-um-futuro-sustentavel/ciencia-para-a-sociedade>>.

VELHO, L. **Cuidado com os rankings científicos**. Disponível em: <<http://www.prometeu.com.br/bb-lea.htm>>.

VANTAGEPOINT. **The VantagePoint**. Disponível em:

<<https://www.thevantagepoint.com/>>. Acesso em: 11 dez. 2020.

VIANA, L. C. S. et al. Genealogia acadêmica e sua relação com a trajetória do conhecimento na Escola Nacional de Saúde Pública Sergio Arouca, Fundação Oswaldo Cruz. **Transinformação**, v. 31, 2019.

WERSIG, G.; NEVELING, U. The phenomena of interest to Information Science. **Journal of Chemical Information and Modeling**, v. 9, p. 127–140, 1975.

YAMAMOTO, E. **USP é a universidade brasileira mais internacional**. Disponível em: <<https://jornal.usp.br/institucional/usp-e-a-universidade-mais-internacional-da-america-latina/>>. Acesso em: 30 jan. 2020.

APÊNDICE A – EXECUÇÃO DO SOFTWARE AGLATTES

A programação do AGLattes foi desenvolvida na linguagem *Python*, para ser executado no sistema operacional Linux e o computador utilizado para a sua execução e realização desta pesquisa foi Dell com 8GB de memória RAM e Processador Intel Core i5-2400, com sistema operacional Windows 10.

Desta forma, para executar o *software* AGLattes no computador citado, foi necessário instalar uma máquina virtual, a VirtualBox da Oracle, disponível no site: <https://www.virtualbox.org/>. Posteriormente, foi instalado o sistema operacional Linux, optando pela versão Ubuntu 18.04.2 LTS, disponível no site: <https://ubuntu.com/download/desktop>.

Um ponto importante para a execução do programa é a verificação e garantia de uma boa conexão com a internet.

O *software* AGLattes foi disponibilizado em um arquivo compacto, contendo 17(dezessete) arquivos, sendo dez blocos de instruções (*scripts*), que podem ser executados individualmente ou simultaneamente. Os desenvolvedores disponibilizaram as instruções para que o *software* seja executado no arquivo *README.md*, porém algumas não foram descritas, sendo possível a execução após análise do código. Estas informações que não estão disponíveis e as mais relevantes foram incluídas neste trabalho.

Dentre os arquivos, dois são de extrema importância para a execução do programa: *sementes* e *cfg_sementes.py*, que deverão ser configurados antes da execução do programa. Devido a maior agilidade, esses arquivos foram configurados no Windows 10, para posteriormente, serem utilizados na execução do programa no Linux.

No arquivo *sementes* é onde estão relacionados todos os *idlattes* dos docentes que se desejava buscar os dados. Trata-se de uma numeração única e que consta no currículo de toda pessoa que se registra na Plataforma Lattes, como pode ser observado na Figura 1.

Figura 1: Currículo de um dos docentes do PPGCEM/UFSCar

Jose Angelo Rodrigues Gregolin

Endereço para acessar este CV: <http://lattes.cnpq.br/7842651773105125>

ID Lattes: **7842651773105125**

Última atualização do currículo em 17/04/2017

Professor Sênior do Departamento de Engenharia de Materiais, pesquisador do Núcleo de Informação Tecnológica em Materiais (NIT/Materiais) da Universidade Federal de São Carlos (UFSCar), com atuação em prospecção tecnológica, inteligência competitiva e gestão tecnológica em produtos, materiais, processos e energia. Vinculado à UFSCar desde 1983, tendo sido Coordenador do NIT/Materiais e do Programa de Pós-Graduação em Ciência, Tecnologia e Sociedade, Vice-Coordenador do Programa de Pós-Graduação em Ciência e Engenharia de Materiais, Chefe do Departamento de Engenharia de Materiais e Coordenador do Curso de Graduação em Engenharia de Materiais. Doutor em Engenharia Mecânica pela Unicamp (1990) e Engenheiro de Materiais pela UFSCar (1976). (Texto informado pelo autor)

Fonte: (LATTES, 2019)

Em seguida, os *idlattes* foram salvos em um arquivo *.txt*, quais devem estar dispostos um por linha, sem a necessidade de qualquer caractere para separá-los, como mostra a Figura 2.

Figura 2: Disposição das sementes no arquivo.txt

```

7511722134943618
6360362248108490
9754332336954137
4499231813051400
3527528295399928
3259846931278607
8141224513975606
0896060959622431
8699821828310072
9231627080617037
5443002033733395
2176215981291453
6576844591900987
4100119317525940
1055167132036400
1025598529469393
1240692926327687
9848311210578810
7924187202036614
2134503900832349
2410385643716370
7737833809468759
7842651773105125
6653223814715498
1807399214239200
7592090152400440
4000441989868775
9470914895315270
8882038118634925
5373845785326215
7287108960864123
0333912025909549
8246109223926480
8276650236213537
5839043594908917
7904227795652627
0172837599844991

```

Fonte: Elaborado pela autora

Após foi configurado o arquivo *cfg_sementes.py*. A configuração deste arquivo determina o comportamento do programa.

