

**UMA PLATAFORMA PARA EXPLORAÇÃO DA
GENEALOGIA ACADÊMICA BRASILEIRA**

JOÃO MARCOS MARTINS DA COSTA COTA

**UMA PLATAFORMA PARA EXPLORAÇÃO DA
GENEALOGIA ACADÊMICA BRASILEIRA**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais como requisito parcial para a obtenção do grau de Mestre em Ciência da Computação.

ORIENTADOR: ALBERTO H. F. LAENDER

COORIENTADOR: RAQUEL O. PRATES

Belo Horizonte

Janeiro de 2021

Martins da Costa Cota, João Marcos.

C843u

Uma plataforma para exploração da genealogia acadêmica brasileira [manuscrito] / João Marcos Martins da Costa Cota. - 2021.

xvii, 91 f. il.

Orientador: Alberto Henrique Frade Laender

Coorientadora: Raquel Oliveira Prates

Dissertação (mestrado) - Universidade Federal de Minas Gerais, Instituto de Ciências Exatas, Departamento de Ciência da Computação

1. Computação – Teses. 2. Visualização de dados – Teses. 3. Bancos de dados – Teses 4. Genealogia acadêmica – Teses I. Laender, Alberto Henrique Frade. II. Prates, Raquel Oliveira. III. Universidade Federal de Minas Gerais; Instituto de Ciências Exatas, Departamento de Ciência da Computação. IV. Título.

CDU 519.6*71(043)



UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

FOLHA DE APROVAÇÃO

Uma Plataforma para Exploração da Genealogia Acadêmica Brasileira

JOÃO MARCOS MARTINS DA COSTA COTA

Dissertação defendida e aprovada pela banca examinadora constituída pelos Senhores:

PROF. ALBERTO HENRIQUE FRAIDE LAENDER - Orientador
Departamento de Ciência da Computação - UFMG

PROFA. RAQUEL OLIVEIRA PRATES - Coorientadora
Departamento de Ciência da Computação - UFMG

PROF. THIAGO MAGELA RODRIGUES DIAS
Departamento de Informática Gestão e Design - CEFET MG

PROF. ADO JÓRIO DE VASCONCELOS
Departamento de Física - UFMG

PROF. JOSÉ PALAZZO MOREIRA DE OLIVEIRA
Departamento de Informática Aplicada - UFRGS

Belo Horizonte, 20 de Janeiro de 2021.

Esta Dissertação dedico aos meus pais, minha fonte de inspiração da pessoa que me tornei hoje, João Anastácio e Maria da Conceição pelo apoio incondicional, sem o qual não seria possível a realização deste trabalho e da plataforma envolvida nele.

Agradecimentos

Aos meus pais, João Anastácio e Maria da Conceição, pelo apoio, dedicação e incentivo para mim, e aos meus estudos, desde sempre. Aos meus irmãos Mateus e Paulo que com conversa e apoio não mediram esforços para que eu chegasse até essa etapa da vida.

Ao professor e orientador Alberto Henrique Frade Laender e à professora e co-orientadora Raquel Oliveira Prates, pela paciência, orientação e incentivo que tornaram possível a conclusão desta dissertação. Agradeço também por terem me apresentado ao projeto de pesquisa que desenvolvi, a técnicas de desenvolvimento e análise de software, e por me permitirem ter acesso a diferentes ferramentas computacionais que foram importantes para a minha formação científica, além de ampliar meus horizontes além da fronteira do conhecimento científico.

Aos professores e colegas do Departamento de Ciência da Computação da Universidade Federal de Minas Gerais pela valiosa contribuição à minha formação durante as disciplinas cursadas.

Aos amigos do LBD e do Synergia, com os quais tive o prazer de conviver durante esse período, pela agradável convivência, discussões e colaboração durante a construção deste trabalho; em especial gostaria de agradecer ao Geraldo e ao Thiago Magela pela valiosa contribuição e pelas discussões científicas que culminaram em um trabalho colaborativo. Gostaria de agradecer também a Daniel Freitas, Issa, Mariana e Taiar, pelas lições aprendidas e pelo apoio ao meu trabalho, bem como ao professor Geraldo Robson Mateus pelas oportunidades. Um agradecimento especial aos participantes (anônimos) das entrevistas que foram realizadas nesta dissertação. Gostaria de pedir desculpa àqueles que eu não citei, mas não queria fazer isso e esquecer de alguém.

Aos amigos, que conheci neste departamento, no período da UFMG, e aos que carregou comigo desde o período do CEFET – não ousarei citar os nomes para não esquecer injustamente algum – pela amizade, apoio e união que foram cruciais em momentos difíceis dessa construção. Àqueles que se tornaram importantes, obrigado Elias e Daniel Reis pelo apoio e parcerias.

À minha companheira Fran pela ajuda na dissertação e pelo apoio, carinho, compreensão e incentivo. Por ter me acolhido e ajudado no momento mais desafiador deste período.

Gostaria de agradecer também aos meus familiares, aos meus tios e tias, e aos meus primos e primas, bem como à Geisiane por seus conselhos e orientações para a vida. Obrigado por terem acendido a luz do caminho do querer e serem mais-e-mais no meu engradecimento pessoal, intelectual e mental, e assim tornaram possível a conclusão deste trabalho.

Agradeço imensamente aos envolvidos no Programa de Pós-Graduação em Ciência da Computação DA UFMG pela ajuda para a conclusão desta importante etapa da minha formação.

Por fim, ao Synergia pela bolsa inicial e posterior contratação, e ao Inter. Agradeço aos órgãos de fomento CNPq, CAPES e FAPEMIG, como também à UFMG, pelo apoio financeiro dado ao desenvolvimento desta pesquisa.

*“The sun is a wondrous body. Like a magnificent father!
If only I could be so grossly incandescent!”*
(Solaire of Astora)

Resumo

Ao longo da história, muitos pesquisadores têm dado contribuições notáveis à ciência, não apenas avançando o conhecimento, mas também formando novos cientistas. Atualmente, identificar e estudar a formação de pesquisadores ao longo dos anos é uma tarefa desafiadora, pois os atuais repositórios de teses e dissertações são catalogados de forma descentralizada em diferentes bibliotecas digitais, muitas delas de escopo limitado. Nesta dissertação, damos um primeiro passo para a construção de um grande repositório que registre a genealogia acadêmica brasileira. Para isso foram coletados dados da plataforma Lattes do CNPq e desenvolvido um arcabouço computacional para gerar as árvores genealógicas acadêmicas dos pesquisadores brasileiros a partir deles, fornecendo ainda dados adicionais resultantes de uma série de análises que descrevem as principais propriedades dessas árvores. Nosso esforço identificou aspectos interessantes relacionados à formação acadêmica dos pesquisadores brasileiros, os quais destacam a importância da geração e catalogação de suas árvores genealógicas acadêmicas.

Palavras-chave: Visualização de dados, bancos de dados, genealogia acadêmica, plataforma Lattes, plataforma Árvore da Ciência.

Abstract

Throughout history, many researchers have made remarkable contributions to science, not only advancing knowledge, but also forming new scientists. Currently, identifying and studying the formation of researchers over the years is a challenging task, as the current repositories of theses and dissertations are cataloged in a decentralized manner in different digital libraries, many of them with limited scope. In this MSc dissertation, we take a first step towards building a large repository to record the Brazilian academic genealogy. For that, we collected data from the CNPq Lattes platform and developed a computational framework in order to generate the academic genealogy trees of Brazilian researchers from them, also providing additional data resulting from a series of analyses that describe the main properties of such trees. Our effort has identified interesting aspects related to the academic background of the Brazilian researchers, which highlight the importance of generating and cataloging their academic genealogy trees.

Keywords: Data visualization, data bases, Lattes platform, Science Tree.

Lista de Figuras

1.1	Ancestrais acadêmicos da Professora Joan Redwing, atual líder de um grupo de pesquisa da Universidade Estadual da Pennsylvania no Estados Unidos (https://sites.psu.edu/redwing/academic-family-tree , acesso em Abril de 2020) e descendente direta de Antoine Bacquerel, físico francês que recebeu o Prêmio Nobel de Física em 1903.	4
2.1	Representação dos ancestrais acadêmicos do Prof. Weichang Du da Universidade de New Brunswick, Canadá.	12
2.2	Representação dos ancestrais acadêmicos do Prof. Stacy McGaugh da Case Western Reserve University separados pelo nível de formação. Imagem disponível em http://astroweb.case.edu/ssm/academicgeneology.html , acesso em Abril de 2020.	13
2.3	Página da árvore genealógica acadêmica dos pesquisadores da área de meteorologia mantida pela Universidade do Estado da Flórida, disponível em http://moe.met.fsu.edu/familytree , acesso em Abril de 2020.	14
2.4	Crescimento dos dados ao longo dos anos na plataforma Mathematics Genealogy Project (imagem obtida no <i>website</i> do projeto, https://www.genealogy.math.ndsu.nodak.edu/growth_image.php , acesso em Abril de 2020).	15
2.5	Página de busca avançada da plataforma Mathematics Genealogy Project, em que é possível realizar buscas pelos nomes do pesquisador, instituição, ano de obtenção dos títulos e local de atuação, acesso em Abril de 2020.	16
2.6	Após a consulta pelo pesquisador, feita de forma simplificada ou avançada, é exibida uma tela com os seus dados presentes na plataforma Mathematics Genealogy Project, acesso em Abril de 2020.	16
2.7	Lista com a contagem da frequência dos totais de orientações, presente na plataforma Mathematics Genealogy Project, acesso em Abril de 2020.	17

2.8	Lista dos 25 matemáticos com maior número de orientações (filhos acadêmicos), presentes na plataforma Mathematics Genealogy Project, acesso em Abril de 2020.	17
2.9	Lista dos 25 matemáticos com maior número de descendentes diretos e indiretos, presentes na plataforma Mathematics Genealogy Project, acesso em Abril de 2020.	18
2.10	Página inicial da plataforma The Academic Family Tree. Acesso em Abril de 2020.	19
2.11	Árvore do Prof. José Palazzo Moreira de Oliveira da Universidade do Rio Grande do Sul, pesquisador da área de Ciência da Computação, que mostra os seus ascendentes e descendentes acadêmicos. Acesso em Abril de 2020. .	20
2.12	Página com dados detalhados do pesquisador selecionado. Acesso em Abril de 2020.	20
2.13	Busca da distância entre dois pesquisadores com o propósito de mostrar o caminho de orientações entre eles. Acesso em Abril de 2020.	21
2.14	Exibição do perfil de um pesquisador na plataforma Acácia. Acesso em Abril de 2020.	23
2.15	Exibição da Análise de Áreas disponível na plataforma Acácia, que mostra os pesquisadores listados em ordem decrescente em relação à quantidade de descendentes diretos e indiretos. Acesso em Abril de 2020.	24
2.16	Visualização de dois níveis da árvore de um pesquisador na plataforma The Gold Tree. Acesso em Abril de 2020.	25
2.17	Visualização dos dados referentes à orientação de doutorado de um descendente acadêmico escolhido aleatoriamente. Acesso em Abril de 2020.	25
3.1	Fluxo de dados da nova versão da plataforma Árvore da Ciência.	29
3.2	Esquema de comunicação entre os componentes da plataforma Árvore da Ciência.	30
3.3	Visualização da página de dados específicos de um pesquisador na plataforma <i>Science Tree</i>	32
3.4	Visualização da árvore de um pesquisador na plataforma <i>Science Tree</i> em que é possível identificar o pesquisador em destaque, bem como seus orientadores e orientados.	32
3.5	Busca pelo nome abreviado do pesquisador Luciano Antonio Digiampietri que retorna um resultado único.	34

3.6	Interface da plataforma Árvore da Ciência mostrando a árvore do pesquisador Luciano Antonio Digiampietri e a aba de dados extraídos do seu Currículo Lattes.	35
3.7	Árvore do pesquisador Luciano Antonio Digiampietri expandida a partir do nodo do seu orientador e mostrando a sua aba de métricas. Note que no gráfico da figura é possível clicar nas barras para obter a quantidade de orientações de mestrado e doutorado que elas indicam.	36
3.8	Página do Prof. Crodowaldo Pavan mostrando o primeiro nível de sua árvore genealógica acadêmica e dados gerais sobre a sua formação acadêmica.	41
3.9	Árvore do Prof. Crodowaldo Pavan expandida até o seu terceiro nível para visualização de um caminamento específico pela sua linhagem acadêmica. Logo abaixo da árvore é mostrada a aba com as suas respectivas métricas e o seu gráfico de orientações concluídas por ano.	42
4.1	Cabeçalho do questionário utilizado para avaliação da plataforma Árvore da Ciência.	48
4.2	Seção do questionário para identificação do respondentes.	49
4.3	Seção do questionário com questões a respeito da plataforma.	50
4.4	Página de agradecimento pelo preenchimento do questionário.	53
4.5	Gênero dos respondentes.	54
4.6	Idade dos respondentes agrupada por faixas de dez anos.	54
4.7	Total de respondentes por nível de formação.	55
4.8	Quantidade de respondentes por ocupação profissional.	56
4.9	Total de respondentes por área de formação.	56
4.10	Agrupamento dos respondentes de acordo com as respostas às questões relativas às características visuais dos dados da plataforma.	58
4.11	Agrupamento dos respondentes de acordo com as respostas às questões relativas aos aspectos visuais e aos filtros disponíveis na plataforma.	59
4.12	Agrupamento dos respondentes de acordo com as respostas às questões relativas à facilidade de uso e aprendizado da plataforma.	60
4.13	Opinião dos respondentes em relação à questão “Você teve dificuldade para entender o fluxo de funcionamento da plataforma?”.	61
4.14	Agrupamento dos respondentes de acordo com as respostas às questões relativas à utilidade da plataforma no dia-a-dia dos pesquisadores e para exploração das árvores genealógicas acadêmicas.	62
4.15	Gráfico referente às sugestões propostas para ampliação das informações gerais sobre a plataforma.	63

4.16	Sugestões propostas para melhoria da versão atual da plataforma Árvore da Ciência.	64
4.17	Sugestões propostas para ampliação do escopo da plataforma Árvore da Ciência.	66
4.18	Problemas relatados sobre a plataforma Árvore da Ciência.	68
4.19	Comentários adicionais sobre a plataforma Árvore da Ciência.	69
4.20	Acessos por estado no período de 1 de junho de 2019 a 1 de novembro de 2020. Dados coletados do serviço de estatísticas de tráfego do <i>Google Analytics</i>	78

Lista de Tabelas

2.1 Tabela comparativa dos principais recursos presentes nas quatro plataformas abordadas.	27
--	----

Sumário

Agradecimentos	vi
Resumo	ix
Abstract	x
Lista de Figuras	xi
Lista de Tabelas	xv
1 Introdução	1
1.1 Análise da Genealogia Acadêmica	2
1.2 Trabalhos Relacionados	4
1.3 Objetivo e Contribuições da Dissertação	8
1.4 Organização da Dissertação	9
2 Iniciativas Similares	11
2.1 Mathematics Genealogy Project	14
2.2 The Academic Family Tree	18
2.3 Plataforma Acácia	22
2.4 The Gold Tree	23
2.5 Considerações Finais	26
3 Descrição Funcional da Plataforma Proposta	28
3.1 Arquitetura Geral da Plataforma	29
3.2 Interface de Usuário	30
3.3 Detalhes de Implementação	36
3.3.1 Aplicação <i>Front-end</i>	37
3.3.2 Aplicação <i>Back-end</i>	38
3.4 Estudo de Caso	39

3.5	Considerações Finais	43
4	Avaliação com Usuários	46
4.1	Avaliação por Questionário	47
4.1.1	Perfil dos Respondentes	53
4.1.2	Avaliação Geral da Plataforma	57
4.1.3	Modo de Interação com a Plataforma	59
4.1.4	Utilidade da Plataforma	60
4.1.5	Comentários dos Respondentes	62
4.2	Avaliação por Entrevistas	69
4.2.1	Perfil dos Participantes	70
4.2.2	Tarefas e Perguntas	70
4.2.3	Resultados da Avaliação	74
4.3	Considerações Finais	77
5	Conclusões e Trabalhos Futuros	80
5.1	Conclusões	80
5.2	Trabalhos Futuros	82
	Referências Bibliográficas	83
	Apêndice A Termo de Consentimento da Avaliação por Entrevistas	90

Capítulo 1

Introdução

Desde a década de 1930 sociólogos estudam o comportamento da sociedade sob muitos aspectos e, em especial, frente à relação do ser humano com a informação, utilizando, dentre outras ferramentas, as redes complexas, ou seja, grafos. Grafos podem ser entendidos como um conjunto de nós conectados por arestas. No livro “*Networks, Crowds, and Markets: Reasoning about a Highly Connected World*” [Easley & Kleinberg, 2010], os autores analisam, de forma detalhada e por meio de diversas formas de visualização, diferentes redes complexas por eles abordadas, procurando avaliar o impacto delas em várias aplicações. Como resultado, foi observado que o modelo de grafo escolhido para tratar adequadamente uma rede não é único e depende de fatores como formato e volume dos dados, informações relevantes consideradas e também do modo de exibição adotado. Essas observações evidenciam que a complexidade de um grafo é acompanhada por um aumento no custo computacional e que a escolha do método para análise da rede é, portanto, estratégica.

O desenvolvimento da Ciência da Computação e áreas correlatas permitiu a evolução da capacidade de processamento e armazenamento dos computadores, possibilitando que novos resultados de pesquisa das mais variadas áreas possam ser analisados de forma detalhada e eficiente. O volume de dados atualmente disponível, no entanto, requer uma alta capacidade de armazenamento e organização a fim de gerar informação acessível [Dhar, 2013; van der Aalst, 2016]. Como exemplos de repositórios com tais características, podemos citar a DBLP Computer Science Bibliography¹, hospedada na Universität Trier, na Alemanha; a MEDLINE² (sigla em inglês para Sistema Online de Busca e Análise de Literatura Médica) que é o repositório bibliográfico da Biblioteca

¹<https://dblp.uni-trier.de>, acesso em Abril de 2020

²<https://pubmed.gov/>, acesso em Abril de 2020.

Nacional de Medicina dos Estados Unidos; e o Portal SOL³ que é o portal de publicações e conteúdos da Sociedade Brasileira de Computação (SBC). Especificamente, diversos pesquisadores se utilizam de ferramentas computacionais para tratar dados coletados dessas bibliotecas digitais, visando a realização de trabalhos que exploram e analisam a produção científica nacional e internacional [Abramo et al., 2018; Cunningham, 2001; Dawson et al., 2014; Iglič et al., 2017; Lane, 2010; Lima et al., 2015; Liu et al., 2005; Lopes et al., 2012; Silva et al., 2016; Su et al., 2018].

Outra fonte importante de dados acadêmicos disponível no país é a Plataforma Lattes do CNPq⁴, utilizada nesta dissertação. Essa plataforma é um esforço do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), tendo sido criada com o objetivo de integrar, em um único sistema de informação, currículos de pesquisadores de diferentes níveis de formação, grupos de pesquisa, e instituições de ensino e pesquisa do país. A importância dessa plataforma, reconhecida internacionalmente [Lane, 2010], pode ser evidenciada pela sua pluralidade de funções, que cobrem desde o gerenciamento de dados curriculares até o cumprimento de pré-requisitos para a submissão de propostas de projeto às várias agências de fomento à pesquisa, bem como para apoio à formulação das políticas do Ministério da Ciência, Tecnologia e Inovações (MCTI) e de outros órgãos governamentais da área. Devido à abrangência dos seus dados, essa plataforma tem sido usada também como fonte para diversos projetos de pesquisa que abordam a produção científica brasileira [Araújo et al., 2014; Digiampietri et al., 2015, 2017, 2019; Dores et al., 2017; Lima et al., 2015; Mena-Chalco et al., 2014; Pessoa Junior et al., 2020; Rossi & Mena-Chalco, 2014; Silva et al., 2016; Tuesta et al., 2015].

1.1 Análise da Genealogia Acadêmica

Uma maneira relevante de caracterizar o impacto de um pesquisador é por meio de sua genealogia acadêmica [Damaceno et al., 2019; Hirshman et al., 2016; Sugimoto et al., 2011]. O termo genealogia acadêmica ou científica refere-se à organização de uma árvore genealógica de cientistas e estudiosos estruturada de acordo com as relações de orientação acadêmica (mestrado e doutorado) ao longo de suas respectivas carreiras [Jackson, 2007]. Neste contexto, a genealogia acadêmica de um pesquisador pode influenciar até mesmo os resultados de uma pesquisa específica. Por exemplo, Hirshman et al. [2016], ao analisarem a literatura médica relacionada à cirurgia de gli-

³<https://sol.sbc.org.br/index.php/indice>, acesso em Abril de 2020.

⁴<http://lattes.cnpq.br>, acesso em Abril de 2020

oma⁵, concluíram que ela tem sido bastante influenciada por genealogias médicas, ou seja, os autores dessa literatura geralmente fazem parte de uma equipe médica cujos integrantes foram formados ao longo dos anos por alguns de seus próprios membros.

A Figura 1.1 mostra uma representação parcial da genealogia acadêmica do cientista Antoine Becquerel⁶, físico francês responsável por estudos que culminaram na descoberta do fenômeno da radioatividade. Segundo a sua árvore genealógica acadêmica, Antoine Becquerel deu origem a uma rede de pesquisadores voltados para estudos envolvendo radioatividade e produção de energia. A sua primeira descendente acadêmica foi a também renomada cientista Marie Curie, que por sua vez gerou novos descendentes pela orientação de Emile Henriot. Esse movimento se repete até a formação da Professora Joan M. Redwing em 1994. Baseando-se apenas nesse ramo da árvore apresentado na Figura 1.1, observa-se uma linhagem de pesquisadores de grande prestígio formada a partir de Antoine Becquerel e que inclui a Professora Joan M. Redwing, atualmente vinculada ao Departamento de Ciência e Engenharia de Materiais da Penn State University, cuja árvore, de acordo com o sítio Academic Family Tree, possui pelo menos 11 descendentes acadêmicos⁷. Assim, conhecer a linhagem acadêmica de um pesquisador é importante porque permite mostrar não só as suas influências, mas também conhecer mais a respeito de sua formação acadêmica e, até mesmo, prever possíveis desdobramentos futuros de sua linha de pesquisa.

Por sua riqueza de informação, além de sua crescente confiabilidade e abrangência, a plataforma Lattes tornou-se um instrumento indispensável para mapear os pesquisadores brasileiros e a sua produção acadêmica. Desse modo, ela se apresenta como uma fonte de dados natural para o desenvolvimento de uma plataforma para visualização e exploração da genealogia acadêmica dos pesquisadores da comunidade científica brasileira, pois possibilita acesso não apenas a dados específicos relativos à carreira acadêmica e à produção científica desses pesquisadores, como também registra todas as suas orientações concluídas, tornando possível a construção de árvores genealógicas acadêmicas semelhantes à que foi mostrada na Figura 1.1.

Apesar dos inúmeros trabalhos que procuram analisar a dinâmica da produção científica em termos das redes de colaboração científica [Adams, 2012; Barabási et al., 2002; Coccia & Wang, 2016; Digiampietri et al., 2019; Mena-Chalco et al., 2014; Pessoa Junior et al., 2020], ainda são poucos aqueles que visam estudar a genealogia acadêmica e o processo de mentoria. Dentre os pioneiros, Jackson [2007] reporta os primeiros

⁵Glioma é um tumor de células gliais, células que protegem, nutrem e dão suporte aos neurônios (see <https://pt.wikipedia.org/wiki/Glioma>, acesso em Abril de 2020).

⁶https://en.wikipedia.org/wiki/Henri_Becquerel, acesso em Abril de 2020

⁷<https://academictree.org/etree/tree.php?pid=480318>, acesso em Abril de 2020

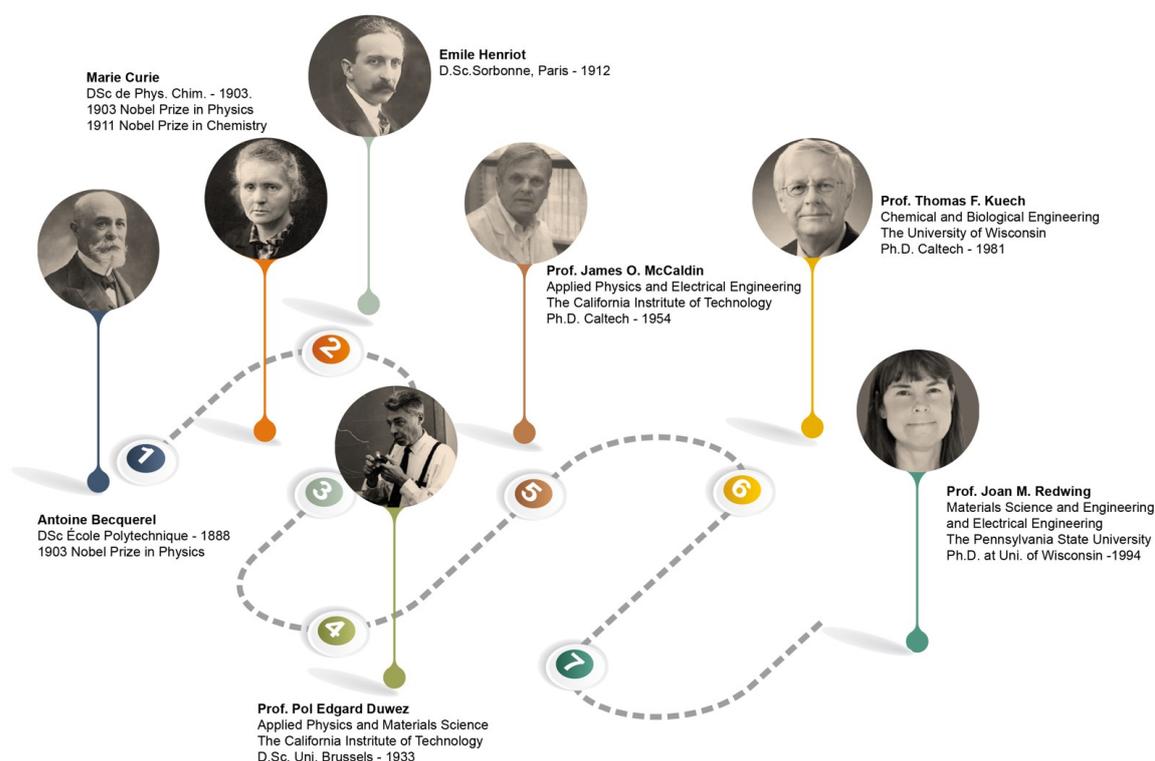


Figura 1.1. Ancestrais acadêmicos da Professora Joan Redwing, atual líder de um grupo de pesquisa da Universidade Estadual da Pennsylvania no Estados Unidos (<https://sites.psu.edu/redwing/academic-family-tree>, acesso em Abril de 2020) e descendente direta de Antoine Bacquerel, físico francês que recebeu o Prêmio Nobel de Física em 1903.

esforços realizados para a construção da árvore genealógica da Matemática e que deu origem ao *Mathematics Genealogy Project*⁸. Inicialmente o projeto contou com a colaboração de 3.500 pesquisadores que enviaram seus dados por meio de correspondências pessoais, contando o projeto atualmente com mais de 256.000 nomes, evidenciando a relevância de se integrar, em uma mesma plataforma, dados de orientadores e orientados de diversas áreas do conhecimento [Chang, 2003; Hart & Cossuth, 2013; Malmgren et al., 2010; Tuesta et al., 2015].

1.2 Trabalhos Relacionados

Com o passar dos anos, tem-se observado um crescimento dos trabalhos de pesquisa que procuram analisar a dinâmica e o impacto da produção científica. Alguns desses trabalhos têm como objetivo analisar redes específicas de colaboração científica, tendo como foco principalmente a produção acadêmica dos pesquisadores envolvidos

⁸<https://www.genealogy.math.ndsu.nodak.edu/>, acesso em Abril de 2020

ou dos seus grupos de pesquisa [Benevenuto et al., 2015; Delgado-Garcia et al., 2014; Digiampietri et al., 2015; Dores et al., 2016; Mena-Chalco et al., 2014; Montoya et al., 2018; Newman, 2001, 2004; Lima et al., 2015; Sonnenwald, 2007; Strehl et al., 2016]. Essas análises são, em geral, baseadas em dados coletados em grande maioria de fontes abertas como a DBLP Computer Science Bibliography e a Networked Digital Library of Theses and Dissertations⁹ (NDLTD), bem como a partir de dados obtidos da MEDLINE e da plataforma Lattes do CNPq por meio de coletas específicas.

Outros trabalhos têm procurado analisar aspectos distintos da produção acadêmica como a transformação da pesquisa científica e a sua relação com a evolução da pós-graduação [Almeida & Guimarães, 2017], além das tendências observadas nas pesquisas realizadas [Montoya et al., 2018] e o impacto acadêmico proporcionado pelos próprios pesquisadores [Heinisch & Buenstorf, 2018; Rossi et al., 2019]. Nesses estudos são levados em consideração não apenas os aspectos relacionados ao crescimento e à mutabilidade dos grupos de pesquisa, mas também à evolução e ao impacto de campos específicos do conhecimento [Heinisch & Buenstorf, 2018; Hirshman et al., 2016].

Mais recentemente, vários trabalhos têm enfatizado o estudo da genealogia acadêmica visando analisar a capacidade de um pesquisador em formar novos cientistas, seja em sua área de atuação ou em colaboração com pesquisadores de áreas distintas. Em um trabalho pioneiro, Dores et al. [2016] apresentaram os resultados de um estudo sobre a genealogia acadêmica internacional realizado a partir de dados coletados da NDLTD - Networked Digital Library of Theses and Dissertations. Essa biblioteca digital armazena dados de teses e dissertações defendidas em instituições acadêmicas de todo o mundo e possui como principal característica o fato de armazenar esses dados em formato aberto, o que possibilita coletá-los diretamente do seu repositório. Dos mais de 4,5 milhões de registros coletados da coleção de teses e dissertações da NDLTD, apenas 638.812 foram efetivamente considerados após um processo de padronização e limpeza dos dados. Esses registros receberam ainda um tratamento para desambiguação dos nomes dos pesquisadores, de modo que as árvores genealógicas acadêmicas¹⁰ pudessem ser geradas com a maior acurácia possível. Uma caracterização geral das árvores obtidas mostrou que elas são geralmente mais largas do que profundas e também que árvores com maior número de nós são em sua maioria as mais profundas, refletindo de certa forma a maior experiência dos pesquisadores mais próximos da raiz.

Com o objetivo de reconstruir as relações de orientação acadêmica dos pesquisadores brasileiros, Dores et al. [2017] realizaram a coleta e o processamento de 222.674

⁹<http://www.ndltd.org>, acesso em Abril de 2020

¹⁰Conforme Dores et al. [2017], consideramos uma árvore genealógica acadêmica como sendo um grafo dirigido acíclico.

currículos de pesquisadores com doutorado da plataforma Lattes. Entre os resultados relatados, os autores identificaram 903.183 vértices (orientadores e orientados), 1.144.051 arestas (orientações), 70.610 árvores e 22.061 componentes distintos, que podem ser entendidos como árvores conectadas, por alguma relação de orientação ou coorientação em comum. As árvores reportadas tinham em média 40,19 nós, apresentando uma largura média de 3,81 nós. As dez maiores árvores possuíam mais de cinco mil nós, embora 80% delas com menos de 20 nós. Quanto à distribuição da profundidade das árvores, cerca de 50% delas tinham apenas um nível, significando que muitos pesquisadores ainda não possuíam uma maior experiência de orientação. Uma importante contribuição derivada desse trabalho foi a implementação de uma primeira versão da plataforma *Árvore da Ciência* (então denominada *Science Tree*) de modo a possibilitar a exploração das árvores geradas a partir dos currículos coletados da plataforma Lattes [Dores, 2017], o que permitiu identificar os desafios que precisavam ser tratados para o desenvolvimento de uma interface mais funcional para exploração da genealogia acadêmica dos pesquisadores brasileiros.

Em um trabalho com objetivos similares, Damaceno et al. [2017, 2019] também propuseram uma estratégia para extração de árvores genealógicas acadêmicas¹¹ a partir de dados da plataforma Lattes, mas que, ao contrário do trabalho de Dores et al. [2017], considera os currículos de todos os pesquisadores com mestrado e doutorado que possuam indicações explícitas ou implícitas a relacionamentos de orientação acadêmica. Para fins de comparação, Damaceno et al. [2019] identificaram, a partir de 737.919 currículos de mestres e doutores coletados, 1.111.544 vértices, correspondentes a pesquisadores distintos, e 1.208.398 arestas, que representam relacionamentos de orientação entre tais pesquisadores. A validação dessas árvores se deu manualmente por meio de uma comparação amostral entre os dados existentes nos currículos dos pesquisadores e aqueles mostrados nas árvores geradas, tendo sido detectado um número final de descendentes maior do que aquele registrado pelos próprios pesquisadores em seus currículos. Isso ocorreu devido ao fato de alguns pesquisadores não terem registrado em seus currículos tais orientações que, por sua vez, foram indicadas nos currículos de seus orientados, caracterizando uma informação implícita de orientação. Esse trabalho também mostrou que a idade acadêmica dos pesquisadores apresenta uma influência direta em suas métricas e também nas métricas de suas respectivas áreas, indicando haver também um certo grau de interdisciplinaridade desde o início do registro da mentoria acadêmica formal no Brasil e que a geração do conhecimento no país tem ocorrido de forma colaborativa.

¹¹Nesse trabalho, os autores denominam as árvores usadas para representar a genealogia acadêmica de grafos, já que tais tais árvores correspondem a grafos dirigidos conforme definido anteriormente.

Uma outra vertente dos trabalhos relacionados à genealogia acadêmica inclui aqueles que propõem métricas específicas para que se possa aferir o desempenho dos pesquisadores em relação à evolução de suas árvores, como é o caso do *índice genealógico* proposto por Rossi et al. [2017]. Esse índice, que pode ser aplicado para medir a influência acadêmica de pesquisadores, é usado para caracterizar árvores genealógicas acadêmicas em termos de sua topologia. Ao ser acumulado ao longo de vários níveis da árvore, ele possibilita uma forma efetiva de mensurar e analisar a influência acadêmica de um orientador na comunidade científica em que ele está inserido.

Rossi et al. [2018] também utilizam várias métricas para avaliar a capacidade de orientação de um orientador acadêmico (docente ou pesquisador), como *fecundidade* (número de dependentes ou descendentes diretos de um orientador acadêmico), *fertilidade* (capacidade de um orientador acadêmico gerar novos orientados), *descendência* (quantidade de descendentes diretos e indiretos de um orientador acadêmico, ou seja, total de nós de sua árvore), *número de gerações* (maior distância hierárquica entre um orientador acadêmico e seus descendentes) e *relacionamentos* (número de relacionamentos entre orientações na descendência de um orientador acadêmico). Essas métricas visam analisar dados de genealogia acadêmica de modo a obter características relativas ao desempenho de orientação de um acadêmico. Nesse trabalho, as métricas foram aplicadas a dados extraídos das plataformas *Mathematics Genealogy Project* e *The Academic Family Tree*. A partir dos resultados obtidos, foi possível observar um padrão evolutivo das comunidades acadêmicas principalmente na relação orientador-orientado.

Uma outra estratégia para análises relacionadas à genealogia acadêmica envolve a utilização de técnicas de aprendizado de máquina para identificar orientadores entre coautores (frequentes) nas publicações de estudantes de doutorado, uma vez que, segundo seus autores [Heinisch & Buenstorf, 2018], a taxa de acerto dessas relações é estatisticamente maior quando comparada à de outros métodos. Nessa abordagem é realizada primeiramente a coleta de dados sobre dissertações, teses e publicações. Em seguida é estabelecida a relação entre os autores utilizando algoritmos de aprendizado de máquina. Por fim, é realizado o cálculo de métricas tais como fecundidade¹², número de publicações e número de citações com o intuito de estabelecer uma previsão sobre os doutorandos tornarem-se orientadores. Essa abordagem foi aplicada a dados de pesquisadores das áreas de engenharia e física na Alemanha, que indicaram que um maior tempo de experiência do orientador não necessariamente produz uma quantidade relevante de futuros orientadores, sugerindo ainda que trabalhar com bons orientadores no

¹²Nesse trabalho, fecundidade é definida como a probabilidade de um doutorando tornar-se orientador.

início da carreira é fundamental para o futuro acadêmico de estudantes de doutorado.

Ainda em relação a desempenho, produtividade e impacto científico de pesquisadores, Liu et al. [2018] realizaram um estudo com o objetivo de descobrir a correlação entre a produtividade dos orientadores e a qualidade dos orientados formados na área de Ciência da Computação. Para tal, foram realizadas análises quantitativas e qualitativas dos currículos de pesquisadores com no mínimo cinco publicações em um espaço de dez anos, a partir de dados coletados da DBLP, tendo os seus resultados sido comparados com dados do AMINER¹³. Foi observado que pesquisadores em início de carreira são responsáveis por uma alta produção acadêmica associada a uma grande quantidade de orientações, seguindo-se um período de sustentação e, por fim, um período em que a produção acadêmica decresce. Mostrou-se também que pesquisadores talentosos tendem a formar orientados mais qualificados se comparado a pesquisadores com produção acadêmica mediana. No entanto, é importante ressaltar que a tendência observada nesse trabalho refere-se apenas à área de Ciência da Computação e, portanto, não deve ser generalizada para outras áreas do conhecimento.

1.3 Objetivo e Contribuições da Dissertação

A construção de árvores genealógicas acadêmicas apresenta desafios variados tais como a coleta dos dados que muitas vezes não são apresentados em um formato padronizado, o tratamento dos dados coletados para que possam ser devidamente aplicados para a geração das árvores e, particularmente, a desambiguação de nomes em razão da existência de homônimos que dificultam a correta identificação dos pesquisadores (orientadores e orientados) [Dores et al., 2016; Jackson, 2007; Malmgren et al., 2010; Tuesta et al., 2015]. Assim, faz-se necessária a criação de ferramentas que permitam a exploração dessas árvores e apresentem resultados que possam gerar informação de valor para o usuário por meio de métricas capazes de expressar em termos quantitativos e qualitativos a relevância das orientações identificadas.

Conforme será mostrado no Capítulo 3, embora a interface adotada por Dores [2017] para visualização das árvores genealógicas acadêmicas geradas a partir de dados da plataforma Lattes tenha sido importante para a depuração e melhoria do algoritmo então proposto para a criação dessas árvores, o esquema de visualização em si apresentava uma série de limitações para os seus usuários. Dentre as principais limitações, não era possível realizar uma navegação hierárquica pelas árvores, cuja representação era muito simples, impedindo uma exploração ampla das relações acadêmicas entre os

¹³<https://www.aminer.org>, acesso em Abril de 2020

pesquisadores. Além disso, a ausência de métricas que pudessem caracterizar a evolução das árvores dos pesquisadores, como número de descendentes diretos e indiretos, ou mesmo a sua profundidade (número de gerações), impedia qualquer tipo de análise mais específica sobre a formação de novos pesquisadores.

Nesse sentido, esta dissertação teve como objetivo desenvolver uma nova versão da plataforma *Árvore da Ciência* de modo a possibilitar uma exploração mais ampla da genealogia acadêmica brasileira. Para tal, ela compreendeu as seguintes etapas:

- Projeto e implementação de uma plataforma com recursos gráficos específicos para uma melhor visualização e exploração das árvores genealógicas acadêmicas geradas a partir de dados dos pesquisadores brasileiros coletados da plataforma Lattes do CNPq;
- Disponibilização dessa nova versão da plataforma para utilização pública na *World Wide Web*;
- Realização de uma extensa avaliação com a participação de usuários de diversas áreas do conhecimento visando validar a usabilidade da nova versão da plataforma *Árvore da Ciência*. Essa avaliação se deu em duas etapas, a primeira realizada por meio de um questionário enviado para 5834 potenciais usuários, dos quais 286 retornaram as suas respostas, e a segunda realizada presencialmente com um grupo seletivo de usuários formado por sete pesquisadores com diferentes níveis de carreira acadêmica.

1.4 Organização da Dissertação

Os demais capítulos desta dissertação estão organizados da seguinte forma:

- O Capítulo 2 apresenta uma breve descrição de algumas iniciativas similares para exploração da genealogia acadêmica;
- O Capítulo 3 apresenta uma descrição funcional da plataforma proposta, caracterizando a sua arquitetura geral e a interface de usuários proposta, além de abordar os principais detalhes de sua implementação e apresentar um estudo de caso ilustrando a sua utilização;
- O Capítulo 4 descreve a metodologia adotada e os resultados da avaliação da plataforma realizada com usuários de diferentes perfis acadêmicos;

- Finalmente, o Capítulo 5 apresenta as conclusões da dissertação e sugere algumas direções para trabalhos futuros.

Capítulo 2

Iniciativas Similares

A construção de árvores genealógicas acadêmicas tem sido abordada por diferentes grupos de pesquisa, como uma ferramenta alternativa para auxiliar na compreensão da formação de novos pesquisadores nas diversas áreas do conhecimento [Anil et al., 2018; Damaceno et al., 2017, 2019; Digiampietri et al., 2019; Dores et al., 2016, 2017; Jackson, 2007; Rossi et al., 2019]. Assim, este capítulo tem por finalidade descrever algumas iniciativas específicas, pessoais e institucionais, que procuram registrar o processo de formação acadêmica de determinadas áreas do conhecimento por meio de árvores genealógicas acadêmicas.

Um primeiro exemplo é o caso do Prof. Weichang Du, da Faculdade de Ciência da Computação da *University of New Brunswick*, Canadá, que possui em sua página pessoal uma representação parcial de sua genealogia acadêmica¹, ilustrada na Figura 2.1. Nela podem ser observados apenas dados básicos de cada um de seus ancestrais acadêmicos, como nome, título obtido, ano de titulação e um fato relevante relacionado ao pesquisador mencionado. Tal página, entretanto, não permite qualquer tipo de interação e nem apresenta qualquer outro dado específico.

Outra iniciativa interessante de ser mencionada é a genealogia acadêmica do Prof. Stacy McCaugh, da Universidade Case Western Reserve, representada em sua página institucional². Conforme ilustrado na Figura 2.2, ela evidencia como se deu a sua formação, indicando os seus ancestrais acadêmicos em todos os seus níveis de formação.

¹<http://www.cs.unb.ca/~wdu/webmain/familytree.html>, Acesso em Novembro de 2020

²<http://astroweb.case.edu/ssm/academicgeneology.html>, Acesso em Novembro de 2020.

Academic Family Tree

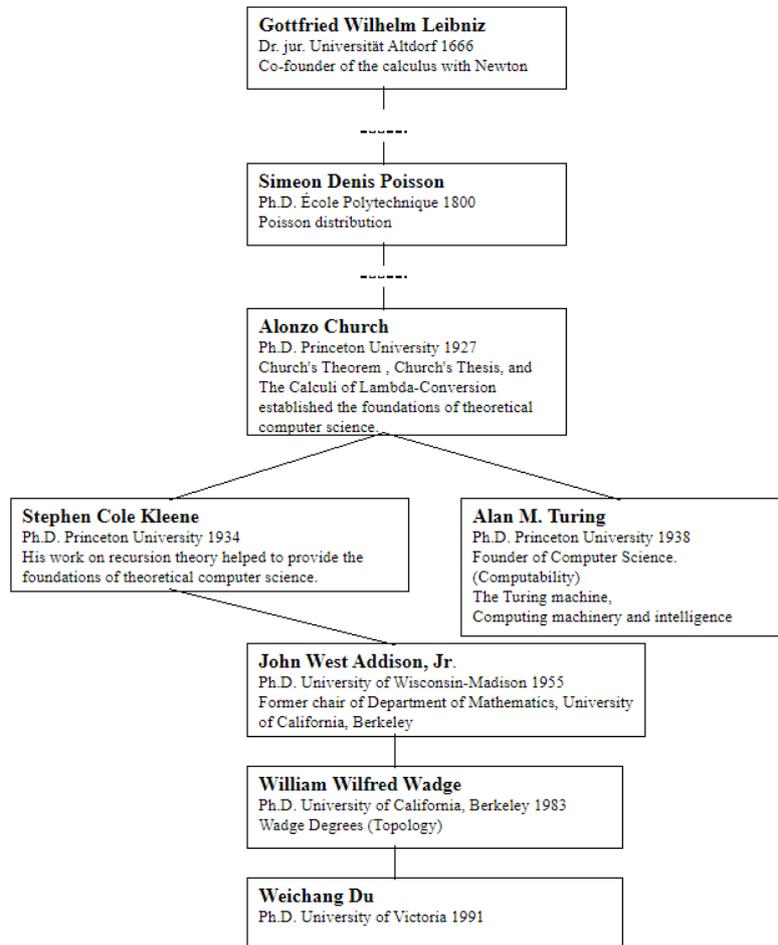


Figura 2.1. Representação dos ancestrais acadêmicos do Prof. Weichang Du da Universidade de New Brunswick, Canadá.

A Universidade do Estado da Flórida, também apresenta uma iniciativa interessante que é o sítio da árvore genealógica acadêmica dos meteorologistas americanos³, mantido pela Sociedade Meteorológica Americana⁴ [Hart & Cossuth, 2013]. Nesse sítio é possível explorar as árvores de diferentes comunidades de meteorologistas americanos. Além disso, é possível a um pesquisador solicitar a sua inclusão na árvore por meio de uma função específica (*Add Yourself/Someone Else*) como ilustrado na Figura 2.3. Entretanto, o sítio não oferece uma forma automática de geração ou visualização das árvores, de modo que elas são compiladas de tempos em tempos e disponibilizadas em formato de imagem.

³<http://moe.met.fsu.edu/familytree>, acesso em Abril de 2020

⁴<https://annual.ametsoc.org/index.cfm/2020/programs/conference-at-a-glance/academic-family-tree/>, acesso em Abril de 2020

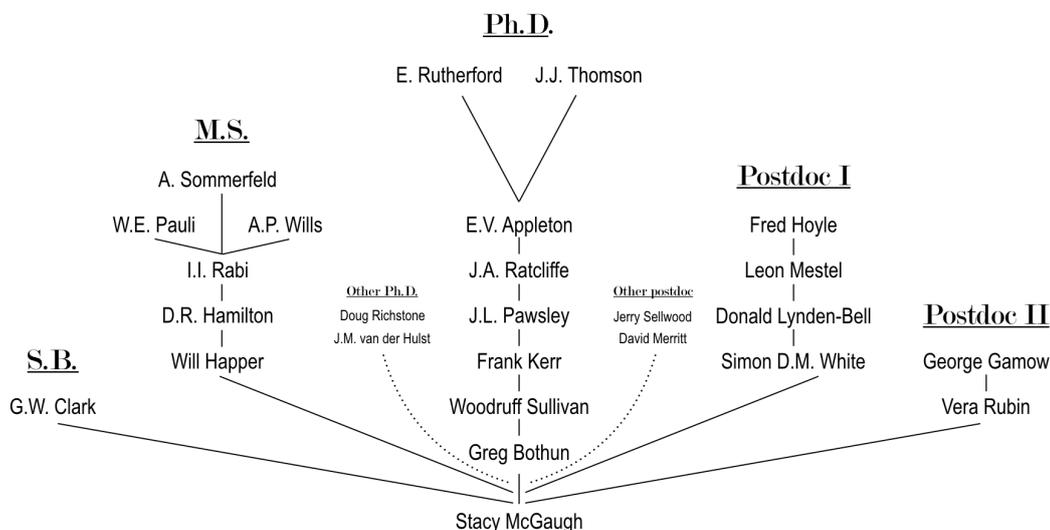


Figura 2.2. Representação dos ancestrais acadêmicos do Prof. Stacy McGaugh da Case Western Reserve University separados pelo nível de formação. Imagem disponível em <http://astroweb.case.edu/ssm/academicgeneology.html>, acesso em Abril de 2020.

Vale salientar ainda que a busca por uma representação ou até mesmo a investigação sobre as relações orientador-orientado não é recente. Tais iniciativas são comumente iniciadas por comunidades acadêmicas específicas no sentido de analisar, classificar e documentar essas relações de orientação. No entanto, esses esforços até então concentravam-se em áreas específicas, como a Matemática [Jackson, 2007] e a Neurociência [David & Hayden, 2012] que tiveram continuidade e são apresentadas neste capítulo, ou eram restritas a uma determinada comunidade, como no caso da retrospectiva da carreira de proeminentes físicos americanos [Chang, 2003].

Além dessas iniciativas pessoais ou relacionadas a departamentos e sociedades acadêmicas, algumas plataformas foram criadas especificamente com o intuito de possibilitar a exploração da genealogia acadêmica de forma mais ampla. Assim, nas demais seções deste capítulo são descritas as seguintes plataformas: Mathematics Genealogy Project⁵, The Academic Family Tree⁶, The Gold Tree⁷ e Acácia⁸. A escolha dessas plataformas se deu principalmente por serem as iniciativas com maior quantidade de pesquisadores cadastrados e também pela sua relevância para a área ou comunidade científica que representam.