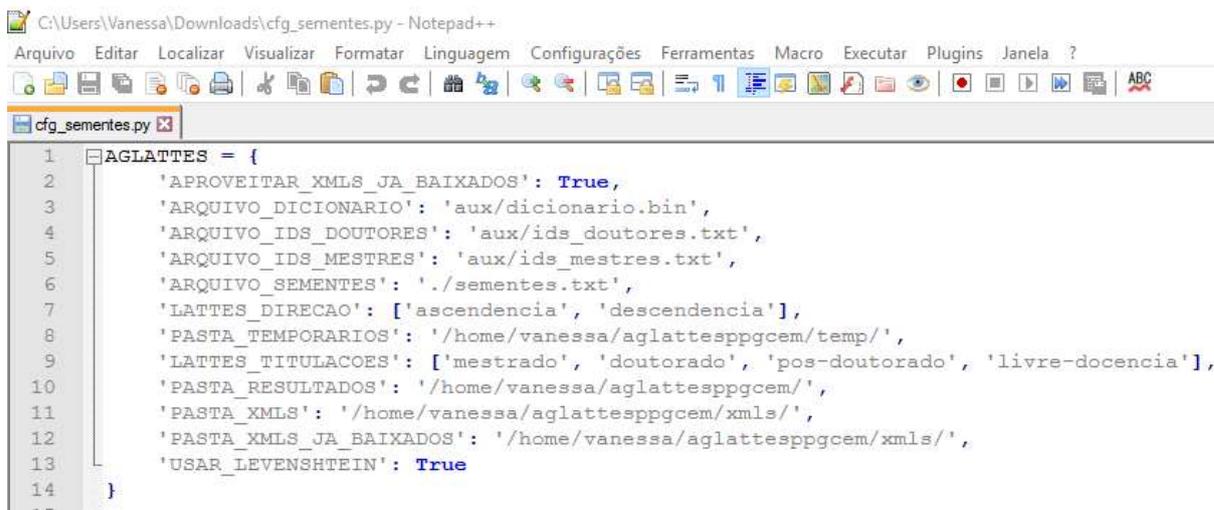
No destino dos resultados, deverão ser criadas duas pastas, uma para armazenar os arquivos no formato XML e outra para armazenar os arquivos temporários. De acordo com que está programado no código, foram criadas a pasta *temp* e a pasta *xmls*. Os autores não mencionam sobre a necessidade da criação dessas pastas no arquivo *README.md* do programa.

De acordo com Rossi, Damaceno, Mena-Chalco (2018), abaixo estão os arquivos que podem ser configurados de acordo com a necessidade:

- **APROVEITAR_XMLS_JA_BAIXADOS**: Deverá ser configurado com *True* ou *False*, caso pretenda aproveitar currículos em formato XML já baixados.
- **ARQUIVO_SEMENTES**: Nome do arquivo onde constam os códigos identificadores lattes dos pesquisadores que deseja construir o grafo.
- **LATTES_DIRECAO**: Define o sentido de construção do grafo. Diante da necessidade de construir o grafo apenas no sentido descendente é preciso manter apenas o termo 'descendência' na lista de sentidos.
- **LATTES_TITULACOES**: Define os níveis de titulação a serem considerados, tais como `mestrado`, `doutorado`, `pós-doutorado` e `livre-docência`.
- **PASTA_RESULTADOS**: Define a pasta onde os dados serão salvos após execução do programa.
- **PASTA_XMLS**: Define a pasta onde os currículos em formato XML serão baixados.
- **PASTA_XMLS_JA_BAIXADOS**: Define a pasta onde já constam currículos em formato XML já baixados.

Todos os arquivos do AGLattes podem ser visualizados e editados utilizando o bloco de notas do Windows ou qualquer outro programa de edição, como o *Notepad ++*, conforme mostra a Figura 3.