⁵<https://www.genealogy.math.ndsu.nodak.edu/>, Acesso em Novembro de 2020

⁶<https://academicfamilytree.org/>, Acesso em Novembro de 2020

⁷<http://thegoldtree.c3.furg.br/>, Acesso em Novembro de 2020

⁸<http://plataforma-acacia.org>, Acesso em Novembro de 2020



Meteorology / Atmos.
Science
Academic Family Tree

Robert Hart¹ and Josh Cossuth²

¹Department of Earth, Ocean, and Atmospheric Science, FSU
²NPS/ONR

Photos above (L-R): C. Rossby, H. Riehl, J. Simpson, J. Charney, P. Austin, S. Syono, Sir B. J. Mason, V. Bjerknes

[[Home](#) | [Help/FAQ](#) | [Latest Updates](#) | [Full Name List](#) | [Most Wanted](#) | [Individual Trees](#) | [Quotes & Stories](#) | [Add Yourself/Someone Else](#) | [Acknowledgments](#)]

Tree versions and updates:

Version	Format & Link	Date of Update	Viewing Size (Inches)	Scope	Number of entries
20200228-all-large	PDF SVG	28 February 2020	400x400	All Origins "Meteorology"	5944
20200228-all	PDF SVG	28 February 2020	200x200	All Origins "Meteorology"	5944
20160711-allcloudphysics	PDF SVG	11 July 2016	200x200	All Origins "CloudPhysics"	1314
20200228-post1800	PDF SVG	28 February 2020	200x100	Post-1800 "Meteorology"	5542
20160711-post1800cloudphysics	PDF SVG	11 July 2016	200x100	Post-1800 "CloudPhysics"	995
20130213-post1800	PDF	13 February 2013	200x50	Post-1800 "Tropical Met."	1143

We strongly recommend the SVG versions for browsing/viewing.
Please read the important help/notes at the "Help/FAQ" link above before downloading the documents.

Figura 2.3. Página da árvore genealógica acadêmica dos pesquisadores da área de meteorologia mantida pela Universidade do Estado da Flórida, disponível em <http://moë.met.fsu.edu/familytree>, acesso em Abril de 2020.

A respeito das plataformas selecionadas, foram definidos quatro aspectos considerados relevantes para avaliar a usabilidade e a forma de visualização dos dados. São eles: (1) forma de representação da genealogia acadêmica, isto é, se a genealogia do pesquisador é apresentada em formato tabular ou em forma de grafo; (2) quantidade de métricas consideradas para analisar a genealogia de um pesquisador; (3) forma de visualização dos dados, se visual ou gráfica; e (4) foco da plataforma (pesquisadores brasileiros ou de outras nacionalidades). Ao final deste capítulo é apresentada uma tabela resumindo essa análise com o intuito de promover uma comparação entre as plataformas consideradas a partir dos aspectos listados acima. Esses aspectos são importantes pois, como veremos, definem as principais diretrizes para o desenvolvimento e futuro aperfeiçoamento da plataforma Árvore da Ciência.

2.1 Mathematics Genealogy Project

O Mathematics Genealogy Project [Jackson, 2007] é um dos mais antigos esforços no sentido de prover uma plataforma para visualização das árvores genealógicas acadêmicas de uma comunidade científica específica. Conforme descrito em sua página, essa plataforma tem como objetivo compilar dados sobre os matemáticos de todo o mundo e

também possibilitar que as instituições que realizam pesquisas na área de matemática cadastrem os seus pesquisadores para que possam ter os seus dados divulgados amplamente. Atualmente a plataforma conta com dados de mais de 253 mil matemáticos ao redor do mundo, conforme mostra o gráfico da Figura 2.4.

A plataforma *Mathematics Genealogy Project* apresenta os seus dados em forma textual. As buscas na plataforma podem ser realizadas a partir do nome do matemático em questão ou ainda de uma maneira mais refinada por meio de recursos avançados, conforme mostrado na Figura 2.5. Os dados disponíveis sobre um pesquisador são: nome, titulação, instituição, título de sua dissertação ou tese, nome(s) do(s) orientador(es) e a relação dos seus orientados. Na tabela que mostra a relação de orientados de um orientador (Figura 2.6), eles são apresentados com o nome no formato *<sobrenome, primeiro nome>*, seguido do nome da instituição em que realizaram o mestrado ou doutorado, do ano de conclusão e da quantidade de seus descendentes, além de ser possível ordenar os orientados cronologicamente ou por sobrenome.

Por fim, a plataforma apresenta algumas métricas gerais que refletem os seus próprios dados, além de estatísticas específicas, como: a distribuição dos pesquisadores por total de orientações (Figura 2.7), a lista dos 25 pesquisadores com o maior número de orientações (Figura 2.8) e a lista dos 25 pesquisadores com maior número de descendentes diretos e indiretos (Figura 2.9). Como pode ser observado, as métricas disponíveis nessa plataforma são bastante restritas, limitando assim as análises que podem ser feitas a partir de seus dados.

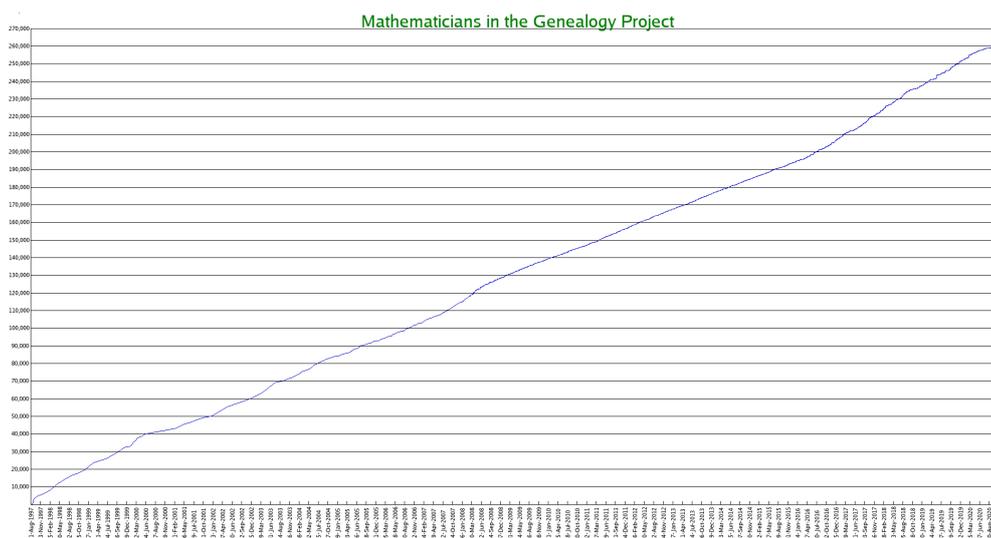


Figura 2.4. Crescimento dos dados ao longo dos anos na plataforma Mathematics Genealogy Project (imagem obtida no *website* do projeto, https://www.genealogy.math.ndsu.nodak.edu/growth_image.php, acesso em Abril de 2020).

Mathematics Genealogy Project

Search the database

256304 records as of 26 April 2020
View the [growth](#) of the genealogy project

Thank you for visiting the Mathematics Genealogy Project web page. We hope you find the information here both interesting and helpful. While you are visiting this site, please keep in mind that this is an ongoing project. Due to the difficulty in obtaining the type of organized data we need, and the varying sources of data, this project is continuously changing.

Not all fields are required to search for a particular mathematician. Partial matches will work as well. See examples and tips below.

First/Given Name

Middle Name

Last/Family Name

Name of School

Year of Degree

Thesis Keyword

Country

Math Subject Class

Examples

To search for all Mathematicians who received their degree from the University of Wisconsin in 1950:
Name of School: University of Wisconsin
Year of Degree: 1950

Figura 2.5. Página de busca avançada da plataforma Mathematics Genealogy Project, em que é possível realizar buscas pelos nomes do pesquisador, instituição, ano de obtenção dos títulos e local de atuação, acesso em Abril de 2020.

Mathematics Genealogy Project

Roger Meyer Temam
[MathSciNet](#)

Ph.D. Université de Paris 1967

Dissertation: *Sur La Stabilité Et La Convergence De La Methode Des Pas Fractionnaires*

Advisor: [Jacques-Louis Lions](#)

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

Name	School	Year	Descendants
Abergel, Frédéric	Université Paris-Sud XI - Orsay	1986	3
Adamy, Karine	Université Paris-Sud XI - Orsay	2008	
Assila, R.	Université Paris-Sud XI - Orsay	1985	
Bouilhac, Isabelle	Université Paris-Sud XI - Orsay	1991	
Bousquet, Arthur	Indiana University	2015	
Brauner, Claude-Michel	Université Paris-Sud XI - Orsay	1972	70
Brefort, Bernadette	Université Paris-Sud XI - Orsay	1986	
Bruneau, Charles-Henri	Université Paris-Sud XI - Orsay	1989	14
Cabello, Jean	Université Paris-Sud XI - Orsay	1989	
Cabral, Marco	Indiana University	2002	
Calgare, Caterina	Université Paris-Sud XI - Orsay	1997	
Chattot, Jean-Jacques	Université Paris-Sud XI - Orsay	1979	3

Figura 2.6. Após a consulta pelo pesquisador, feita de forma simplificada ou avançada, é exibida uma tela com os seus dados presentes na plataforma Mathematics Genealogy Project, acesso em Abril de 2020.

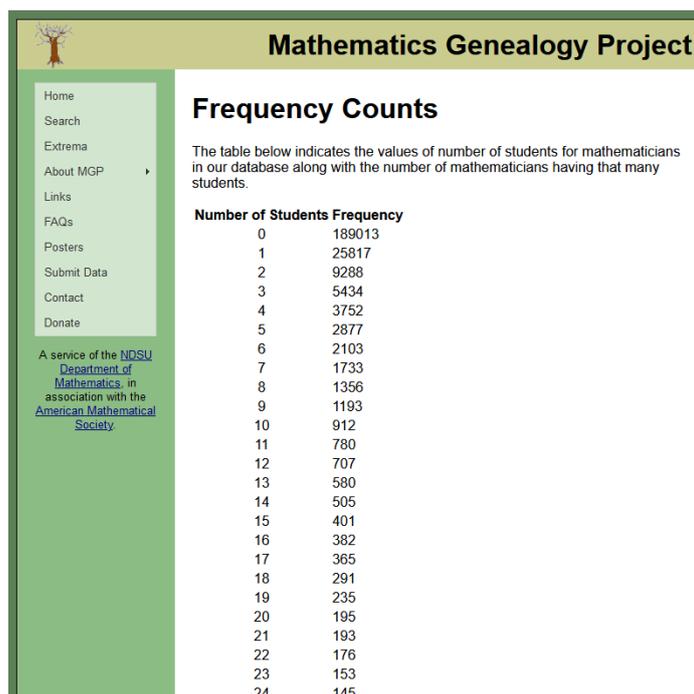


Figura 2.7. Lista com a contagem da frequência dos totais de orientações, presente na plataforma Mathematics Genealogy Project, acesso em Abril de 2020.

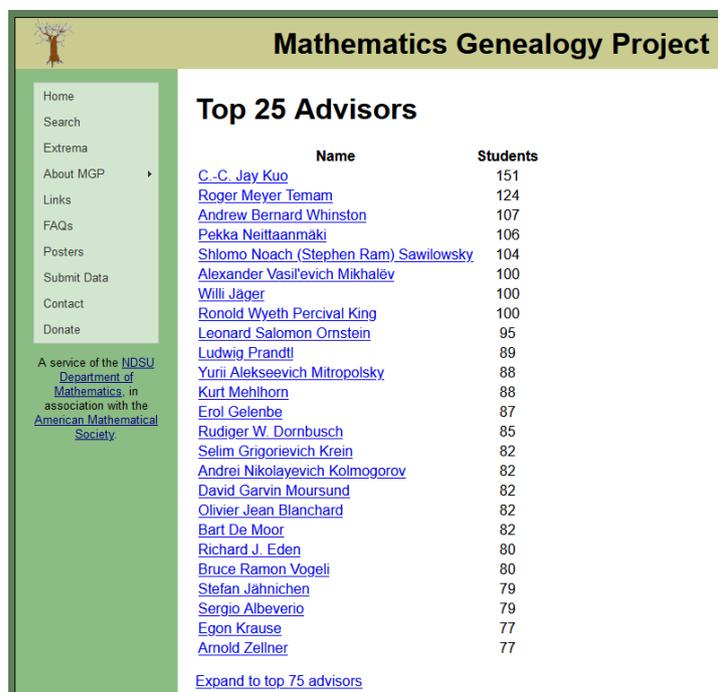
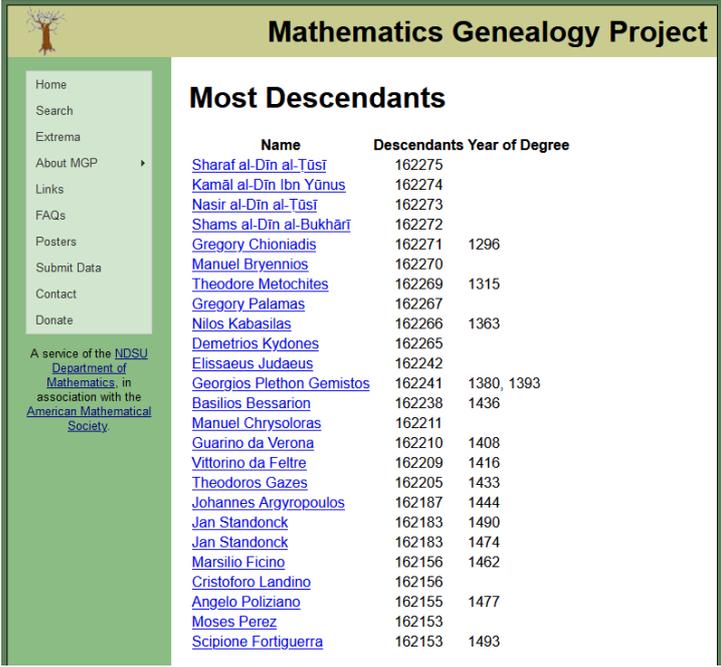


Figura 2.8. Lista dos 25 matemáticos com maior número de orientações (filhos acadêmicos), presentes na plataforma Mathematics Genealogy Project, acesso em Abril de 2020.



Mathematics Genealogy Project

Most Descendants

Name	Descendants	Year of Degree
Sharaf al-Din al-Tusi	162275	
Kamal al-Din Ibn Yunus	162274	
Nasir al-Din al-Tusi	162273	
Shams al-Din al-Bukhari	162272	
Gregory Chioniadis	162271	1296
Manuel Bryennios	162270	
Theodore Metochites	162269	1315
Gregory Palamas	162267	
Nilos Kabasilas	162266	1363
Demetrios Kydones	162265	
Elissaeus Judaeus	162242	
Georgios Plethon Gemistos	162241	1380, 1393
Basilios Bessarion	162238	1436
Manuel Chrysoloras	162211	
Guarino da Verona	162210	1408
Vittorino da Feltre	162209	1416
Theodoros Gazes	162205	1433
Johannes Argyropoulos	162187	1444
Jan Standonck	162183	1490
Jan Standonck	162183	1474
Marsilio Ficino	162156	1462
Cristoforo Landino	162156	
Angelo Poliziano	162155	1477
Moses Perez	162153	
Scipione Fortiguerra	162153	1493

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

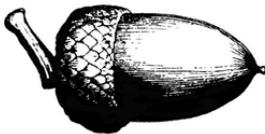
Figura 2.9. Lista dos 25 matemáticos com maior número de descendentes diretos e indiretos, presentes na plataforma Mathematics Genealogy Project, acesso em Abril de 2020.

2.2 The Academic Family Tree

A plataforma *The Academic Family Tree* originou-se de outra iniciativa similar, denominada *Neurotree*⁹. Assim, a versão atual dessa plataforma é uma generalização da plataforma original que, iniciada em janeiro de 2005, tinha como objetivo documentar as relações de orientação na área de Neurociência [David & Hayden, 2012], exibindo-as no formato de uma árvore genealógica. Logo após a criação da *Neurotree*, os seus mantenedores perceberam que a Neurociência possuía relações com outras áreas da ciência e que essas relações poderiam ser tão interessantes quanto as da própria Neurociência, fazendo com que novas áreas fossem incorporadas à plataforma que passou a ser denominada *The Academic Family Tree*.

A página inicial dessa plataforma, mostrada na Figura 2.10, relaciona todos os outros sítios que compõem a plataforma, agrupados por afinidade de suas áreas do conhecimento. Nessa página é possível que seja realizada uma busca geral na plataforma ou que se navegue por uma determinada área para realizar a busca por algum pesquisador específico daquela área escolhida. Em números, a plataforma possui hoje mais de 748 mil pesquisadores com mais de 687 mil conexões entre eles.

⁹<http://neurotree.org>, Acesso em Novembro de 2020



The Academic Family Tree
Building a single, interdisciplinary academic genealogy

Participating fields:

<p>Neuroscience Experimental Psychology Education Linguistics Primateology Communication Sciences</p> <p>Economics Public Policy Architecture City Planning Sociology Advertising Organizational Communication Ingestive Behavior Drosophila Genetics Fission Yeast Genetics Mycology and Fungal Genetics Plant Biology Microbiology Human Genetics Telomere and Telomerase Cell and Gene Therapy Cell Biology</p>	<p>History Philosophy Anthropology Political Science Law Literature Music Theology Mathematics Computer Science Engineering Astronomy Physics Crystallography Chemistry Physiology Anatomy Developmental Biology Evolutionary Biology Terrestrial Ecology Environmental Engineering Marine Ecology NEW! - Marine Mammal Science - NEW! Oceanography Meteorology Geoscience Geography</p>	<p>Browse - Search: <input type="text"/></p> <p>Is your research area missing from this list? Learn more about us, or contact us at admin at neurotree.org to find out how to set up a tree for your field.</p> <hr/> <p>What's new:</p> <p>Aug 17, 2019: Our Open Collective account is live! You can now make a contribution of any size to keep this site running. Read more here.</p> <p>Jan 1, 2019: New publication based on Academic Family Tree data: Intellectual Synthesis in Mentorship Determines Success in Academic Careers</p> <p>Feb 28, 2018: We've released a preprint on the synthesis of training by graduate and postdoctoral mentors: https://www.biorxiv.org/content/early/2018/02/28/273888. Your comments are welcome!</p> <p>More news ...</p>
--	--	--

Figura 2.10. Página inicial da plataforma The Academic Family Tree. Acesso em Abril de 2020.

Como resultado da busca por um pesquisador é exibida a sua árvore, conforme mostrada na Figura 2.11, no caso o Prof. José Palazzo Moreira de Oliveira da Universidade do Rio Grande do Sul. Nela estão presentes os seus ancestrais acadêmicos (seus orientadores diretos e os orientadores deles). A partir do pesquisador selecionado, é possível realizar interações com comandos específicos localizados do lado direito da árvore para expandir os nós horizontalmente e navegar verticalmente por meio deles, mostrando os seus descendentes diretos (orientados) e os descendentes acadêmicos deles. Para cada pesquisador é exibido também o nome da instituição em que atua.

Ao clicar sobre o nome de um pesquisador é aberta uma nova página contendo dados mais detalhados a respeito dele, que incluem, além do seu nome, do nome da instituição onde atua e de sua área de pesquisa, o endereço do seu sítio pessoal, a lista dos seus ascendentes (*parents*) e descendentes (*children*) diretos, e a lista de suas publicações, obtidas de fontes como PubMed¹⁰ e Google Scholar¹¹, conforme mostrado na Figura 2.12.

¹⁰<https://pubmed.ncbi.nlm.nih.gov/>

¹¹<https://scholar.google.com/>

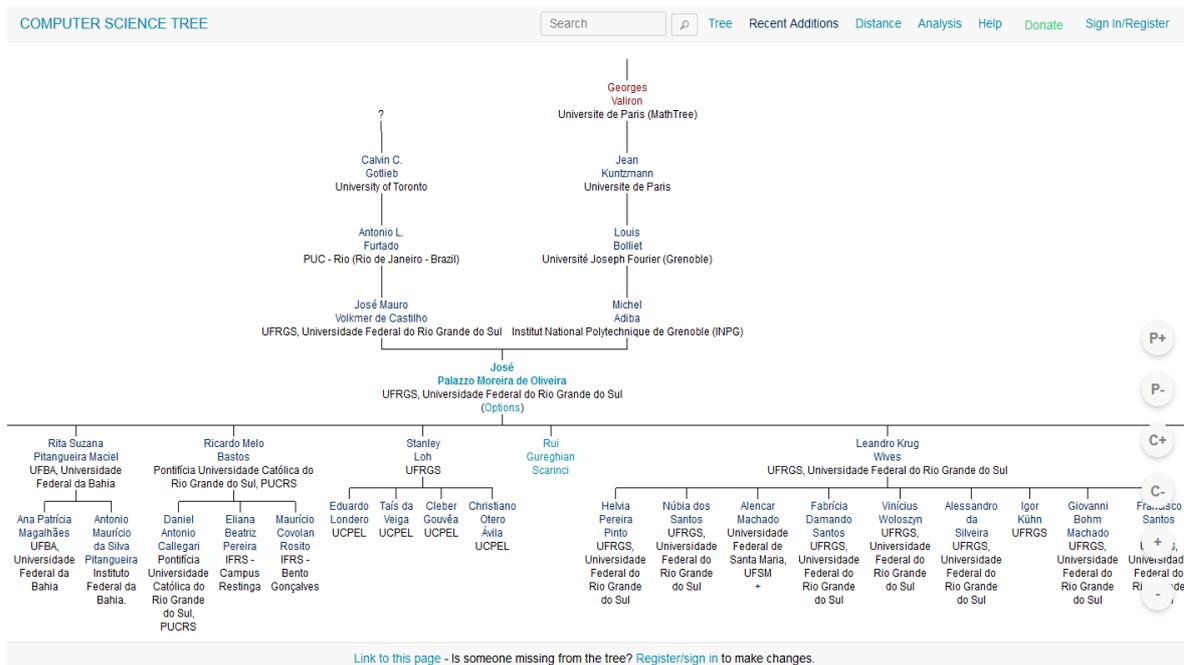


Figura 2.11. Árvore do Prof. José Palazzo Moreira de Oliveira da Universidade do Rio Grande do Sul, pesquisador da área de Ciência da Computação, que mostra os seus ascendentes e descendentes acadêmicos. Acesso em Abril de 2020.

The screenshot shows a web page for 'COMPUTER SCIENCE TREE' featuring a profile for José Palazzo Moreira de Oliveira, Dr. The profile includes a photo, a bio, and affiliations: 1969- Instituto de Informática UFRGS, Universidade Federal do Rio Grande do Sul; Area: Information Systems, Conceptual Modeling, Databases; Website: http://www.palazzo.pro.br/Wordpress/; Google: "José Palazzo Moreira de Oliveira". The bio states he is a full professor at UFRGS. Below the profile are sections for 'Parents' and 'Children'. The 'Parents' section lists José Mauro Volkmer de Castilho (grad student, 1973-1976, UFRGS) and Michel Adiba (grad student, 1982-1984, Institut National Polytechnique de Grenoble (INPG)). The 'Children' section lists Guilherme Medeiros Machado (grad student, 2014-), Luis Otávio Campos Alvares (grad student, 1979-1982, UFRGS), and Duncan Dubugras Alcoba (grad student, 1988-1985, UFRGS). A 'Publications' section lists several works, including 'Semiautomatic Process Model Verification Method based on Process Modeling Guidelines Proceedings of the 19th International Conference On Enterprise Information Systems, Porto, Portugal, April 26-29, 3: 274-281' and 'The collaboration network of the Brazilian Symposium on Databases - 30 editions of history Journal of the Brazilian Computer Society, Springer, 23: 10-1-10-16'.

Figura 2.12. Página com dados detalhados do pesquisador selecionado. Acesso em Abril de 2020.

The screenshot shows the 'COMPUTER SCIENCE TREE' website interface. At the top, there is a search bar and navigation links: Tree, Recent Additions, Distance, Analysis, Help, Donate, and Sign In/Register. The main section is titled 'Distance'. On the left, there are two input fields: 'Person 1:' with 'Alan M. Turing' and 'Person 2:' with 'Mirella M. Moro'. Below these are several checkboxes: 'Include research assistantships', 'Include postdoctoral positions', 'Include research scientist positions', 'Include collaborations', and 'Backward connections only (common ancestor)'. A 'Find Connection' button is located below the checkboxes. On the right side, the results are displayed under the heading 'Connection from Alan M. Turing to Mirella M. Moro:'. It shows '13 steps:' followed by a vertical list of names and their affiliations, connected by vertical lines. The list starts with Alan M. Turing (University of Cambridge) and ends with Mirella M. Moro (University of California, Los Angeles). Intermediate names include Alonzo Church, Michael O. Rabin, Dov Gabbay, Azriel Levy, Abraham Robinson, Paul Dienes, and Georges Valiron. Some entries are marked as '(grad student for)' or '(trained grad student)'.

Figura 2.13. Busca da distância entre dois pesquisadores com o propósito de mostrar o caminho de orientações entre eles. Acesso em Abril de 2020.

Com relação a métricas de desempenho acadêmico e outros dados de interesse, a plataforma *The Academic Family Tree* inclui também funções específicas para se determinar a distância entre dois pesquisadores, conforme mostrado no exemplo da Figura 2.13. Como pode ser observado, à esquerda são inseridos os nomes dos dois pesquisadores para os quais deseja-se obter a sua distância acadêmica e à direita é exibida a lista de pesquisadores que estabelecem essa relação. Além disso, a plataforma oferece várias outras facilidades que permitem analisar a relação entre pesquisadores ou compará-los entre si. Dentre as facilidades disponíveis é possível, por exemplo, verificar o relacionamento entre pesquisadores com perfis similares, buscar pesquisadores por instituição ou com colaboradores em comum, e avaliar a evolução de uma comunidade acadêmica em relação a outra existente na plataforma.

Por ser uma plataforma colaborativa, os pesquisadores interessados em ter seus perfis nela cadastrados podem conseguir isso de duas maneiras. A primeira, e mais comum, é por meio do seu próprio cadastramento na plataforma. Após o preenchimento de dados básicos, como *e-mail* e senha, é disponibilizada uma página em que o usuário pode inserir os seus próprios dados e, até mesmo, cadastrar perfis de outros pesquisadores, permitindo que esses pesquisadores também possuam um perfil na plataforma, mesmo que não tenham sido cadastrados por eles mesmos.

Por fim, podemos destacar algumas características relevantes dessa plataforma para a análise da genealogia acadêmica de um pesquisador, como a visualização gráfica

de sua árvore. A plataforma também disponibiliza uma página pessoal para cada pesquisador, capaz de efetuar o cruzamento de dados com outras plataformas, como, por exemplo, a PubMed. Outras características mais específicas como, por exemplo, informações e estatísticas de âmbito mais geral a respeito da plataforma e gráficos estatísticos sobre novos pesquisadores também são disponíveis.

2.3 Plataforma Acácia

A plataforma Acácia¹² tem o intuito de documentar as relações formais de orientação no contexto dos programas de pós-graduação brasileiros. Ao utilizar dados da plataforma Lattes do CNPq é realizado um tratamento para obter grafos de genealogia acadêmica, em que cada vértice representa um pesquisador e cada aresta uma relação de orientação concluída entre orientador e orientado. Mantida pelo grupo de pesquisa liderado pelo Prof. Jesús Mena-Chalco da UFABC, o lançamento da plataforma Acácia ocorreu em Abril de 2018. Detalhes do seu algoritmo para geração de grafos de genealogia acadêmica foram apresentados por Damaceno et al. [2017, 2019].

Um exemplo da página de um pesquisador gerada pela plataforma Acácia é ilustrado na Figura 2.14, que mostra o perfil do Prof. José Palazzo Moreira de Oliveira. Como pode ser visto, a plataforma foca principalmente nos dados de formação acadêmica e orientação do pesquisador, coletados da plataforma Lattes. Outro ponto a ser observado é que pode ser feita uma navegação simples envolvendo os descendentes e ascendentes do pesquisador, porém sem qualquer representação gráfica a respeito dessas relações de orientação. Entretanto, a plataforma Acácia não provê qualquer informação a respeito de coorientações, ou mesmo sobre ascendentes ou descendentes, de pesquisadores que não tenham sido efetivamente cadastrados na plataforma Lattes.

Por outro lado, a plataforma Acácia apresenta vários dados sobre o pesquisador, como grande área e área de atuação¹³, titulação, instituição, seus ascendentes e descendentes diretos. Além disso, ela provê sete métricas específicas sobre a genealogia acadêmica do pesquisador, que são: *Descendência* (número de descendentes diretos), *Índice Genealógico* (maior número de descendentes que possuem a mesma quantidade de descendentes), *Fecundidade* (quantidade de descendentes diretos), *Fertilidade* (quantidade de descendentes diretos que possuem pelo menos um descendente), *Gerações* (número de gerações ou altura da árvore), *Relações* (quantidade de conexões ou arestas na descendência do acadêmico) e *Primos* (quantidade de acadêmicos que

¹²<http://plataforma-acacia.org>, acesso em Abril de 2020

¹³Definidas conforme padrão estabelecido pelo CNPq.

possuem avós em comum e pais diferentes). Todas essas informações são compiladas a partir do currículo do pesquisador escolhido.

Plataforma Acácia
Genealogia Acadêmica do Brasil

Pesquise mestres e doutores atuantes no Brasil

Acadêmicos: 1.272.590
Relações de orientação: 1.404.109

Home César Lattes Análise de Áreas Quem Somos Perguntas Frequentes Sobre

Plataforma Acácia
Genealogia Acadêmica do Brasil

Pesquise mestres e doutores atuantes no Brasil

Acadêmicos: 1.272.590
Relações de orientação: 1.404.109

Digitte o nome do pesquisador

José Palazzo Moreira De Oliveira

Ascendentes Descendentes

Grande Área¹: Ciências Exatas E Da Terra
Área¹: Ciência Da Computação
Instituição¹: Universidade Federal Do Rio Grande Do Sul

Descendência (Ds)¹: 395
Índice Genealógico (IG)¹: 8
Fecundidade (Fc)¹: 71
Fertilidade (Ft)¹: 16
Gerações (G)¹: 3
Relações (R)¹: 397
Primos (Pr)¹: 0

Data dos dados¹: 13/6/2019

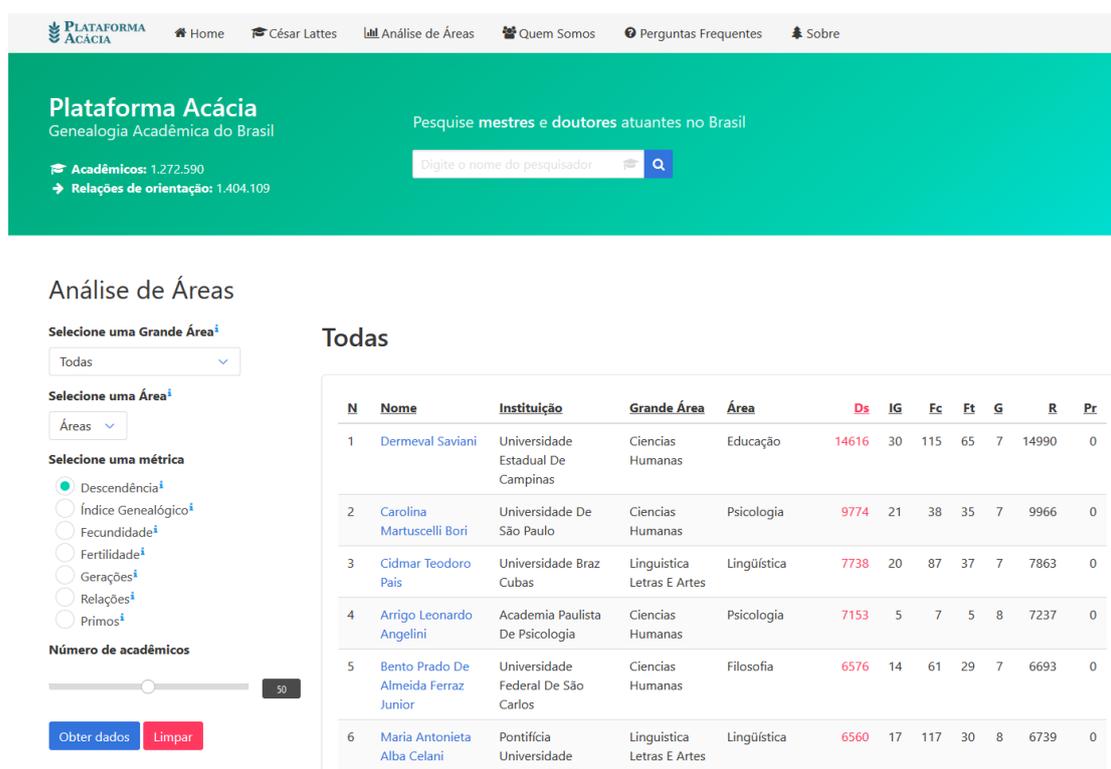
N	Nome	Orientações	Ds	IG	Fc	Ft	G	R	Pr
1	Adriana Jouris	M 2011	0	0	0	0	0	0	101
2	Alencar Machado	D 2015	2	0	2	0	1	2	277
3	Aliomar Mariano Rêgo	M 1990	0	0	0	0	0	0	101
4	Ana Carla Macedo Da Silva	M 2002	0	0	0	0	0	0	168
5	Ana Marilza Pernas Fleischmann	D 2012	3	0	3	0	1	3	157
6	Andre Luis Costa De Oliveira	M 1998	0	0	0	0	0	0	101

Figura 2.14. Exibição do perfil de um pesquisador na plataforma Acácia. Acesso em Abril de 2020.

Essa plataforma também apresenta uma análise por áreas conforme pode ser mostrado pela Figura 2.15. Nesse contexto podem ser realizadas buscas por grande área e refinadas de acordo com a área ou quantidade de registros a serem retornados, havendo necessidade de escolher qual a métrica dentre as sete citadas anteriormente que deseja ser evidenciada. No caso específico da Figura 2.15 foi realizada uma busca envolvendo todas as grandes áreas, tendo sido selecionados os 50 pesquisadores com maior quantidade de descendentes diretos ou indiretos.

2.4 The Gold Tree

A plataforma *The Gold Tree* tem por finalidade gerar uma visualização da genealogia acadêmica dos pesquisadores brasileiros que possuem seus currículos cadastrados na plataforma Lattes do CNPq [Madeira et al., 2020]. Segundo dados disponíveis no próprio sítio da plataforma, foram utilizados dados de mais de 570 mil dissertações e teses para realizar o cruzamento desses dados com os currículos. A plataforma foi desenvolvida pelo Grupo de Pesquisa em Gerenciamento de Informações da Universidade Federal do Rio Grande, tendo sido parcialmente financiada pela Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul.



Plataforma Acácia
Genealogia Acadêmica do Brasil

Pesquise mestres e doutores atuantes no Brasil

Acadêmicos: 1.272.590
Relações de orientação: 1.404.109

Análise de Áreas

Selecione uma Grande Área¹
Todas

Selecione uma Área¹
Áreas

Selecione uma métrica

- Descendência¹
- Índice Genealógico¹
- Fecundidade¹
- Fertilidade¹
- Gerações¹
- Relações¹
- Primos¹

Número de acadêmicos
50

Obter dados Limpar

Todas

N	Nome	Instituição	Grande Área	Área	Ds	IG	Fc	Ft	G	R	Pr
1	Dermeval Saviani	Universidade Estadual De Campinas	Ciencias Humanas	Educação	14616	30	115	65	7	14990	0
2	Carolina Martuscelli Bori	Universidade De São Paulo	Ciencias Humanas	Psicologia	9774	21	38	35	7	9966	0
3	Cidmar Teodoro Pais	Universidade Braz Cubas	Linguística Letras E Artes	Linguística	7738	20	87	37	7	7863	0
4	Arrigo Leonardo Angelini	Academia Paulista De Psicologia	Ciencias Humanas	Psicologia	7153	5	7	5	8	7237	0
5	Bento Prado De Almeida Ferraz Junior	Universidade Federal De São Carlos	Ciencias Humanas	Filosofia	6576	14	61	29	7	6693	0
6	Maria Antonieta Alba Celani	Pontifícia Universidade Católica De São	Linguística Letras E Artes	Linguística	6560	17	117	30	8	6739	0

Figura 2.15. Exibição da Análise de Áreas disponível na plataforma Acácia, que mostra os pesquisadores listados em ordem decrescente em relação à quantidade de descendentes diretos e indiretos. Acesso em Abril de 2020.

A plataforma *The Gold Tree* inclui em sua página inicial um campo de busca pelo nome do pesquisador. Ao realizar uma busca, o usuário pode informar, além do nome do pesquisador, a quantidade de níveis que se deseja visualizar em sua árvore. Conforme pode ser evidenciado na Figura 2.16, essa plataforma exhibe apenas a árvore de orientações do pesquisador, seus níveis, ano de obtenção do título e o grau obtido. Ao clicar no tipo de orientação de um pesquisador, indicado pela letra D (doutorado) ou M (mestrado) seguida pelo ano de conclusão, é aberta uma janela com dados específicos sobre aquela orientação, tais como nome do orientado, nome do orientador, título da dissertação ou tese, e ano de conclusão, conforme ilustrado pela Figura 2.17. Porém, a formatação vertical de suas árvores torna a visualização da genealogia acadêmica de um pesquisador pouco natural. Apesar de possibilitar uma forma simples de navegação, a plataforma possui algumas limitações, já que fornece poucos dados sobre o pesquisador, não provê métricas que indiquem o impacto de sua genealogia acadêmica e não possibilita o caminhar pelas árvores de forma interativa, pois apenas exhibe as árvores em um formato vertical limitado ao filtro de profundidade escolhido previamente.

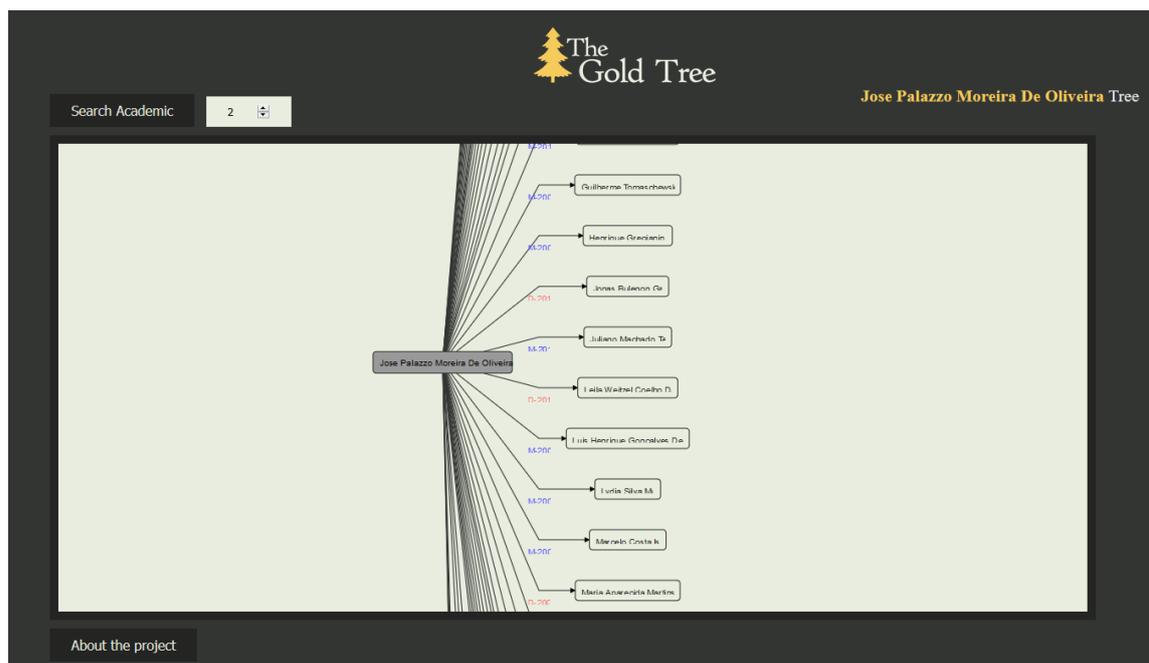


Figura 2.16. Visualização de dois níveis da árvore de um pesquisador na plataforma The Gold Tree. Acesso em Abril de 2020.

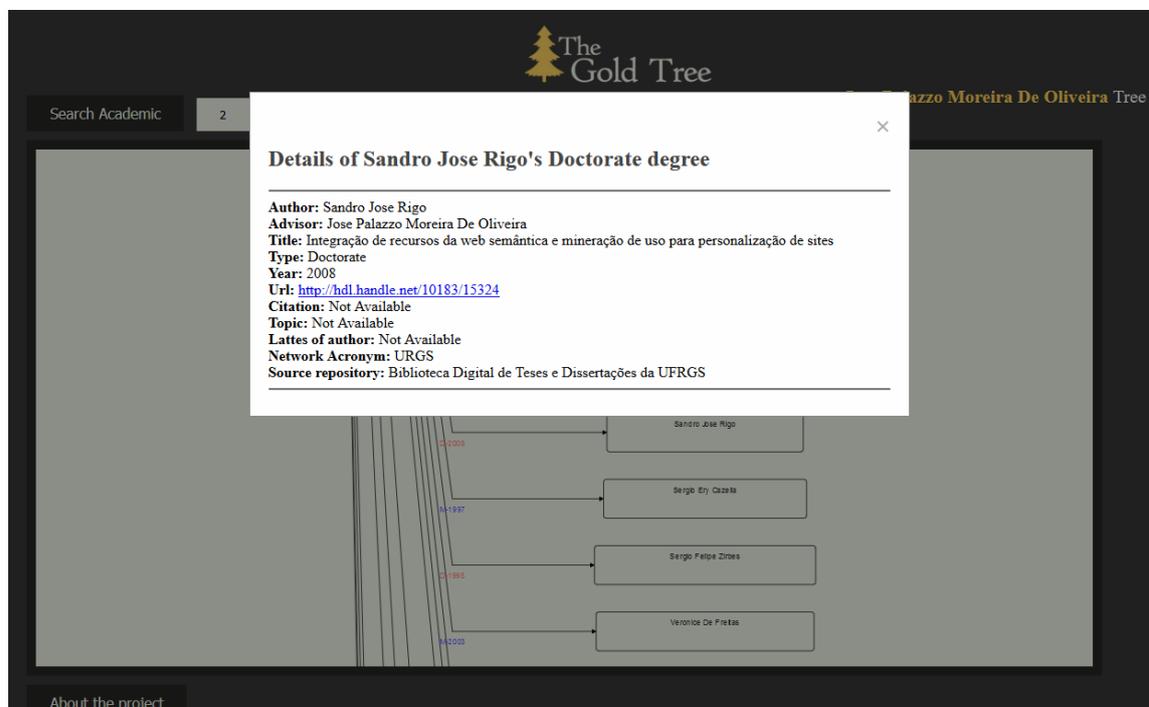


Figura 2.17. Visualização dos dados referentes à orientação de doutorado de um descendente acadêmico escolhido aleatoriamente. Acesso em Abril de 2020.

2.5 Considerações Finais

As plataformas apresentadas neste capítulo podem ser comparadas principalmente em relação aos dados utilizados para geração das árvores genealógicas acadêmicas. As plataformas *Acácia* e *The Gold Tree* utilizam dados referentes à formação acadêmica no Brasil, coletados da plataforma Lattes do CNPq. A plataforma *Mathematics Genealogy Project* foca nos pesquisadores de uma área específica, a Matemática, abrangendo pesquisadores de vários países. Finalmente, a plataforma *The Academic Family Tree* provê um grande arcabouço que cobre diversas outras áreas do conhecimento de forma integrada, abrangendo também pesquisadores de diversos países.

No que diz a respeito à forma de interação e visualização dos dados, a plataforma *The Academic Family Tree* é a que provê mais recursos aos seus usuários. Além de possibilitar interatividade na forma de uma árvore genealógica, ela também provê uma série de métricas que podem ser usadas para comparação dos pesquisadores, sendo também capaz de extrair dados de outras fontes, como Google Acadêmico e PubMed. A plataforma *Acácia* também possui métricas e indicadores de produção por área e por pesquisador, mas carece de uma melhor interatividade para visualização da genealógica acadêmica dos pesquisadores, já que o caminhamento se dá apenas do orientador para os seus orientados. A plataforma *The Gold Tree*, por outro lado, conta com uma interação na forma de um grafo horizontal e provê apenas alguns dados gerais sobre os pesquisadores. Por fim, a plataforma *Mathematics Genealogy Project* apresenta os dados dos pesquisadores somente em formato textual, o que limita bastante a visualização de suas linhagens acadêmicas.

As características comuns de cada plataforma discutidas neste capítulo estão resumizadas na Tabela 2.1. A partir dessa tabela é possível observar que a forma de representação da genealogia varia entre tabular e gráfica. Já em relação às métricas, a plataforma *The Academic Family Tree* é a que provê o maior número de métricas, possibilitando uma análise mais detalhada acerca dos seus dados. Por outro lado, a plataforma *The Gold Tree* não provê qualquer métrica, limitando a análise da genealogia acadêmica apenas ao processo de orientação. A forma de visualização dos dados varia entre textual ou gráfica (barras e linhas), sendo que a plataforma *The Academic Family Tree* é a que se mostra mais versátil neste quesito.

No entanto, as quatro plataformas apresentam em comum limitações no que se refere ao modo de interação do usuário. O foco de cada plataforma também é variado, sendo que tanto a plataforma *Acácia* quanto a *The Gold Tree* lidam com dados provenientes de um repositório nacional, no caso a plataforma Lattes do CNPq, enquanto que as plataformas *Mathematics Genealogy Project* e *The Academic Family*

Plataforma	Representação da Genealogia (Tabular ou Gráfica)	Quantidade de Métricas Disponíveis	Forma de Visualização dos Dados (Textual ou Gráfica)	Foco da Plataforma
Mathematics Genealogy Project	Tabular	2	Textual	Mundial
The Academic Family Tree	Ambas	15	Ambas	Mundial
Acácia	Tabular	7	Textual	Nacional
The Gold Tree	Gráfica	0	Gráfica	Nacional

Tabela 2.1. Tabela comparativa dos principais recursos presentes nas quatro plataformas abordadas.

Tree registram dados de instituições de diversos países do mundo para compor o seu próprio repositório, sendo em geral alimentadas pelos próprios pesquisadores ou por terceiros, por exemplo, instituições interessadas em manter as árvores de seus docentes ou pesquisadores.

Capítulo 3

Descrição Funcional da Plataforma Proposta

Conforme mencionado anteriormente, uma importante contribuição originada do estudo pioneiro sobre a genealogia acadêmica brasileira realizado por Dores [2017] foi a implementação de uma primeira versão da plataforma *Árvore da Ciência*, então denominada *Science Tree*. Entretanto, a interface de usuário proposta para aquela versão da plataforma era muito limitada em termos funcionais e não oferecia maiores facilidades para exploração das árvores genealógicas acadêmicas. Apesar disso, ela se mostrou bastante útil não só para a validação do processo de geração dessas árvores, mas também para a identificação dos principais aspectos funcionais que deveriam nortear o desenvolvimento de uma nova versão da plataforma, de modo a propiciar aos seus usuários não apenas uma interface mais funcional, mas também mais abrangente a fim de possibilitar uma ampla exploração da genealogia acadêmica brasileira a partir de dados coletados da plataforma Lattes.

A Figura 3.1 descreve o fluxo de dados proposto para geração das árvores genealógicas acadêmicas na nova versão da plataforma *Árvore da Ciência*. Como pode ser observado, a partir do nome de um pesquisador, é feita uma consulta ao repositório de currículos coletados da plataforma Lattes para extração dos dados necessários para geração de sua respectiva árvore genealógica acadêmica. Extraídos os dados, é gerada a árvore do pesquisador com suas respectivas métricas. Finalmente, a árvore gerada pode ser visualizada juntamente com dados específicos sobre o pesquisador, como nome, formação acadêmica e instituição de vínculo, métricas que expressam diferentes aspectos de sua genealogia acadêmica e um gráfico temporal de suas orientações concluídas conforme registradas na plataforma Lattes. Vale ressaltar que todos esses dados são mantidos nos repositórios da própria plataforma *Árvore da Ciência* e ficam disponíveis publicamente para acesso pelos seus usuários por meio da World Wide Web.

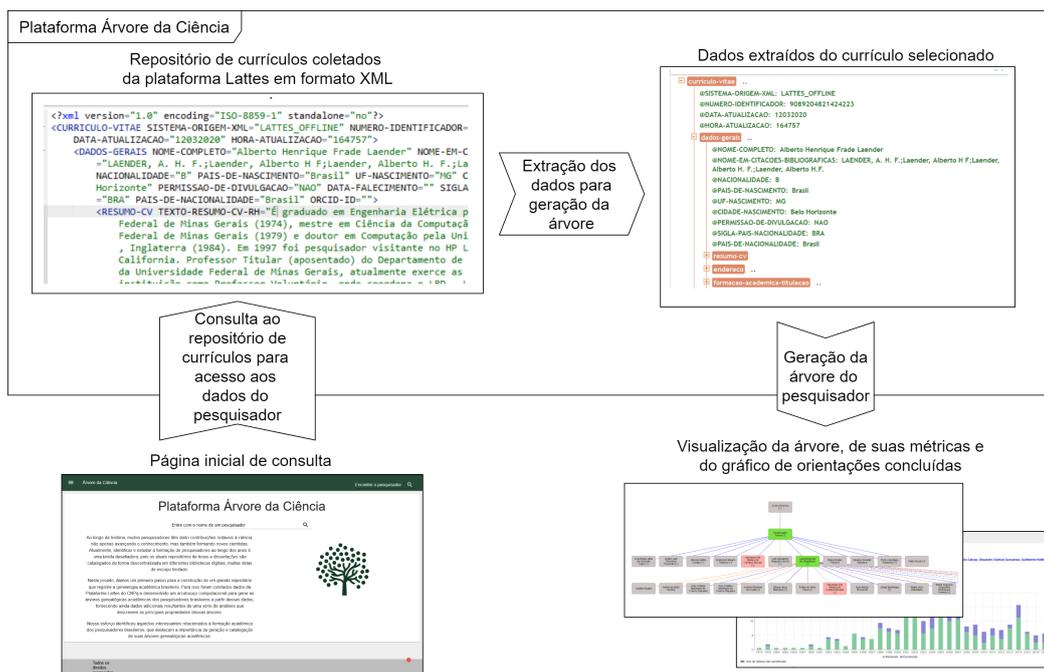


Figura 3.1. Fluxo de dados da nova versão da plataforma Árvore da Ciência.