Figura 3: Configuração do arquivo `cfg_sementes.py` no Notepad ++



```

1  AGLATTES = {
2      'APROVEITAR_XMLS_JA_BAIXADOS': True,
3      'ARQUIVO_DICIONARIO': 'aux/dicionario.bin',
4      'ARQUIVO_IDS_DOUTORES': 'aux/ids_doutores.txt',
5      'ARQUIVO_IDS_MESTRES': 'aux/ids_mestres.txt',
6      'ARQUIVO_SEMENTES': './sementes.txt',
7      'LATTES_DIRECAO': ['ascendencia', 'descendencia'],
8      'PASTA_TEMPORARIOS': '/home/vanessa/aglattesppgcm/temp/',
9      'LATTES_TITULACOES': ['mestrado', 'doutorado', 'pos-doutorado', 'livre-docencia'],
10     'PASTA_RESULTADOS': '/home/vanessa/aglattesppgcm/',
11     'PASTA_XMLS': '/home/vanessa/aglattesppgcm/xmls/',
12     'PASTA_XMLS_JA_BAIXADOS': '/home/vanessa/aglattesppgcm/xmls/',
13     'USAR_LEVENSHTEIN': True
14 }
15

```

Fonte: Elaborado pela autora

Após configuração, o *software* foi executado no Linux. Este procedimento deve ser realizado no Terminal de comandos. Optou-se por executar todos os comandos simultaneamente, através do arquivo `main.sh`.

Todo o processo de extração de dados da Plataforma Lattes através do *software* AGLattes, leva cerca de onze minutos para ser completado e esse tempo pode variar de acordo com o número de sementes e com a velocidade da internet. Ao terminar a execução, ainda no Terminal de Comandos, o *software* informa a geração do grafo final, encerrando o processo.

Como resultado, o *software* fornece dois arquivos principais e muitos outros reunidos em duas pastas como pode ser observado na Figura 4.

Figura 4: Resultados do *software* AGLattes

Nome	Data de modificaç...	Tipo	Tamanho
temp	22/01/2020 20:11	Pasta de arquivos	
xmls	22/01/2020 11:50	Pasta de arquivos	
grafo.gdf	22/01/2020 11:51	Arquivo GDF	1.715 KB
metricas	22/01/2020 11:51	Arquivo de Valore...	126 KB

Fonte: Elaborado pela autora

De acordo com Rossi, Damaceno, Mena-Chalco (2018) um desses arquivos trata-se de um grafo (em formato GDF) de genealogia acadêmica em que cada vértice representa um pesquisador e cada aresta (direcionada) uma relação de orientação entre dois pesquisadores. O outro arquivo é em formato CSV (`metricas.csv`). O primeiro representa um grafo composto por vértices e arestas. O segundo contém as métricas genealógicas de cada vértice. É através do grafo que a árvore genealógica é construída.

De acordo com os autores, os atributos de ambos os arquivos (vértices e arestas do GDF e métricas do CSV) estão relacionados a seguir:

- Atributos dos pesquisadores: Ano da primeira livre docência, Ano do primeiro pós-doutorado realizado Bairro da instituição profissional atual, CEP da instituição profissional atual, Cidade da instituição profissional atual, Cidade de nascimento, Código identificador do pesquisador (idlattes), Data de última atualização do currículo, Se é ascendente de semente, Se é descendente de semente, Se é livre docente, Se se trata de uma das sementes pesquisadas, Logradouro da instituição profissional atual, Maior titulação, Menor ano da maior titulação, Nacionalidade, Nome completo, Nome da instituição profissional atual, Órgão na instituição profissional atual, País da instituição profissional atual, País de nacionalidade, País de nascimento, Possui pós-doutorado, Primeira área de atuação, Primeira especialidade de atuação, Primeira grande área de atuação, Primeira subárea de atuação, Sigla do país de nacionalidade, Unidade Federativa da instituição profissional atual e Unidade Federativa de nascimento.
- Atributos das relações de orientação: Ano de início da titulação, Ano de obtenção da titulação, Código do aluno orientado, Código do professor orientador, Curso da titulação obtida, Instituição de obtenção da titulação, Nome do aluno orientado, Nome do professor orientador, Título da tese, Tipo de orientação realizada, Titulação obtida pelo aluno.
- Atributos do arquivo de métricas genealógicas: Código identificador do pesquisador, Descendência, Descendência inversa, Fecundidade, Fecundidade inversa, Fertilidade, Fertilidade inversa, Gerações, Gerações inversa, Índice genealógico, Primos, Primos inverso, Relações e Relações inversas.

Os outros dois arquivos referem-se as pastas criadas na configuração prévia

a execução do *software*, de acordo com a sua programação, onde são depositados os dados de todos os descendentes dos docentes do PPGCEM/UFSCar. Ambos os arquivos contêm praticamente os mesmos dados, porém o que diferencia é o tipo de arquivo, sendo um em formato *CSV* e o outro em formato *XML*.

. É importante ressaltar que na pasta *xm/s* há muitos arquivos, porém os únicos que devem ser utilizados para análise, de acordo com Damasceno (2019) são os últimos arquivos de arestas e vértices, pois apresentam os dados com maior tratamento.