3.1 Arquitetura Geral da Plataforma

Esta seção descreve detalhes gerais da arquitetura da plataforma Árvore da Ciência. Uma vez que trata-se de um sistema computacional constituído por vários componentes, ela carece de uma definição estrutural de cada um desses componentes e de como eles se relacionam entre si. A Figura 3.2 apresenta esses componentes e descreve como eles se comunicam entre si.

Conforme mostrado, a plataforma possui duas aplicações principais: a aplicação *Front-end* e a aplicação *Back-end*. A aplicação *Front-end* estabelece a interface com os usuários, estruturando os dados que possibilitarão a esses usuários interagir com a plataforma, por exemplo, para que possam navegar pelas árvores genealógicas acadêmicas visualizando os dados dos respectivos pesquisadores, bem como para que possam ter acesso aos valores de determinadas métricas que expressam a estrutura das árvores e caracterizam o desempenho acadêmico desses pesquisadores. Já a aplicação *Back-end* intercepta as solicitações da aplicação *Front-end* para realizar buscas no Repositório de Dados da plataforma, de modo a calcular o valor das métricas genealógicas e gerar outros dados de interesse dos usuários finais, como o gráfico de orientações concluídas.

Os currículos em formato XML, coletados *offline* da Plataforma Lattes, são re-

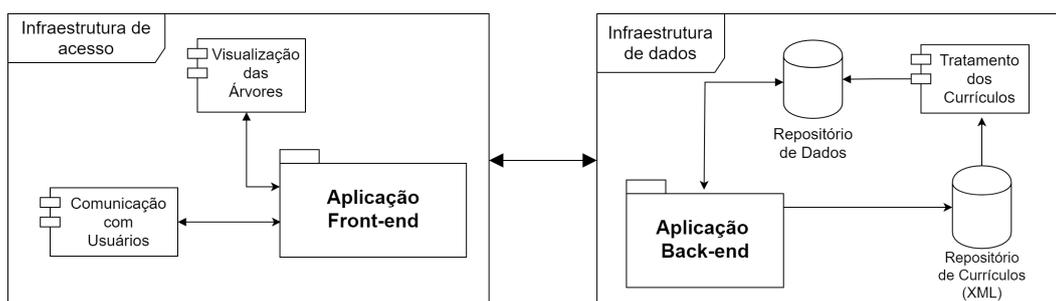


Figura 3.2. Esquema de comunicação entre os componentes da plataforma Árvore da Ciência.

presentados pelo componente Repositório de Currículos (XML)¹. Esses currículos são usados pelo módulo de Tratamento dos Currículos para gerar, utilizando um formato intermediário e padronizado, o Repositório de Dados que armazena os dados pessoais básicos e aqueles que descrevem as relações de orientação dos pesquisadores com título de Doutor cadastrados na plataforma Lattes do CNPq e que são usados para a construção de suas respectivas árvores genealógicas acadêmicas de acordo com o algoritmo proposto por Dores et al. [2017].

Após o armazenamento dos dados relevantes extraídos dos currículos coletados, eles são consumidos diretamente pela aplicação *Back-end* que interpreta as consultas que o usuário realiza por meio da aplicação *Front-end*. Essa é a forma de o usuário comunicar-se com a plataforma para ter acesso aos dados de um pesquisador de seu interesse e visualizar a sua árvore genealógica acadêmica. A aplicação *Back-end* é ainda capaz de acessar diretamente o Repositório de Currículos (XML) para complementar dados sobre o pesquisador em questão conforme solicitação do usuário, como, por exemplo, o resumo do seu currículo por ele informado.

3.2 Interface de Usuário

A interface de usuário da plataforma Árvore da Ciência foi concebida a partir de um *design* minimalista, com poucos elementos visuais e textos objetivos. Essas características auxiliam no entendimento dos dados apresentados aos usuários, além de diminuir o tempo de navegação. Além disso, como a plataforma disponibiliza várias métricas que permitem quantificar diferentes aspectos da genealogia acadêmica dos pesquisadores, é importante que a sua interface permita aos seus usuários obter uma visão mais

¹Os dados utilizados para a geração da versão da plataforma desenvolvida nesta dissertação foram obtidos a partir de coletas realizadas pela ferramenta Lattes DataXplorer [Dias, 2016].

analítica do desempenho de cada um deles em relação ao papel que exercem como orientadores.

Assim, a interface foi implementada empregando uma linguagem e bibliotecas de programação atuais que serão descritas na próxima seção. A forma de acesso a ela se dá por meio de um dispositivo eletrônico (computador, *smartphone* ou *tablet*) conectado à Internet e com acesso a um navegador. A plataforma está hospedada no sítio <http://sciencetree.net> e não exige qualquer cadastramento prévio de seus usuários para a sua utilização.

A genealogia acadêmica vem se tornando um foco de estudo cada vez mais popular e a sua representação é comumente realizada em formato de árvores como exemplificado por algumas das plataformas apresentadas no capítulo anterior. Os trabalhos nessa área têm como objetivo coletar e estruturar dados de um repositório de teses e dissertações para desenvolver ferramentas amigáveis que possibilitem a visualização e a exploração das árvores geradas [Liu et al., 2017; Sheng et al., 2019]. Neste contexto, uma primeira versão da plataforma *Árvore da Ciência* foi implementada a partir do trabalho de Dores [2017]. Essa primeira versão possuía uma interface bem simples e tinha como objetivo validar o algoritmo por ele proposto para geração das árvores genealógicas acadêmicas dos pesquisadores brasileiros a partir de dados coletados da plataforma Lattes.

Duas telas da então plataforma *Science Tree* são ilustradas nas Figuras 3.3 e 3.4. A Figura 3.3 mostra a página de um pesquisador com dados específicos sobre a sua formação acadêmica, como título da Tese de Doutorado e da Dissertação de Mestrado, nome dos seus orientadores, ano de obtenção dos respectivos títulos acadêmicos, denominação de suas respectivas áreas e sub-áreas do conhecimento segundo a classificação do CNPq e, por fim, os anos de obtenção dos respectivos títulos. Por outro lado, a Figura 3.4 mostra uma interface com poucos recursos para uma interação visual e que permitia aos seus usuários realizar apenas uma simples navegação pelas árvores geradas, isto é, sem a possibilidade de expandir na mesma tela os vários níveis da árvore genealógica acadêmica do pesquisador. Dessa forma, a cada interação do usuário, uma nova árvore era gerada a partir de um dos nodos da árvore precedente, caso o pesquisador correspondente tivesse concluído alguma orientação.

Assim, a plataforma *Science Tree* carecia de maiores facilidades que possibilitassem uma exploração mais ampla dos dados disponíveis. Por exemplo, o uso de uma métrica como *número de orientações por ano* poderia auxiliar o usuário a estabelecer uma correlação direta com o impacto do pesquisador selecionado na formação de novos pesquisadores em uma determinada área do conhecimento. Essa é uma correlação importante, geralmente utilizada pelas agências de fomento à pesquisa como CAPES e CNPq, pois permite não só avaliar a produtividade de um pesquisador, como também

Home Search Contact About

Alberto Henrique Frade Laender

Institution: Universidade Federal De Minas Gerais
Webpage: Unknown

Tree Information

Thesis Title: A Framework For The Definition And Manipulation Of Database Views By End Users
Supervisor: Peter M Stocker
Awarded Degree: PhD
Institution: University Of East Anglia
Research Area: ciencias_exatas_e_da_terra
Research Sub-Area: & ciência da computação
Year: 1984
[Edit this profile](#)
[Download Alberto Henrique Frade Laender's tree.](#)

Thesis Title: Especificação E Projeto De Um Sistema Integrado Para Aplicação Correção E Análise De Provas
Supervisor: Ivan Moura Campos
Awarded Degree: Master
Institution: Universidade Federal De Minas Gerais
Research Area: ciencias_exatas_e_da_terra
Research Sub-Area: & ciência da computação
Year: 1979

Figura 3.3. Visualização da página de dados específicos de um pesquisador na plataforma *Science Tree*.

Home Search Contact About

Alberto Henrique Frade Laender

Institution: Universidade Federal De Minas Gerais
Webpage: Unknown

Tree Information

Supervised Students: 69

The diagram shows a central node for Alberto Henrique Frade Laender, with lines connecting to a large number of other nodes representing supervised students. Above the central node, two nodes represent supervisors: Peter Stocker and Ivan Campos.

Figura 3.4. Visualização da árvore de um pesquisador na plataforma *Science Tree* em que é possível identificar o pesquisador em destaque, bem como seus orientadores e orientados.

direcionar recursos a um programa de pós-graduação ou a um grupo de pesquisa no país. Apesar dessas limitações, a plataforma *Science Tree* mostrou-se uma ferramenta importante para induzir a solução de problemas relacionados tanto ao tratamento dos currículos coletados como à ambiguidade inerente aos dados disponíveis nesses currículos ao gerar as árvores genealógicas acadêmicas dos pesquisadores brasileiros.

Assim, como parte desta dissertação, foi desenvolvida uma nova interface de usuário com diversas melhorias visando não só explorar de forma mais ampla as árvores geradas a partir de dados extraídos de currículos coletados da plataforma Lattes do CNPq, mas também proporcionar aos seus usuários diversas métricas e dados adicionais para uma análise mais ampla do desempenho dos pesquisadores em termos de orientação acadêmica. Para isso, utilizou-se o arcabouço computacional já existente em nosso grupo de pesquisa e alguns dos componentes de software desenvolvidos para tratamento de dados na plataforma *Science Tree*.

Dentro desse contexto, o novo ambiente desenvolvido possibilita uma exploração mais natural das árvores, incluindo novos recursos de interação e visualização que permitem não só expandir ou ocultar determinadas partes (galhos) delas, como também exibir detalhes específicos como métricas topológicas ou dados referentes à carreira acadêmica dos pesquisadores, de modo a prover uma melhor experiência ao usuário ao explorar as árvores de seu interesse.

O ambiente proposto possui uma tela inicial (Figura 3.5) com um campo de busca no canto superior esquerdo por meio do qual o nome de um pesquisador de interesse pode ser inserido. Ao realizar uma busca, os nomes dos pesquisadores que satisfazem total ou parcialmente o nome inserido são recuperados e exibidos em seguida. Clicando em um desses nomes, a árvore do pesquisador correspondente é carregada na tela. No caso específico do exemplo da Figura 3.5, um único nome foi retornado, o do pesquisador Luciano Antonio Digiampietri.

A Figura 3.6 ilustra a árvore genealógica acadêmica do pesquisador Luciano Antonio Digiampietri, gerada ao clicar o seu nome no resultado da busca ilustrada na Figura 3.5. A cor das arestas define o nível das respectivas orientações, laranja (mestrado) e azul (doutorado). Assim, no caso deste pesquisador, todas as suas orientações são de mestrado. Na sua árvore é possível observar ainda que ele foi orientado pelo Prof. João Carlos Setubal e coorientado pela Profa. Claudia Maria Bauzer Medeiros (arestas tracejadas representam coorientações). As abas abaixo, conforme acionadas, apresentam dados específicos sobre o pesquisador extraídos do seu Currículo Lattes ou métricas referentes à sua árvore mostrada acima. Também é possível realizar outras ações como movimentar os nós na horizontal e clicar em um nó para que os dados mostrados nas abas abaixo sejam alterados de acordo com o pesquisador selecionado.

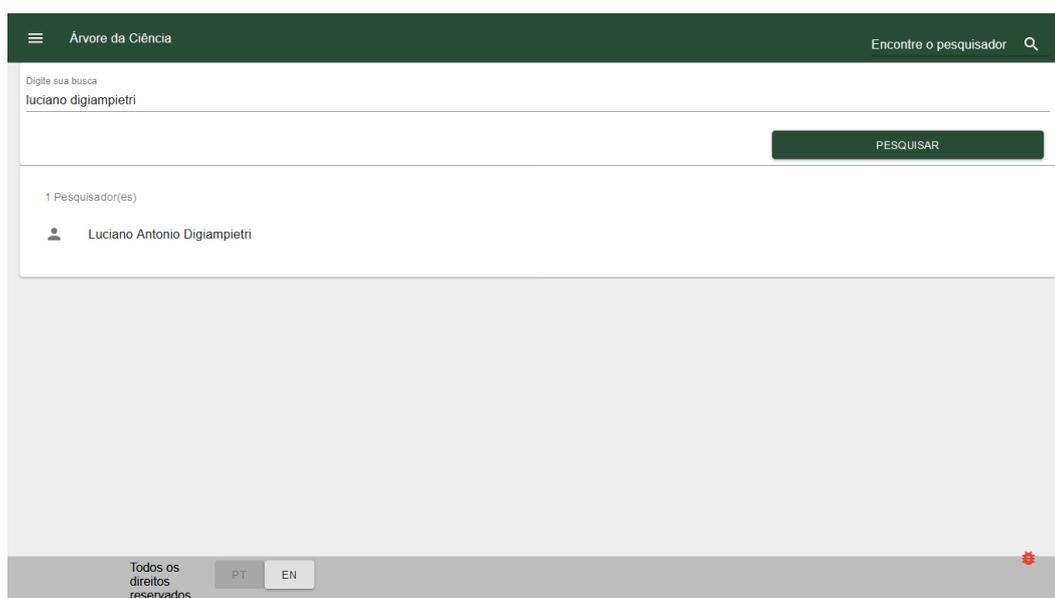


Figura 3.5. Busca pelo nome abreviado do pesquisador Luciano Antonio Digiampietri que retorna um resultado único.

O caminhamento, que é a expansão dos nodos de uma árvore, ocorre quando é realizado um duplo clique em um nodo que possui essa capacidade, simbolizada pelo sinal de soma entre colchetes ([+]). Assim, por exemplo, ao realizar um duplo clique no nodo do Prof. João Carlos Setubal na Figura 3.6, obtém-se como resultado a árvore da Figura 3.7 que mostra que ele possui outros descendentes acadêmicos além do Luciano Antônio Digiampietri. Note que o nodo expandido tem a sua cor alterada para indicar a expansão ocorrida.

A Figura 3.7 mostra também a aba de métricas do pesquisador Luciano Antônio Digiampietri. Nela é possível observar cinco das seis métricas disponíveis. A *profundidade* de sua árvore é igual a um, indicando que seus orientados ainda não formaram novos descendentes. Já o *número de descendentes diretos e indiretos* indica o tamanho da árvore do pesquisador que, no caso do pesquisador Luciano Antônio Digiampietri, é igual à sua *largura* (12), uma vez que nenhum de seus descendentes possui outros orientados.

Finalmente, a última métrica corresponde ao *número médio de orientações concluídas por ano* e indica que o pesquisador Luciano Antonio Digiampietri orientou em média, de acordo com dados então coletados da plataforma Lattes, dois alunos de mestrado por ano. Os valores absolutos do número de orientações concluídas por ano, bem como o nível de formação (mestrado ou doutorado) são sumarizados no histograma mostrado logo abaixo, sendo a cor verde utilizada para indicar a formação no nível de mestrado e a cor azul a formação no nível de doutorado. Note que à

medida que o usuário desloca o cursor sobre o histograma, os valores correspondentes ao número total de orientações de mestrado e doutorado correspondentes a cada ano são realçados. Uma descrição detalhada das métricas consideradas pode ser obtida pelo *link* disponível abaixo do histograma.

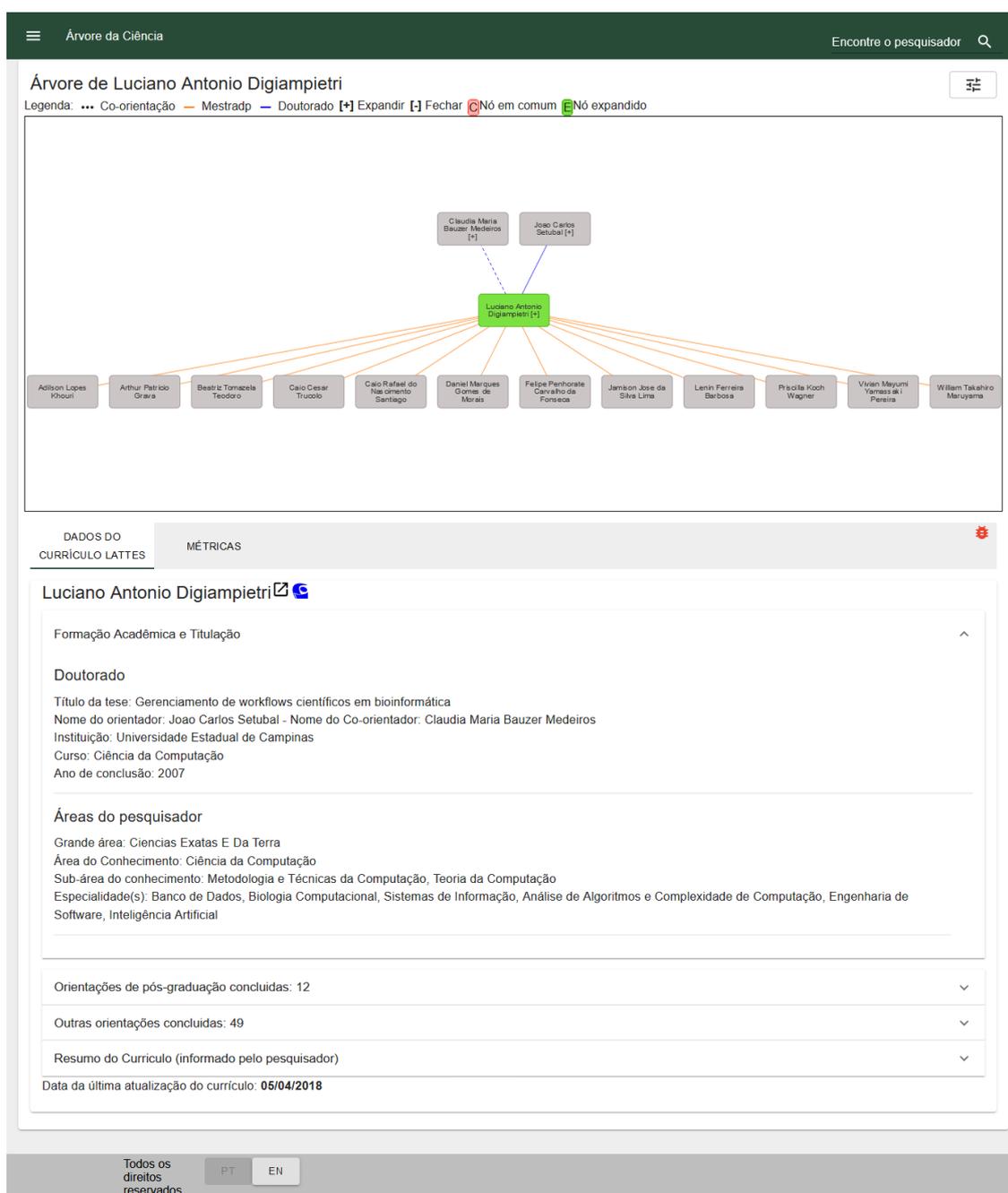


Figura 3.6. Interface da plataforma Árvore da Ciência mostrando a árvore do pesquisador Luciano Antonio Digiampietri e a aba de dados extraídos do seu Currículo Lattes.

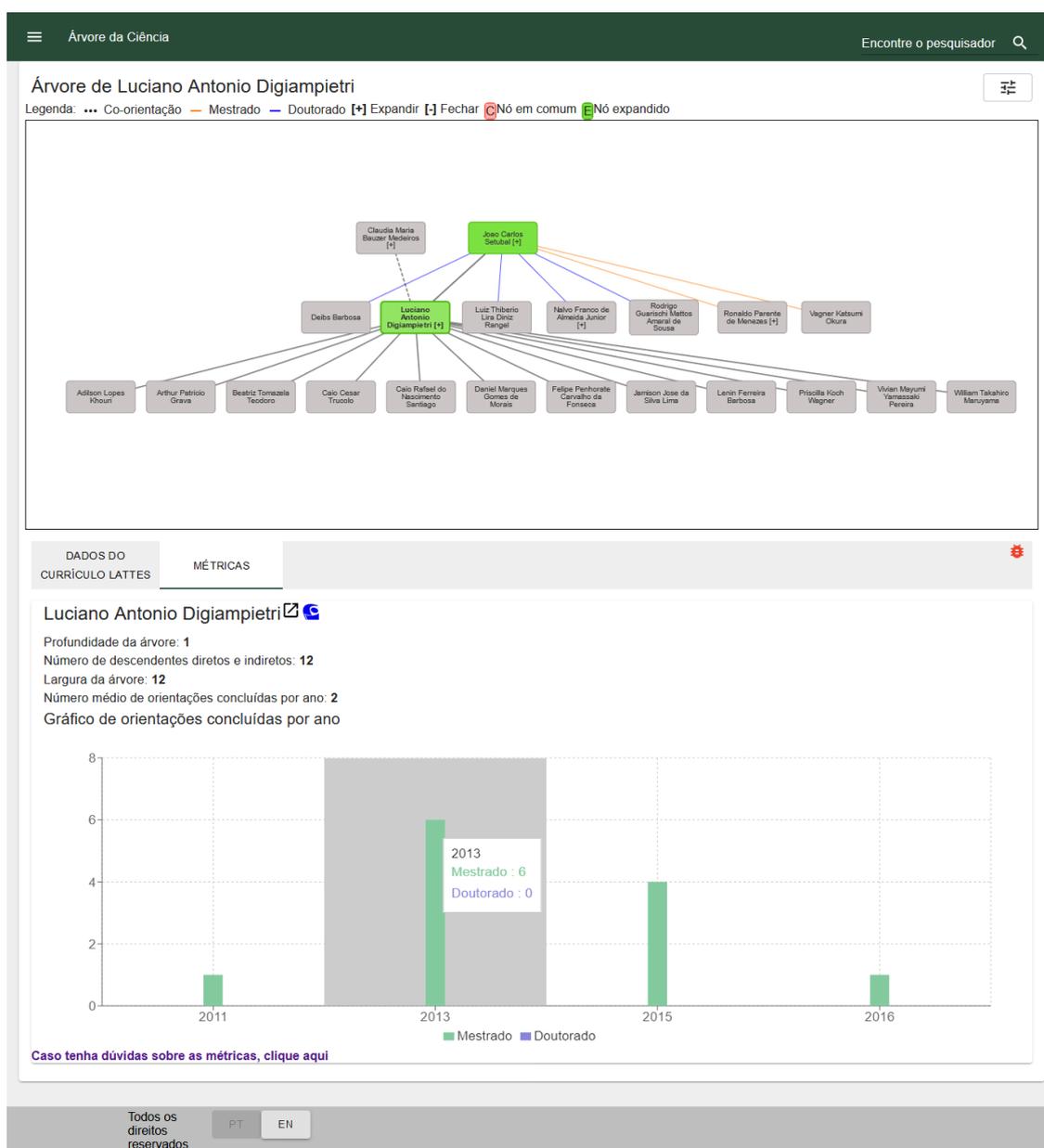


Figura 3.7. Árvore do pesquisador Luciano Antonio Digiampietri expandida a partir do nodo do seu orientador e mostrando a sua aba de métricas. Note que no gráfico da figura é possível clicar nas barras para obter a quantidade de orientações de mestrado e doutorado que elas indicam.

3.3 Detalhes de Implementação

Esta seção aborda os principais detalhes relacionados à implementação da plataforma Árvore da Ciência. Inicialmente, é apresentado o componente que serviu como principal motivador desta dissertação, a aplicação *Front-end*, que é responsável pelo processo de interação com os usuários e pela visualização dos dados tratados pela plataforma, in-

cluindo as métricas geradas a partir deles. Em seguida, é descrita a aplicação *Back-end* que tem como principal função suportar a aplicação *Front-end* ao acessar os repositórios de dados para gerar as árvores a serem exibidas aos usuários e extrair os dados necessários para compor as duas abas (DADOS DO CURRÍCULO e MÉTRICAS) que complementam a tela principal da interface, de acordo com o esquema de comunicação descrito anteriormente na Figura 3.2.

O esforço de desenvolvimento foi realizado em um computador com sistema operacional *Windows* e a plataforma encontra-se hospedada em um servidor com sistema operacional Linux. A sua implantação foi realizada utilizando o ambiente Docker² de modo a facilitar a alocação dos recursos adicionais utilizados, como, por exemplo, o sistema de gerência de bancos de dados adotado, no caso o MariaDB³, e as ferramentas para tratamento dos documentos XML.

3.3.1 Aplicação *Front-end*

Para agregar valor à plataforma *Árvore da Ciência*, foi necessário desenvolver uma interface de usuário que, além de possibilitar a visualização das árvores de modo intuitivo, também apresentasse informações relevantes aos seus usuários a partir dos dados coletados da plataforma Lattes e das métricas derivadas das árvores geradas. Conforme será descrito na Subseção 3.3.2, a comunicação da aplicação *Back-end* com a aplicação *Front-end* é realizada de forma assíncrona como explicado por Simpson [2015] para a linguagem de programação *JavaScript*. Nesse trabalho, a linguagem *JavaScript* é utilizada para alterar conteúdos dinamicamente em um sistema de informação disponibilizado por meio da Internet. Tal dinamicidade permite inclusive que sejam feitas requisições a um servidor, as quais, por sua vez, não impedem o carregamento e o funcionamento da página que as tenha solicitado. Para isso, foi utilizada a biblioteca *ReactJs*⁴, implementada em *JavaScript*, que permite a criação de páginas únicas com conteúdos dinâmicos que são renderizadas diretamente no navegador do usuário.

A biblioteca *ReactJs* foi escolhida para desenvolver a interface da plataforma *Árvore da Ciência* pelas facilidades oferecidas para desenvolvimento de código, permitindo a reutilização de objetos criados, a integração com outras bibliotecas e maior compatibilidade entre diferentes navegadores. Desenvolvida em *JavaScript*, ela é responsável por prover recursos para capturar as interações do usuário com a interface, enviá-las para a aplicação *Back-end* e exibir as respectivas respostas de volta para o usuário.

²<https://www.docker.com/>, Acesso em Novembro de 2020.

³<https://mariadb.org/>, Acesso em Novembro de 2020.

⁴<https://reactjs.org/>, Acesso em Novembro de 2020.

O sistema de *layout* utilizado baseia-se no *Material Design*⁵ que é uma linguagem de *design* comumente usada para gerar visualizações e padronizar elementos em uma tela, como, por exemplo, botões e espaçamento entre textos. A principal vantagem do *Material Design* é que, depois de anos de desenvolvimento, ele é hoje um ambiente para gerenciamento de *layouts* gráficos com responsividade para atender diversos tamanhos de telas, além de possuir uma gama de componentes já estilizados para diversos propósitos, tais como botões, caixas de texto e listas, o que permitiu reduzir enormemente o esforço de desenvolvimento da interface da plataforma Árvore da Ciência.

Para a geração das árvores foi utilizada uma biblioteca que visa principalmente o desenho de redes e grafos, conhecida por *VisJs*⁶. Apesar dessa biblioteca possuir diversos recursos gráficos, ela é utilizada na plataforma Árvore da Ciência apenas para a exibição das árvores genealógicas acadêmicas. A escolha dessa biblioteca se deu pelas facilidades gráficas providas por ela, como *zoom*, cliques em nós e expansão das árvores dinamicamente sem necessidade de recarregamento das respectivas páginas, este um importante recurso para exibir a expansão das árvores dos pesquisadores melhorando a visualização dos seus descendentes.

3.3.2 Aplicação *Back-end*

Para o bom funcionamento da interface proposta para a plataforma Árvore da Ciência, foi necessária a criação de um arcabouço para suporte dessa interface, isto é, uma aplicação que fosse de fácil utilização pelos usuários, interceptasse as requisições corretamente, tivesse acesso aos repositórios de dados e impedisse acessos não autorizados. A utilização desse arcabouço é realizada por um conjunto de interfaces de programação (API's) que constituem a aplicação *Back-end* [Brown & Robinson, 2001].

As árvores genealógicas acadêmicas dos pesquisadores foram geradas a partir do pré-processamento dos currículos no formato XML, utilizando o algoritmo proposto por Dores et al. [2017], escrito em *Python*. Nesta etapa, é acessado o repositório de currículos para identificar, para cada pesquisador, os dados referentes às suas respectivas orientações (título das teses e dissertações orientadas, instituição onde ocorreram as orientações e o ano de concessão dos respectivos títulos). Com base nas análises desses dados, é realizada a desambiguação de nomes para a correta geração das respectivas relações de orientação. Após esse pré-processamento, as relações de orientação geradas são armazenadas no Repositório de Dados na forma de um grafo dirigido. Esses da-

⁵<https://material.io/>, Acesso em Novembro de 2020 (na plataforma é utilizada a implementação disponível em <https://material-ui.com/>)

⁶<https://visjs.org/>, Acesso em Novembro de 2020.

dos ficam disponíveis para serem utilizados para a visualização das respectivas árvores genealógicas acadêmicas e para a geração dos valores das métricas correspondentes.

Para tornar esses dados disponíveis de uma forma rápida e segura, foi desenvolvida uma Web API utilizando a linguagem de programação C# junto com o *framework* .NET Core [Freeman, 2017]. Vale ressaltar que tanto a linguagem C# quanto o *framework* .NET Core são recursos de código aberto e que podem ser usados sem restrições, não dependendo de uma ferramenta ou sistema operacional específicos, o que torna a aplicação desenvolvida multi-plataforma. A aplicação *Back-end* teve o seu desenvolvimento iniciado com a versão 1.1 do *framework* .NET Core, sendo que no momento da escrita desta dissertação a versão utilizada é a 3.1, em relação à qual a plataforma foi atualizada para incluir melhorias de segurança e suporte por um maior período de tempo.

3.4 Estudo de Caso

Esta seção apresenta um estudo de caso que ilustra um acesso hipotético à plataforma *Árvore da Ciência*. Para isso é utilizada a árvore do Prof. Crodowaldo Pavan que possibilita mostrar não só a evolução de sua linhagem acadêmica a partir de dados obtidos do seu currículo disponível na plataforma Lattes, mas também salienta aspectos específicos referentes à sua atuação como pesquisador durante a sua longa carreira acadêmica.

O Prof. Crodowaldo Pavan (1919-2009) foi um dos mais renomados geneticistas brasileiros. Graduado em 1941 em História Natural pela Faculdade de Filosofia, Ciências e Letras na Universidade de São Paulo (USP), ele obteve o doutorado, também pela USP, em 1944 na área de Genética, investigando a adaptação ambiental e a evolução dos peixes cegos das cavernas de Iporanga, no interior de São Paulo. Ele foi também um dos responsáveis por introduzir na biologia o estudo da citogenética da mosca *Rhynchosciara americana*, conhecida por seus cromossomos gigantes, facilitando assim a determinação de locus gênicos. Esse trabalho contribuiu para a demonstração de que a estrutura dos genes e dos cromossomos são alteradas por infecções.

Em sua carreira profissional, o Prof. Pavan foi o responsável pela criação do laboratório de genética celular no Laboratório Nacional de Oak Ridge, EUA. Atuou também como professor na Universidade do Texas em Austin, EUA, até 1975. Sua relevância científica foi reconhecida por diversos prêmios recebidos, entre eles, o Prêmio Nacional de Genética, em 1963, o Prêmio Moinho Santista de Biologia da Fundação Moinho Santista, em 1980 e o Prêmio Alfred Jurzykowski da Academia Nacional de

Medicina, em 1986. Além disso, ele foi um dos poucos brasileiros a se tornar membro da Pontifícia Academia das Ciências, fundada em Roma em 1603 e considerada a primeira academia científica do mundo. O Prof. Pavan foi ainda presidente do CNPq no período de 1986 a 1990, tendo exercido outros cargos em instituições como a FAPESP (1981-1984), a Sociedade Brasileira para o Progresso da Ciência (1981-1986) e o Comitê de Ciência e Tecnologia da OEA (1988-1992).

O currículo do Prof. Pavan foi incluído na plataforma Lattes em 2008, pouco antes de seu falecimento em 2009. Nele consta que ele teria apenas uma única orientação de doutorado. Porém, essa informação não é precisa, pois, além da tardia inclusão de seu currículo na plataforma Lattes, os dados então cadastrados eram incompletos. Apesar disso, como será mostrado a seguir, a árvore genealógica acadêmica do Prof. Pavan na plataforma Árvore da Ciência inclui 6322 descendentes diretos e indiretos. Para a sua geração, foram utilizados não só os dados de seu próprio currículo, mas também de outros 10 currículos coletados da plataforma Lattes, cujos pesquisadores indicaram o Prof. Pavan como orientador, totalizando assim 11 orientações, das quais 10 identificadas pelo algoritmo proposto por Dores et al. [2017] para a geração de árvores genealógicas acadêmicas a partir de dados extraídos de currículos de pesquisadores disponíveis na plataforma Lattes.

A Figura 3.8 mostra a página do Prof. Pavan na plataforma Árvore da Ciência, conforme identificado no canto superior esquerdo, a qual exhibe o primeiro nível de sua árvore, com os nós de seus 11 descendentes acadêmicos diretos, e o conteúdo de sua aba DADOS DO CURRÍCULO LATTES. Essa aba exhibe o seu nome, seguido pelo ícone  que possibilita acessar o mesmo conteúdo em uma nova aba do navegador e pelo ícone da Plataforma Lattes que redireciona o usuário para o currículo original do Prof. Pavan disponível naquela plataforma. Os dados do Prof. Pavan mostrados na seção Formação Acadêmica e Titulação são o título de sua tese, o nome do seu orientador, a instituição em que a tese foi defendida, o ano de conclusão do doutorado, e as suas respectivas grande áreas e áreas do conhecimento conforme classificação do CNPq. Em seguida é mostrado que em seu currículo consta uma única orientação de pós-graduação registrada e também que não foram registradas quaisquer outras formas de orientação. Finalmente, há uma última seção que, quando expandida, apresenta um resumo textual do seu currículo conforme disponível na própria plataforma Lattes.

A Figura 3.9 mostra uma outra instância da página do Prof. Pavan na plataforma Árvore da Ciência que exhibe uma versão mais detalhada de sua árvore genealógica acadêmica com alguns de seus nodos, marcados pelo sinal de subtração ($[-]$), expandidos na cor verde e com a aba de métricas aberta logo abaixo. Examinando-se mais detalhadamente essa outra página, pode-se observar que, a partir da expansão do nodo do seu

Árvore da Ciência Encontre o pesquisador 🔍

Árvore de Crodowaldo Pavan

Legenda: ... Co-orientação — Mestrado — Doutorado [+] Expandir [-] Fechar [C] Não em comum [G] Não expandido

Árvore de relações acadêmicas:

- Andre Dreyfus [*]
- Crodowaldo Pavan [*]
- Luz Edmundo de Magalhães [*]
- Hermione Ely Maira de Campos Boudo [*]
- Antonio Netto Cestari [*]
- Francisco Mauro Salzano [*]
- Wily Beok [*]
- Pedro Henrique Saldanha [*]
- Andre Luiz Parrizhos Feondri [*]
- Newton Goulart Medeira
- Neiva Izabel Pierozzi
- Luiz Alexandre Nogueira de Sa [*]
- Ana Maria Lima de Assis Espin [*]

DADOS DO CURRÍCULO LATTES | MÉTRICAS

Crodowaldo Pavan

Formação Acadêmica e Titulação

Doutorado

Título da tese: OS PEIXES CEGOS DAS CAVERNAS DE IPORANGA E A EVOLUCAO.
 Nome do orientador: ANDRE DREYFUS
 Instituição: Universidade de São Paulo
 Ano de conclusão: 1944

Áreas do pesquisador

Grande área: Ciências Biológicas Ciências Sociais Aplicadas
 Área do Conhecimento: Genética, Comunicação
 Sub-área do conhecimento:
 Especialidade(s):

Orientações de pós-graduação concluídas: 1

Outras orientações concluídas: 0

Resumo do Currículo (informado pelo pesquisador)

Data da última atualização do currículo: 15/09/2008

Todos os direitos reservados | PT | EN

Figura 3.8. Página do Prof. Crodowaldo Pavan mostrando o primeiro nível de sua árvore genealógica acadêmica e dados gerais sobre a sua formação acadêmica.

descendente acadêmico Antonio Netto Cestari, foi possível visualizar uma outra geração de descendentes indiretos do Prof. Pavan, como a pesquisadora Margherita Anna Antonia Maria Barracco que, ao ter o seu nodo expandido, possibilitou a visualização de mais uma geração de descendentes indiretos do Prof. Pavan.

A mesma Figura 3.9 mostra ainda a aba MÉTRICAS da página do Prof. Pavan



Figura 3.9. Árvore do Prof. Crodowaldo Pavan expandida até o seu terceiro nível para visualização de um caminhamento específico pela sua linhagem acadêmica. Logo abaixo da árvore é mostrada a aba com as suas respectivas métricas e o seu gráfico de orientações concluídas por ano.

que apresenta os valores das métricas relativas à sua árvore genealógica acadêmica. Como pode ser visto, a profundidade de sua árvore é igual a 8, o que indica que a partir dele foram formadas oito gerações de novos pesquisadores. Essas oito gerações estão expressas pelo caminho para a orientação mais profunda de sua árvore que inicia-

se com o pesquisador Francisco Mauro Salzano até alcançar o seu último descendente indireto, o pesquisador Expedito Ferreira de Araujo Filho.

Outras métricas importantes para expressar o esforço de formação de um pesquisador são a fertilidade (quantidade de seus descendentes diretos que realizaram pelo menos uma orientação), a largura de sua árvore (quantidade de descendentes diretos) e o número médio de orientações concluídas por ano. No caso do Prof. Pavan, essas métricas são iguais a 9, 11 e 0,29, respectivamente, sendo esta última calculada como sendo a largura de sua árvore dividida pela diferença entre o ano de sua última orientação e o ano de sua primeira orientação, ou seja, $\frac{11}{1993-1955} = 0,29$.

Por fim, o gráfico de orientações concluídas por ano é uma forma complementar de expressar o esforço de orientação de um pesquisador ao mostrar, por meio de um gráfico de barras, no qual a cor azul representa orientações de doutorado e a cor verde orientações de mestrado, o total de orientações por ele concluídas a cada ano. No caso do Prof. Pavan, esse gráfico indica, a partir dos dados extraídos da plataforma Lattes, que ele formou sete alunos de doutorado entre 1955 e 1968, um aluno de mestrado em 1982 e três alunos de doutorado entre 1987 e 1993. Entretanto, o grande impacto do Prof. Pavan na formação acadêmica de novos pesquisadores pode ser medido pelo seu número de descendentes indiretos que totaliza 6316, de modo que os seus 11 orientados contribuíram com a formação de um número extremamente expressivo de mestres e doutores ao longo de suas carreiras acadêmicas.

3.5 Considerações Finais

Analogamente às plataformas Acácia e *The Gold Tree*, a plataforma Árvore da Ciência também utiliza dados extraídos da plataforma Lattes do CNPq para gerar uma representação da genealogia acadêmica dos pesquisadores brasileiros. No entanto, ela foi desenvolvida de modo a prover aos seus usuários uma melhor experiência de visualização dos dados, tanto na forma de árvores genealógicas acadêmicas como também quando apresentados em formato textual.

A plataforma Árvore da Ciência provê cinco métricas que caracterizam aspectos específicos de uma árvore genealógica acadêmica: *profundidade*, *número de descendentes diretos e indiretos*, *fertilidade*, *largura* e *número médio de orientações concluídas por ano*. Além disso, ela também indica o caminho para a orientação mais profunda (somente para os casos em que a profundidade é maior do que um) e disponibiliza um gráfico de orientações concluídas por ano, o que possibilita uma visão clara do impacto dos pesquisadores em relação ao papel por eles desempenhado como orienta-

dores. As métricas consideradas foram escolhidas por serem de fácil compreensão pelos usuários e também suficientemente expressivas para a comparação dos pesquisadores, além de já terem sido academicamente consolidadas, não havendo maior ambiguidade no entendimento delas. Entretanto, a plataforma *Árvore da Ciência* não disponibiliza estatísticas ou provê análises específicas que abranjam de forma mais ampla os seus dados como, por exemplo, no caso de análises segmentadas por instituições ou áreas do conhecimento. Além disso, todas as informações obtidas a partir de suas métricas são derivadas exclusivamente de análises realizadas a partir de dados extraídos dos currículos Lattes dos pesquisadores.

A partir de um *design* funcional específico para a sua finalidade, a plataforma *Árvore da Ciência* visa proporcionar uma melhor interação aos seus usuários, permitindo não só a visualização mas também a exploração das árvores genealógicas acadêmicas. Uma outra característica é a facilidade que ela oferece para a exploração das gerações de orientados, permitindo a navegação pelas árvores genealógicas acadêmicas de forma simples e intuitiva. Dessa forma, é possível aos seus usuários obter uma visão mais clara do papel de um orientador na formação de novos pesquisadores. Ela também provê diferentes métricas que permitem um entendimento das orientações e da contribuição de um pesquisador no cenário nacional de pós-graduação. Por meio dessas métricas é possível comparar diferentes pesquisadores, o que pode ser útil aos programas de pós-graduação ao avaliarem o desempenho de seus orientadores. O fato de os dados utilizados pela plataforma *Árvore da Ciência* serem extraídos a partir dos currículos coletados diretamente da plataforma Lattes garante uma maior assertividade ao processo de geração das árvores, pois não depende da entrada de dados por parte dos pesquisadores, como ocorre em outras plataformas similares, uma vez que, conforme já ressaltado, o algoritmo utilizado para geração das árvores é capaz de identificar orientações explícitas e implícitas a partir dos dados coletados da plataforma Lattes.

Resumindo, com base nos recursos mencionados na tabela comparativa apresentada ao final do Capítulo 2 (Tabela 2.1), a plataforma *Árvore da Ciência* proporciona uma representação da genealogia acadêmica tanto na forma tabular quanto gráfica, como faz a plataforma *The Academic Family Tree*. Ela também provê um bom número de métricas (5) para análise do desempenho acadêmico dos orientadores, compatível com o que oferece a plataforma *Acácia* (7), embora bem aquém do número de métricas providas pela plataforma *The Academic Family Tree* (15).

Embora limitada apenas à análise dos pesquisadores brasileiros, a plataforma *Árvore da Ciência* é a que se mostra mais balanceada em termos dos recursos oferecidos para visualização das árvores e das métricas disponíveis para avaliação do desempenho dos pesquisadores em relação ao processo de orientação acadêmica. Além disso, ela

provê uma interface que se destaca por proporcionar uma navegação intuitiva pelas árvores geradas, possibilitando aos seus usuários uma experiência agradável ao explorar a genealogia acadêmica dos pesquisadores brasileiros, conforme atestam os resultados dos experimentos com usuários relatados no próximo capítulo. Ela também tem os seus dados tratados de forma automatizada, uma vez que consome esses dados a partir de currículos coletados diretamente da plataforma Lattes do CNPq.

Finalmente, outro ponto importante a ser mencionado é a sua arquitetura como sistema computacional, distribuída em dois grupos de aplicações em que a interação do usuário é iniciada pela aplicação *Front-end* e posteriormente transferida de forma assíncrona para a aplicação *Back-end*, ou seja, as interações não dependem de uma espera do usuário, o que proporciona maior fluidez na utilização da plataforma e evita recarregamentos desnecessários de suas páginas. Além disso, a aplicação *Back-end* utiliza uma abordagem baseada em micro-serviços, isto é, a aplicação é separada em pequenos componentes com funções independentes em relação aos dados, melhorando, assim, a performance geral da plataforma.

Capítulo 4

Avaliação com Usuários

O processo de desenvolvimento de software envolve várias etapas como levantamento de requisitos, projeto, implementação, testes e produção [Pressman & Maxim, 2016]. Dentre os testes previstos, estão aqueles que têm por objetivo analisar a perspectiva dos usuários relacionada ao uso do sistema ou aplicação desenvolvida, assim como avaliar a sua funcionalidade e o *design* realizado. No caso particular da plataforma Árvore da Ciência, foram realizadas duas avaliações com usuários que tinham por objetivo coletar dados relacionados tanto à sua percepção do sistema proposto, quanto sobre a sua interação com ele, de modo a indicar as dificuldades e prover sugestões para futuras melhorias [Dumas et al., 1999; Rubin & Chisnell, 2008].

Conforme sugerido por Prates & Barbosa [2003], uma avaliação com usuários é importante para esse tipo de aplicação, pois trata-se de uma plataforma desenvolvida para um público específico que tem o interesse de conhecer não só o seu próprio papel na formação de novos pesquisadores, como também o de outros pesquisadores. Assim, uma avaliação com usuários representativos desse público alvo permite coletar indicadores capazes de mostrar se o sistema desenvolvido atinge o seu objetivo em relação ao apoio que pode oferecer aos seus usuários, bem como na identificação de aspectos dessa interação que possam ser melhorados.

Existem diversas técnicas que permitem avaliar a qualidade de uso de um sistema sob a perspectiva de seus usuários. A seleção dessas técnicas deve levar em consideração o objetivo que se pretende alcançar com a avaliação proposta [Bevan, 2009]. Nesta dissertação, foram aplicadas duas técnicas específicas: uma quantitativa, baseada em questionários, e outra qualitativa, realizada com um grupo de usuários previamente selecionados e conduzida em um ambiente controlado.

A avaliação por questionário teve como objetivo realizar uma coleta de dados mais ampla, focada principalmente na obtenção de indicadores quantitativos de aspectos re-

lacionados à usabilidade da plataforma (facilidade de uso, facilidade de aprendizado, utilidade e satisfação do usuário). Além disso, o questionário utilizado incluía um campo de comentários em que o usuário poderia fazer qualquer observação relativa à plataforma que desejasse. A avaliação com usuários, por outro lado, teve como objetivo obter informações aprofundadas sobre a visão dos participantes sobre a plataforma proposta e a sua experiência ao utilizá-la. Assim, essas duas técnicas são complementares e foram usadas para permitir uma análise ampla da perspectiva dos usuários em relação à plataforma *Árvore da Ciência*, permitindo identificar tanto indicadores de como ela seria recebida pela comunidade científica, como também identificar aspectos relacionados à sua interface e à sua forma de interação que poderiam ser melhorados [Ezzy, 2013; Lazar et al., 2017].

4.1 Avaliação por Questionário

Esta seção descreve como foi realizada a avaliação da plataforma *Árvore da Ciência* por meio de um questionário. Essa avaliação foi realizada através do *Google Forms*¹ que é um serviço de gerenciamento de enquetes a partir do qual seus usuários podem criar questionários genéricos para coleta de dados referentes a pesquisas de mercado, avaliação de produtos, campanhas promocionais, entre outras finalidades. Além das questões propriamente ditas, tais questionários incluem instruções para o seu preenchimento e outras informações necessárias para a finalidade proposta.

A Figura 4.1 mostra o cabeçalho do questionário proposto para a avaliação realizada como parte desta dissertação, o qual inclui uma breve descrição do seu propósito e orientações gerais para o seu preenchimento. Para a coleta de dados, esse questionário foi enviado para uma lista de 5834 *e-mails* de pesquisadores que atuaram no programa dos Institutos Nacionais de Ciência e Tecnologia² (INCTs) em diversas instituições do Brasil. Os INCTs correspondem a 102 centros de pesquisa multicêntricos criados em 2008 com apoio do CNPq e das Fundações de Amparo à Pesquisa de diversos estados do país com o objetivo de desenvolver pesquisa de forma colaborativa envolvendo grupos de diferentes regiões do país. O preenchimento do questionário contou com a participação de 291 respondentes, dos quais 286 foram considerados para análise de suas respostas por terem respondido o questionário em sua totalidade, correspondendo a 4,9% dos pesquisadores contatados por *e-mail*.

¹<https://www.google.com/forms/about>, Acesso em Novembro de 2020.

²http://estatico.cnpq.br/programas/inct/_apresentacao/apresentacao.html, Acesso em Novembro de 2020.

Pesquisa sobre a interface da plataforma Science Tree

Esta pesquisa é parte de um estudo sobre uma plataforma desenvolvida pelo Laboratório de Bancos de Dados do Departamento de Ciência da Computação, UFMG.

O tempo estimado para responder o questionário é de 10 a 15 minutos. Além disso, será necessário interagir com a plataforma Science Tree pelo tempo que o pesquisador desejar.

A seguir apresentamos brevemente a plataforma Science Tree e em seguida explicamos o objetivo deste questionário.

A plataforma Science Tree

Desenvolvido pelo Laboratório de Bancos de Dados do Departamento de Ciência da Computação da UFMG, essa plataforma fornece um ambiente para visualização das árvores genealógicas acadêmicas dos pesquisadores brasileiros, geradas a partir de dados coletados de seus currículos cadastrados na plataforma Lattes do CNPq.

Essa visualização permite a exploração dessas árvores, possibilitando a sua expansão e o caminhar pelos seus vários níveis, bem como a obtenção de dados específicos sobre os pesquisadores e seus orientados.

Sobre este questionário

Antes de responder este questionário, por favor, utilize a plataforma Science Tree disponível em: <https://sciencetree.net>

para que possa então opinar sobre a sua experiência ao interagir com ela. Observe que a plataforma não atua sobre a versão atual do Lattes, mas sobre uma versão que foi coletada anteriormente.

Esta é uma pesquisa de mestrado, conduzida pelo mestrando João Cota (joaocota@dcc.ufmg.br) e orientada pelos professores Alberto Laender e Raquel Prates.

Caso seja de interesse, sinta-se à vontade para entrar em contato pelo e-mail acima, para obter informações sobre a pesquisa ou caso surja qualquer outra dúvida.

Sua participação é anônima e os dados coletados por este questionário serão utilizados apenas para esta pesquisa. Vale ressaltar também que a sua participação é voluntária e não gera nenhum custo ou ganho financeiro. Você também poderá desistir a qualquer momento.

Esta pesquisa está dividida em duas partes. A primeira será para conhecermos melhor quem está participando dela e na segunda você será convidado a responder perguntas sobre a sua interação com a plataforma Science Tree.

Caso concorde em participar da pesquisa clique em Próximo (Next) para dar início a ela.

[Próxima](#)

Nunca envie senhas pelo Formulários Google.

Este conteúdo não foi criado nem aprovado pelo Google. [Denunciar abuso](#) - [Termos de Serviço](#) - [Política de Privacidade](#)

Google Formulários

Figura 4.1. Cabeçalho do questionário utilizado para avaliação da plataforma Árvore da Ciência.

O questionário propriamente dito foi dividido em duas partes. A primeira, de preenchimento obrigatório, teve por objetivo identificar o perfil dos respondentes (Figura 4.2), sendo composta por cinco questões com a finalidade coletar, de forma anônima, os seguintes dados: gênero (três opções, Masculino, Feminino ou Não quero informar), idade (campo livre), área de formação (campo livre), nível de formação (Graduação, Mestrado, Doutorado ou Outro) e ocupação profissional (Estudante de Graduação, Mestrando, Doutorando, Professor em Educação Básica, Professor em Educação Média, Professor Universitário, Pesquisador ou Outro).

A segunda parte do questionário (Figura 4.3) inclui 14 questões de múltipla escolha e uma para resposta aberta que têm por objetivo coletar dados sobre as impressões

Pesquisa sobre a interface da plataforma Science Tree

*Obrigatório

Queremos te conhecer melhor

Qual o seu gênero? *

Feminino

Masculino

Não quero informar

Outro: _____

Qual a sua idade? *

Escolher ▾

Qual a sua área de formação

Sua resposta _____

Qual o seu nível de formação? *

Graduação

Mestrado

Doutorado

Outro: _____

Qual a sua ocupação profissional? *

Estudante de graduação

Mestrando

Doutorando

Professor em Educação Básica

Professor em Educação Média

Professor Universitário

Pesquisador

Outro: _____

[Voltar](#) [Próxima](#)

Nunca envie senhas pelo Formulários Google.

Este conteúdo não foi criado nem aprovado pelo Google. [Denunciar abuso](#) - [Termos de Serviço](#) - [Política de Privacidade](#)

Google Formulários

Figura 4.2. Seção do questionário para identificação do respondentes.

dos usuários quanto à utilização da plataforma. Essas questões foram elaboradas com uma sequência lógica, porém, antes de serem inseridas no questionário, as mesmas questões foram embaralhadas para evitar qualquer viés no momento de serem respondidas. As questões foram apresentadas no questionário de forma aleatória e podem ser agrupadas de acordo com três aspectos importantes para o processo de avaliação da

Pesquisa sobre a interface da plataforma Science Tree

*Obrigatório

Sobre a sua interação com a plataforma Science Tree

Antes de responder as demais perguntas deste questionário utilize a plataforma Science Tree (disponível em <https://sciencetree.net>) para poder avaliar a sua interação com ela.

A plataforma é de fácil utilização *

Discordo Muito

Discordo

Neutro

Concordo

Concordo Muito

Não tenho opinião

Foi fácil aprender a utilizar a plataforma *

Discordo Muito

Discordo

Neutro

Concordo

Concordo Muito

Não tenho opinião

A plataforma é esteticamente agradável *

Discordo Muito

Discordo

Neutro

Concordo

Concordo Muito

Não tenho opinião

A visualização das árvores é clara *

Discordo Muito

Discordo

Neutro

Concordo

Concordo Muito

Não tenho opinião

As cores utilizadas nas páginas e nas legendas não causam confusão. *

Discordo Muito

Discordo

Neutro

Concordo

Concordo Muito

Não tenho opinião

As opções de filtros disponíveis são claras *

Discordo Muito

Discordo

Neutro

Concordo

Concordo Muito

Não tenho opinião

A visualização das métricas referentes ao pesquisador é clara *

Discordo Muito

Discordo

Neutro

Concordo

Concordo Muito

Não tenho opinião

As métricas fornecidas pela plataforma Science Tree sobre a atuação do pesquisador como orientador são úteis. *

Discordo Muito

Discordo

Neutro

Concordo

Concordo Muito

Não tenho opinião

Quando um pesquisador é selecionado, os dados disponibilizados sobre ele são apresentados de forma clara *

Discordo Muito

Discordo

Neutro

Concordo

Concordo Muito

Não tenho opinião

A plataforma proporciona uma navegação intuitiva *

Discordo Muito

Discordo

Neutro

Concordo

Concordo Muito

Não tenho opinião

Esta plataforma é útil no seu dia-a-dia como pesquisador. *

Discordo Muito

Discordo

Neutro

Concordo

Concordo Muito

Não tenho opinião

A visualização da relação orientador-orientado por meio de uma árvore genealógica é útil. *

Discordo Muito

Discordo

Neutro

Concordo

Concordo Muito

Não tenho opinião

A exploração das árvores genealógicas acadêmicas é útil para se ter uma ideia da atuação acadêmica dos pesquisadores. *

Discordo Muito

Discordo

Neutro

Concordo

Concordo Muito

Não tenho opinião

Você teve dificuldade para entender o fluxo de funcionamento da plataforma Science Tree? *

Não

Sim

Em alguns pontos

Por favor, fique à vontade para deixar outros comentários, críticas ou sugestões sobre a sua experiência com a plataforma Science Tree

Sua resposta

Voltar
Próxima

Nunca envie senhas pelo Formulários Google.

Este conteúdo não foi criado nem aprovado pelo Google. [Denunciar abuso](#) - [Termos de Serviço](#) - [Política de Privacidade](#)

Figura 4.3. Seção do questionário com questões a respeito da plataforma.

plataforma *Árvore da Ciência*: percepção geral da plataforma, forma como os dados são dispostos e utilidade da plataforma.

O primeiro grupo de questões, relacionado à percepção geral da plataforma, teve como objetivo avaliar o esforço dispendido pelo usuário ao utilizá-la. Para isto foram propostas as seguintes três questões afirmativas:

- A visualização das métricas referentes ao pesquisador é clara;
- As métricas fornecidas pela plataforma *Árvore da Ciência* sobre a atuação do pesquisador como orientador são úteis;
- Quando um pesquisador é selecionado, os dados sobre ele são apresentados de forma clara.

O segundo grupo de questões teve como objetivo avaliar a forma como a plataforma *Árvore da Ciência* disponibiliza os seus dados e possibilita a navegação pelas árvores genealógicas acadêmicas. Para isso, foram consideradas as seguintes oito questões, das quais as sete primeiras são afirmativas:

- A plataforma é esteticamente agradável;
- A visualização das árvores é clara;
- As cores utilizadas nas páginas e nas legendas não causam confusão;
- As opções de filtros disponíveis são claras;
- A plataforma é de fácil utilização;
- Foi fácil aprender a utilizar a plataforma;
- A plataforma proporciona uma navegação intuitiva;
- Você teve dificuldade para entender o fluxo de funcionamento da plataforma *Árvore da Ciência*?

Finalmente, o terceiro grupo de questões, teve como objetivo avaliar a utilidade geral da plataforma *Árvore da Ciência*. Para isso, foram consideradas as seguintes três questões afirmativas:

- Esta plataforma é útil no seu dia-a-dia como pesquisador;
- A visualização da relação orientador-orientado por meio de uma árvore genealógica é útil;

- A exploração das árvores genealógicas acadêmicas é útil para se ter uma ideia da atuação acadêmica dos pesquisadores.

Ao final do questionário foi adicionada uma pergunta para determinar se o respondente teve dificuldades em entender o fluxo da plataforma. A pergunta era fechada com as seguintes opções de respostas ‘Sim’, ‘Não’ e ‘Em alguns pontos’. Finalmente, a última pergunta do questionário era aberta e opcional, permitindo que o respondente pudesse deixar sugestões, críticas e observações gerais sobre a sua experiência de uso da plataforma. Esse campo foi posteriormente analisado e os respectivos comentários agrupados em categorias de acordo com o seu conteúdo. Os resultados dessa análise estão descritos na Subseção 4.1.5. A Figura 4.4 apresenta a página de agradecimento ao respondente, contendo um campo de *e-mail* caso fosse de seu interesse deixar um contato para recebimento dos resultados da avaliação realizada.

As respostas para as questões de múltipla escolha propostas foram baseadas na escala Likert³. Esta é uma das escalas mais indicadas para pesquisas de opinião, pois combina resultados da área de estatística e psicologia para promover uma profunda e lógica imersão na mente de quem avalia um determinado objeto [Albaum, 1997; Dalmoro & Vieira, 2014; Joshi et al., 2015]. Além disso, questões construídas a partir da escala Likert apresentam uma afirmação auto-descritiva e oferecem como opção de resposta uma escala de pontos com descrições verbais que contemplam extremos, como “concordo muito” e “discordo muito”. Dessa maneira, é possível medir a intensidade de concordância com uma afirmação variando as respostas em diferentes níveis. Outra característica interessante da escala Likert é o fato de dar ao respondente a opção não expressar a sua opinião [Lazar et al., 2017]. No caso do questionário utilizado nesta avaliação, foram considerados cinco níveis de resposta (“Discordo Muito”, “Discordo”, “Neutro”, “Concordo”, e “Concordo Muito”) e uma opção adicional “Não tenho opinião”.

A seguir, são apresentados os resultados obtidos a partir dos questionários preenchidos pelos respondentes. Uma cópia exata desse questionário pode ser obtida por meio da primeira URL disponível no pé desta página⁴. As respostas dos respondentes estão disponíveis de forma anônima por meio da segunda URL disponível no pé desta página⁵, mas também podem ser obtidas contatando-se diretamente este autor.

³Escala proposta por Rensis Likert (1903-1981), professor de psicologia e sociologia, e Diretor do Instituto de Psicologia de Michigan. Acesso em Novembro de 2020.

⁴<https://arvoreciencia.page.link/formulario>

⁵<https://arvoreciencia.page.link/respostas>.

Pesquisa sobre a interface da plataforma Science Tree

Agradecimento

Gostaríamos de agradecer a você que dedicou um tempo para usar a plataforma Science Tree. As respostas deste questionário serão muito úteis para a melhoria da interface da nossa plataforma.

Em caso de dúvidas ou outros comentários sinta-se a vontade para entrar em contato conosco:

Mestrando: João Cota (joacota@dcc.ufmg.br)
Orientadores: Alberto Laender e Raquel Prates

Caso tenha interesse em receber os resultados desta pesquisa, quando a finalizarmos, por favor deixe um email de contato (opcional).

Sua resposta

[Voltar](#) [Enviar](#)

Nunca envie senhas pelo Formulários Google.

Este conteúdo não foi criado nem aprovado pelo Google. [Denunciar abuso](#) - [Termos de Serviço](#) - [Política de Privacidade](#)

Google Formulários

Figura 4.4. Página de agradecimento pelo preenchimento do questionário.

4.1.1 Perfil dos Respondentes

O perfil dos respondentes foi determinado na primeira parte do questionário e mostrou que, dos 286 pesquisadores respondentes, 179 (62,5%) se declararam do gênero masculino, 103 (36,0%) do gênero feminino e 4 (1,5%) optaram por não declarar o gênero, conforme mostrado no gráfico da Figura 4.5.

A idade dos respondentes está distribuída em um espectro que varia de 24 a 75 anos, sendo que o maior número absoluto de respondentes se encontra na faixa etária entre 50 e 59 anos, como mostrado no gráfico da Figura 4.6. Este fato pode indicar que a maior parte dos respondentes é constituída por pesquisadores ainda em atividade e, eventualmente, já consolidados em suas respectivas áreas de pesquisa e com várias orientações concluídas. Outro fato interessante é a existência de 19 respondentes

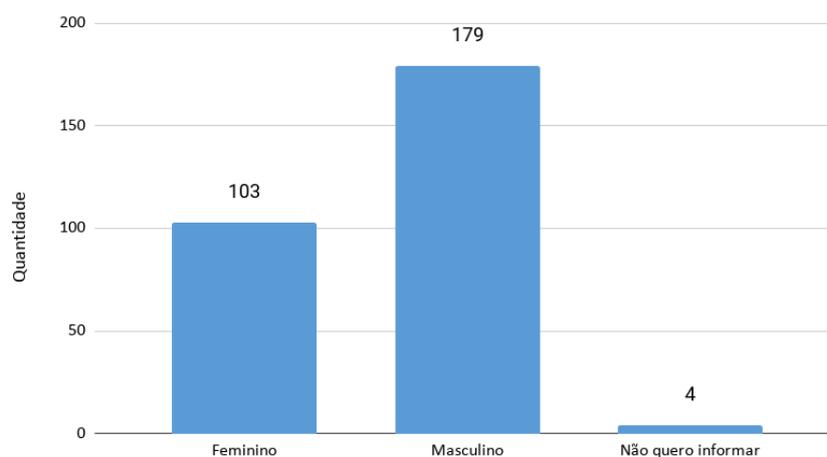


Figura 4.5. Gênero dos respondentes.

(6,6%) com mais de 70 anos, evidenciando que pesquisadores, possivelmente aposentados, também possuem interesse em conhecer as suas próprias árvores genealógicas acadêmicas e, possivelmente, explorar as árvores de outros colegas.

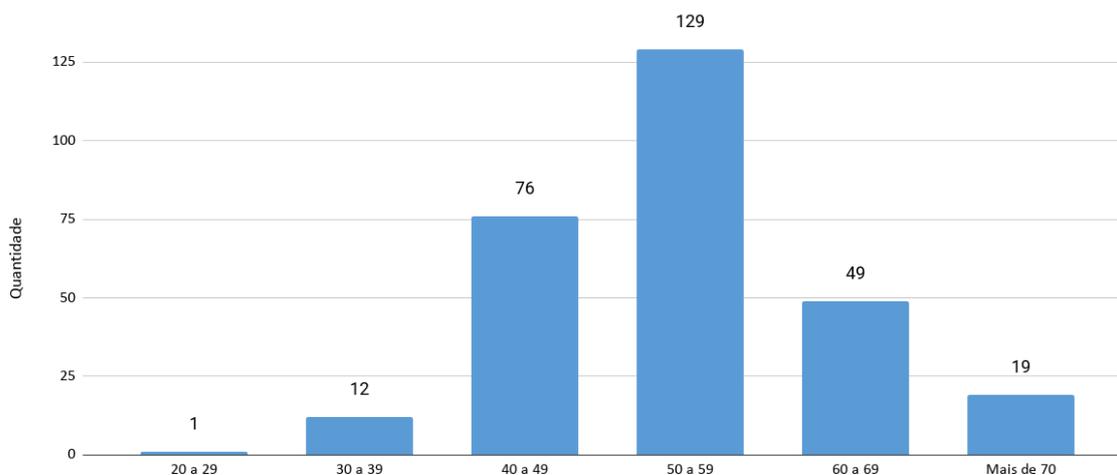


Figura 4.6. Idade dos respondentes agrupada por faixas de dez anos.

A questão referente ao nível de formação dos respondentes, embora também fosse de múltipla escolha, possuía uma opção adicional que podia ser preenchida livremente. A partir das respostas a essa questão, foi possível construir o gráfico da Figura 4.7, o qual mostra que a grande maioria dos respondentes, 242 (84,6%), possui doutorado. Em seguida, aparecem os respondentes que fizeram pós-doutorado, 32 (11,1%), os que possuem livre docência, 7 (2,4%) e os que possuem apenas mestrado, 4 (1,3%). Finalmente, um único respondente (0,3%) era apenas graduado.

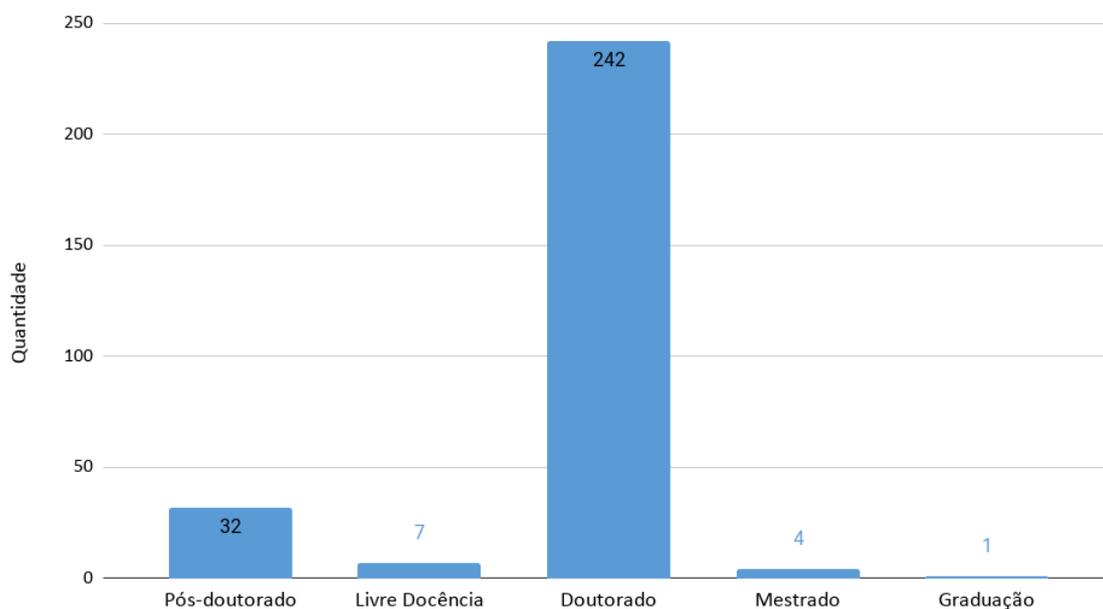


Figura 4.7. Total de respondentes por nível de formação.

O gráfico da Figura 4.8 mostra a distribuição da ocupação profissional dos respondentes. A partir desse gráfico é possível observar que 220 (76,9%) são professores universitários, 54 (18,8%) são pesquisadores e 12 (4,2%) indicaram outras ocupações, sendo 6 (2,1%) técnicos especialistas e 5 (1,8%) pós-graduandos, além de um estudante de graduação.

A área de formação dos respondentes foi obtida pelo preenchimento de um campo específico do questionário proposto. O fato desse campo ser aberto, no entanto, trouxe algumas dificuldades na análise dos dados como, por exemplo, discrepâncias na grafia de uma mesma área do conhecimento, já que alguns respondentes optaram por escrever suas respectivas áreas de forma abreviada. Houve ainda aqueles que preencheram de forma genérica, indicando apenas a grande área em que se insere a sua área de atuação. Para contornar esses problemas, foi realizado um agrupamento considerando apenas aquelas áreas mais mencionadas nas respostas do questionário, conforme mostrado na Figura 4.9. O agrupamento “Outras” corresponde às demais áreas com número de respondentes menor ou igual a quatro, como, por exemplo, Economia, Ecologia, Veterinária e Sociologia. Observando mais detalhadamente a Figura 4.9, é possível notar que o maior número de respondentes ocorre na área de Biologia com um total de 50 (17,5%), contrastando diretamente com as áreas de Humanas e Oceanografia que figuram, cada uma delas, com apenas 5 (1,8%) respondentes. Nas demais áreas o número de respondentes foi de 42 (14,5%) na Engenharia, 41 (14,3%) na Saúde,

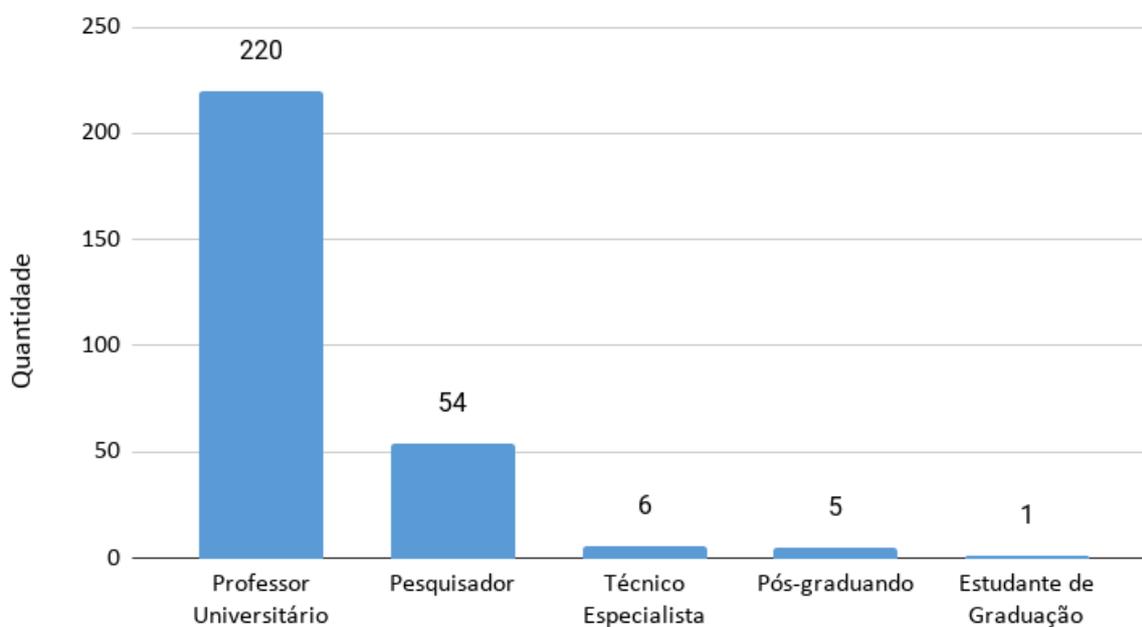


Figura 4.8. Quantidade de respondentes por ocupação profissional.

35 (12,2%) na Física, 25 (8,8%) na Computação, 20 (7,0%) na Química, 15 (5,3%) na Farmácia, 10 (3,5%) na Bioquímica, 8 (2,8%) na Matemática. Outras áreas com menos de cinco respondentes cada contabilizaram um total de 30 (10,5%).

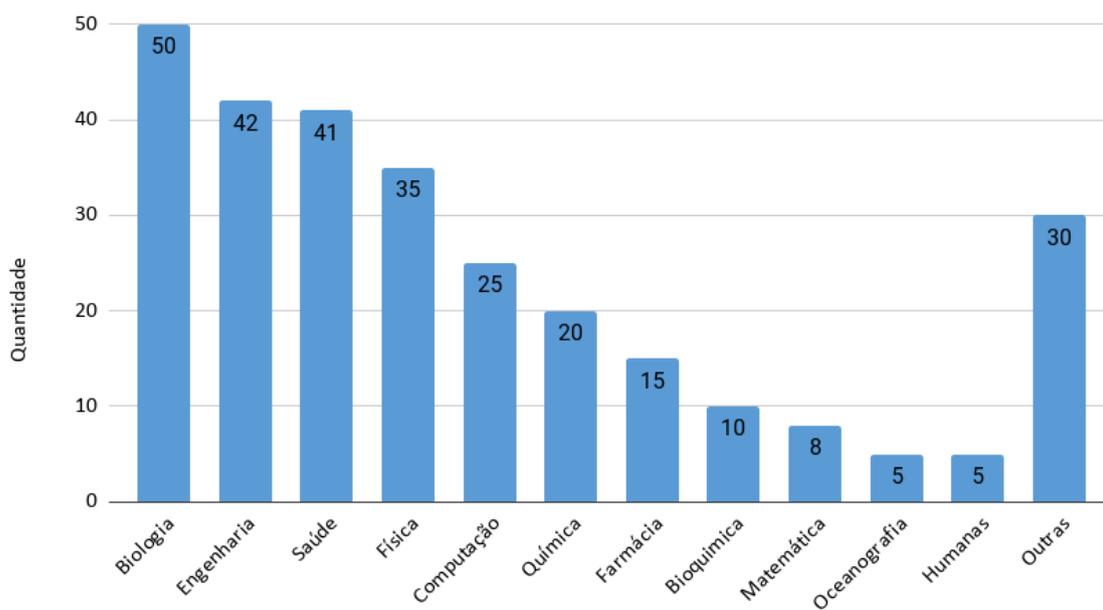


Figura 4.9. Total de respondentes por área de formação.

4.1.2 Avaliação Geral da Plataforma

Uma vez caracterizado o perfil dos respondentes, a próxima etapa foi capturar a impressão deles em relação à apresentação e relevância dos dados mantidos na plataforma *Árvore da Ciência*. As respostas foram avaliadas utilizando a escala Likert, considerando as seguintes seis alternativas: “Concordo Muito”, “Concordo”, “Neutro”, “Discordo”, “Discordo Muito” e “Não tenho opinião”.

Os gráficos da Figura 4.10 mostram o agrupamento dos respondentes de acordo com as respostas ao primeiro grupo de três questões que compõem a segunda parte do questionário referente à percepção dos usuários em relação às características visuais dos dados referentes aos pesquisadores providos pela plataforma (ver página 49). Analisando os gráficos individualmente, é possível observar no primeiro deles que a grande maioria dos respondentes (182, 63,6%) concorda que a visualização das métricas é clara, sendo que 45 (15,7%) concordam muito e 137 (47,9%) apenas concordam. Por outro lado, apenas 45 respondentes (15,7%) possuem alguma discordância em relação a essa característica, dos quais 34 (11,9%) simplesmente discordam e apenas 8 (2,8%) discordam muito, enquanto que 37 (12,9%) são neutros e 22 (7,7%) não têm opinião. Em relação à utilidade das métricas fornecidas pela plataforma (segundo gráfico), observa-se um aumento da aprovação dos respondentes, sendo que 57 (19,9%) deles concordam muito e 138 (48,2%) concordam que as métricas referentes à atuação dos pesquisadores são úteis, enquanto que apenas 23 (8,0%) respondentes se mostram discordantes, dos quais 18 (6,3%) simplesmente discordam e apenas 5 (1,8%) discordam muito. Dos demais respondentes, 53 (18,5%) são neutros e 15 (5,3%) não opinaram a respeito. Esses números indicam a relevância das métricas fornecidas pela plataforma para avaliação da atuação dos pesquisadores de acordo com os respondentes. Finalmente, em relação à clareza da apresentação dos dados sobre os pesquisadores (terceiro gráfico), novamente observa-se uma boa aceitação por parte dos respondentes, dos quais 53 (18,5%) concordam muito e 135 (47,8%) concordam, enquanto 33 (11,5%) discordam e apenas 8 (2,8%) discordam muito. Em relação aos demais respondentes, 47 (16,4%) se mostram neutros e 10 (3,5%) optaram por não opinar.

A avaliação das características visuais relativas à apresentação dos dados gerados pela plataforma foi realizada considerando o segundo grupo de quatro questões da segunda parte do questionário (ver página 49). Os gráficos referentes às respostas dadas a essas quatro questões são mostrados na Figura 4.11. Como pode ser observado no primeiro gráfico, a maioria dos respondentes (162, 56,7%) considera a plataforma esteticamente agradável, sendo que 42 (14,7%) deles concordam muito e 120 (42,0%) concordam, enquanto 52 (18,2%) possuem alguma discordância em relação a esse as-

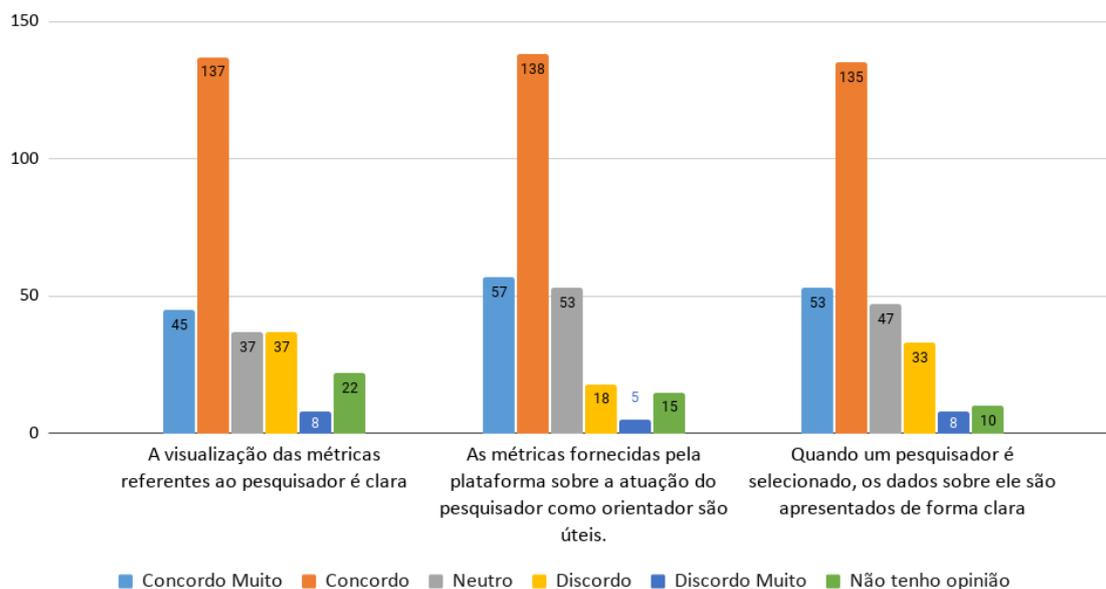


Figura 4.10. Agrupamento dos respondentes de acordo com as respostas às questões relativas às características visuais dos dados da plataforma.

pecto, dos quais 46 (16,1%) discordam e apenas 6 (2,1%) discordam muito. Em relação aos demais respondentes, 67 (23,4%) se mostram neutros e 5 (1,8%) não opinaram. Sobre a clareza da visualização das árvores (segundo gráfico), novamente a maioria dos respondentes (178, 62,2%) a considera satisfatória, dos quais 53 (18,5%) concordam muito e 125 (43,7%) concordam, enquanto que 49 (17,1%) discordam e apenas 10 (3,5%) discordam muito. Considerando os que se mostraram indiferentes, 42 (14,7%) se mostraram neutros e apenas 7 (2,5%) não opinaram. Em relação à afirmação que as cores utilizadas nas páginas e nas legendas não causam confusão aos usuários (terceiro gráfico), novamente a grande maioria dos respondentes (200, 70,0%) concorda com ela, sendo que 60 (21,0%) respondentes concordam muito e 140 (49,0%) apenas concordam, enquanto 24 (8,4%) discordam, e apenas 1 (0,4%) discorda muito. Em relação aos demais respondentes, 56 (19,6%) são neutros e 5 (1,8%) não têm opinião. Finalmente, em relação à clareza das opções disponíveis para os filtros de configuração da plataforma (quarto gráfico), esta é a questão que apresenta a maior discordância entre os respondentes, sendo que 26 (9,1%) concordam muito e 102 (35,7%) concordam, enquanto que 74 (25,9%) discordam e 8 (2,8%) discordam muito. Além disso, 50 (17,5%) respondentes se manifestaram neutros e 26 (9,1%) não opinaram, o que pode levar à conclusão que os filtros disponíveis estão pouco visíveis e por isso não foram devidamente utilizados por eles.

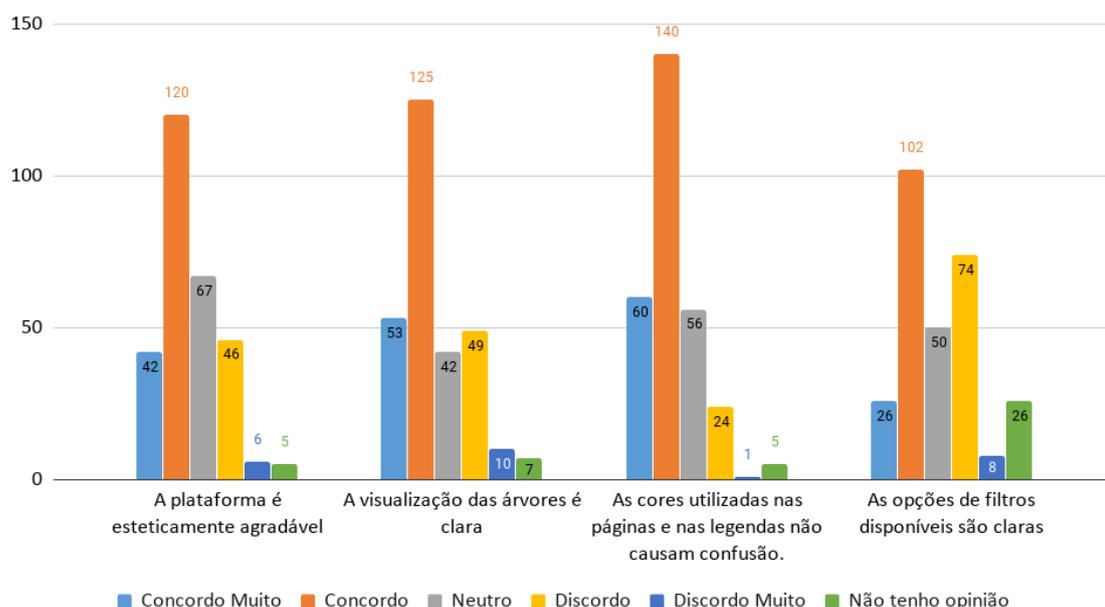


Figura 4.11. Agrupamento dos respondentes de acordo com as respostas às questões relativas aos aspectos visuais e aos filtros disponíveis na plataforma.

4.1.3 Modo de Interação com a Plataforma

As três questões finais da segunda parte do questionário avaliam o modo de interação dos usuários com a plataforma (ver página 49). Os agrupamentos das respostas são mostrados nos gráficos da Figura 4.12. No primeiro gráfico é possível verificar que a grande maioria dos respondentes (221, 77,3%) concorda que a plataforma é de fácil utilização, sendo que 81 (28,3%) deles concordam muito e 140 (49,0%) apenas concordam, enquanto que 22 (7,7%) discordam e apenas 1 (0,04%) discorda muito. Além disso, 37 respondentes (12,9%) se mostraram neutros e 5 (1,8%) não têm opinião. Sobre o aprendizado para utilização da plataforma (segundo gráfico), a grande maioria dos respondentes (243, 85,0%) concorda que foi fácil aprender a utilizá-la, sendo que 105 (36,7%) concordam muito e 138 (48,3%) simplesmente concordam, enquanto que apenas 5 (1,8%) discordam e 4 (1,4%) discordam muito. Os que se mostraram neutros são 19 (6,6%) e apenas 5 (1,8%) não têm opinião formada. Finalmente, em relação à navegação proporcionada pela plataforma (terceiro gráfico), 194 respondentes (67,8%) a consideram intuitiva, dos quais 53 (18,5%) concordam muito e 141 (49,3%) apenas concordam, enquanto que 28 (9,8%) discordam e 6 (2,1%) discordam muito. Por fim, 48 (16,8%) respondentes se mostraram neutros e 10 (3,5%) não têm opinião.

Finalmente, com o objetivo de compreender melhor o entendimento dos respondentes se o fluxo de funcionamento da plataforma é difícil ou não, foi elaborada a se-

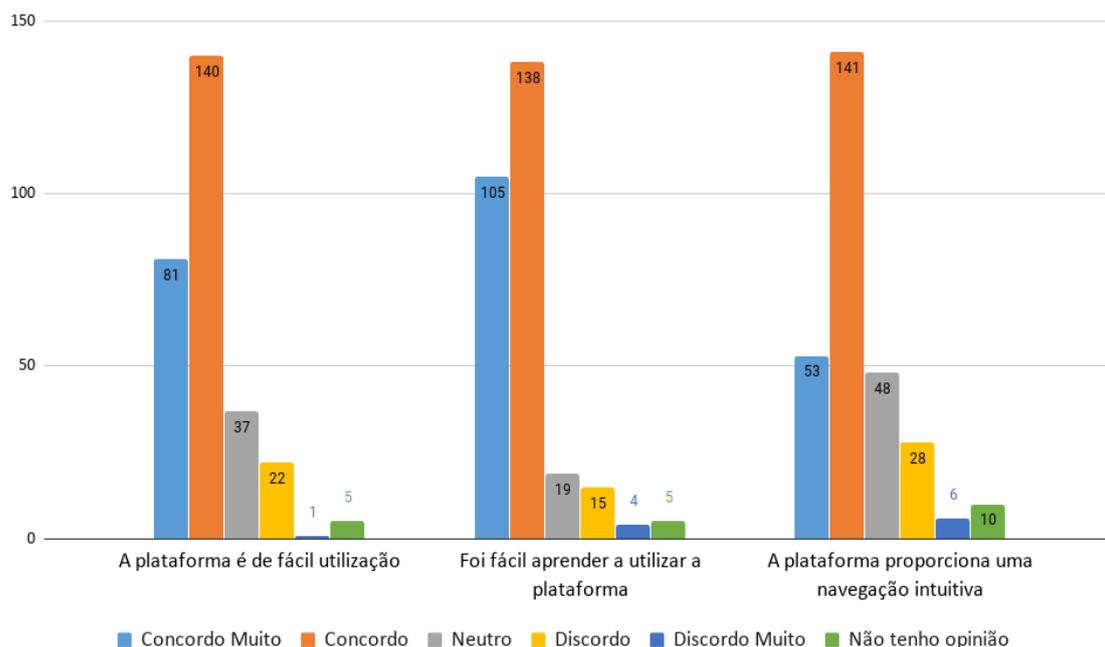


Figura 4.12. Agrupamento dos respondentes de acordo com as respostas às questões relativas à facilidade de uso e aprendizado da plataforma.

guinte questão ao final da segunda parte do questionário: “Você teve dificuldade para entender o fluxo de funcionamento da plataforma *Árvore da Ciência*?”. Diferentemente das demais questões, esta possuía apenas três possibilidades de resposta: *Não*, *Sim* e *Em alguns pontos*. O gráfico da Figura 4.13 mostra que a grande maioria dos respondentes, ou seja 192 deles (67,1%), não teve qualquer dificuldade para entender o fluxo de funcionamento da plataforma, enquanto que apenas 28 (9,8%) apresentaram alguma dificuldade e 66 (23,1%) tiveram dificuldade apenas em alguns pontos específicos. Em suma, esse gráfico mostra que menos de um terço dos respondentes enfrentaram algum tipo de dificuldade em relação ao fluxo de funcionamento da plataforma.

4.1.4 Utilidade da Plataforma

O terceiro grupo de questões da segunda parte do questionário teve como objetivo avaliar a utilidade da plataforma *Árvore da Ciência* para os seus usuários no dia-a-dia de suas atividades acadêmicas, conforme expresso pelos gráficos da Figura 4.14. Como pode ser observado no primeiro gráfico, 125 respondentes (43,7%) consideraram a plataforma útil no seu dia-a-dia como pesquisador, dos quais 33 (11,5%) concordam muito e 92 (32,2%) apenas concordam, enquanto outros 33 (11,5%) possuem alguma discordância e apenas 15 (5,2%) discordam muito. Dentre os demais, 91 (31,8%) se

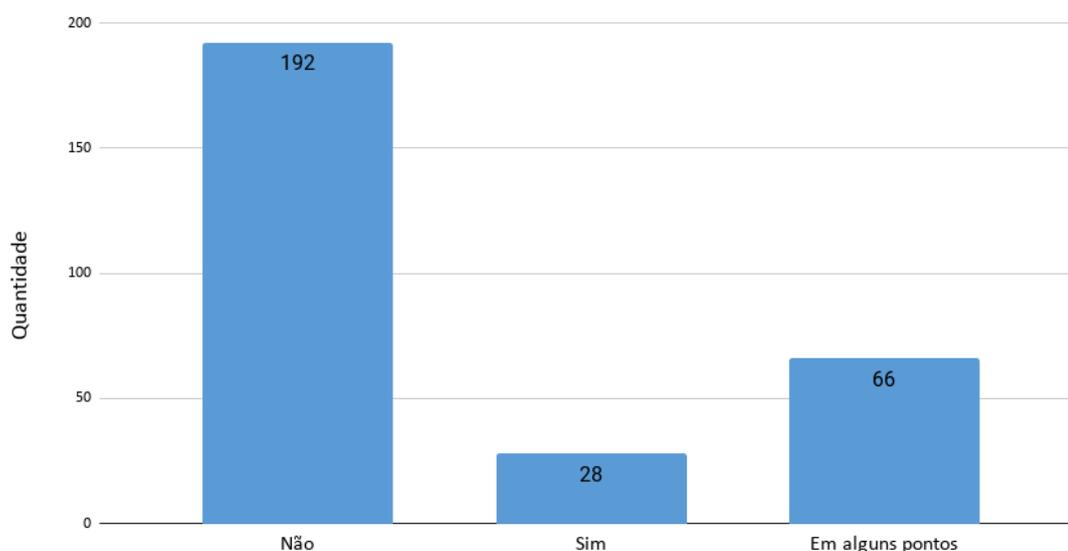


Figura 4.13. Opinião dos respondentes em relação à questão “Você teve dificuldade para entender o fluxo de funcionamento da plataforma?”.

mostraram neutros e apenas 22 (7,7%) não têm opinião.

Em relação à utilidade da visualização da relação orientador-orientado por meio de uma árvore genealógica acadêmica, o segundo gráfico mostra que a grande maioria dos respondentes, 175 (75,2%), concorda que a visualização dessa relação é útil, sendo que 72 (25,2%) concordam muito e 143 (50,0%) concordam, enquanto que 18 (6,3%) possuem alguma discordância e apenas 6 (2,1%) discordam muito. Neutros contam 40 (14,0%) respondentes e apenas 7 (2,5%) não têm opinião. Esses resultados indicam que a visualização da relação orientador-orientando é relevante para a imensa maioria dos respondentes.

Finalmente, em relação à utilidade da exploração de suas árvores genealógicas acadêmicas para entender a atuação dos pesquisadores, o terceiro gráfico mostra que a grande maioria dos respondentes, 223 (78,0%), concorda que tal exploração é útil, sendo que 84 (29,4%) deles concordam muito e 139 (48,6%) concordam, enquanto 18 (6,3%) discordam e apenas 3 (1,1%) discordam muito. Dos demais respondentes, 35 (12,2%) são neutros e 7 (2,4%) não têm opinião. Aqui vale ressaltar que os pesquisadores conhecem suas próprias orientações no dia-a-dia, de modo que a utilidade da plataforma neste caso específico está em conhecer a atuação de outros pesquisadores ao longo de suas trajetórias acadêmicas.

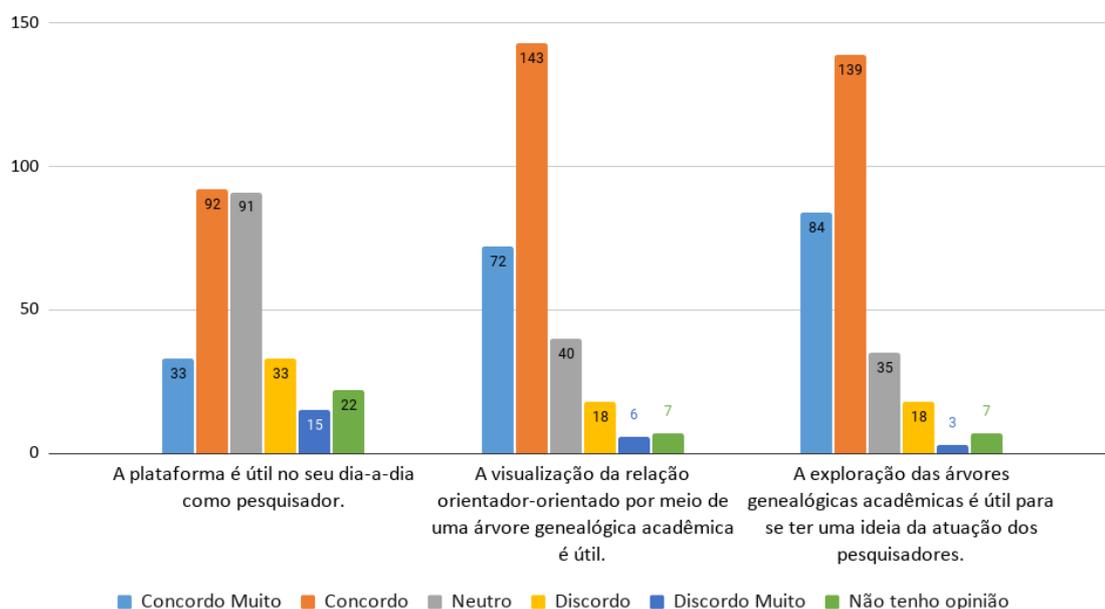


Figura 4.14. Agrupamento dos respondentes de acordo com as respostas às questões relativas à utilidade da plataforma no dia-a-dia dos pesquisadores e para exploração das árvores genealógicas acadêmicas.

4.1.5 Comentários dos Respondentes

Ao final da segunda parte do questionário foi incluído um campo livre para que os respondentes pudessem deixar comentários acerca do uso e da relevância da plataforma. Esse espaço também foi aberto para sugestões, críticas e considerações de outra natureza. Assim, ao final dessa etapa de avaliação, os comentários recebidos foram analisadas a fim de que pudessem ser categorizados para outras considerações.

Dos 286 respondentes do questionário, 147 (51,4%) deixaram comentários com diferentes conteúdos. Porém, 15 desses comentários foram desconsiderados por não apresentarem conteúdo relevante para a avaliação da plataforma. Exemplos desses comentários incluíam *e-mails* deixados pelos respondentes, palavras de incentivo e comentários gerais que nada agregariam à análise proposta. Os 132 comentários restantes (46,2%) foram agrupados em 25 tipos com temáticas semelhantes que, por sua vez, foram reagrupados em quatro categorias específicas: melhorias, sugestões, problemas e comentários adicionais. As quatro categorias identificadas são abordadas nas subseções a seguir.

4.1.5.1 Melhorias Sugeridas

A primeira categoria aborda sugestões específicas que visam melhorar o entendimento do funcionamento da plataforma, como explicações sobre as métricas disponíveis (7 respondentes), instruções sobre determinadas funções (6 respondentes) e informações gerais como a data da coleta de dados mais recente efetuada na plataforma Lattes (5 respondentes). A melhoria mais sugerida foi a inclusão de uma explicação sobre as métricas disponíveis, conforme mostrado pelo gráfico da Figura 4.15, seguida da inclusão de instruções gerais para utilização da plataforma. Entretanto, vale ressaltar que a plataforma já inclui um diagrama explicativo das métricas adotadas para expressar aspectos específicos das árvores geradas, mas que aparentemente passou despercebido por parte dos pesquisadores que participaram de sua avaliação. Os comentários listados a seguir exemplificam duas das melhorias mais sugeridas pelos respondentes: “A ideia é excelente e a plataforma é muito útil e de fácil uso. Creio que uma página com instruções de uso seria útil. Confesso que não sei se explorei todas as informações da plataforma [...]” e “A data na qual os dados foram obtidos do Lattes deveria ser informada. [...]”.

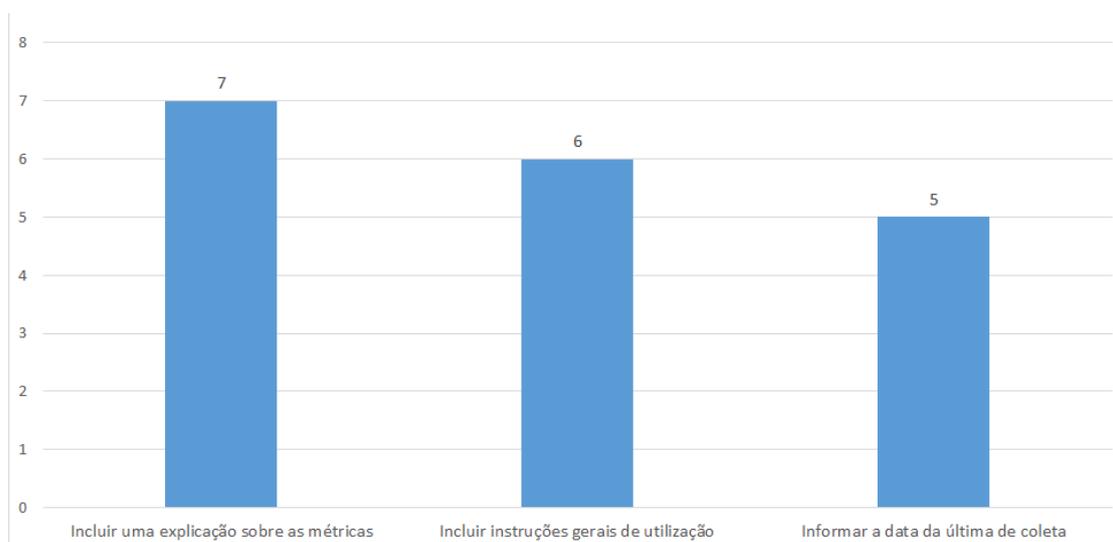


Figura 4.15. Gráfico referente às sugestões propostas para ampliação das informações gerais sobre a plataforma.

4.1.5.2 Sugestões Propostas

As sugestões propostas constituem a segunda categoria com maior número de comentários por parte dos respondentes. Entretanto, é importante ressaltar que muitas dessas sugestões extrapolam o escopo da plataforma *Árvore da Ciência*, que visa registrar a

genealogia acadêmica dos pesquisadores brasileiros e não aspectos específicos de suas carreiras. As sugestões dentro do escopo da plataforma foram classificadas em oito categorias de acordo com a Figura 4.16.

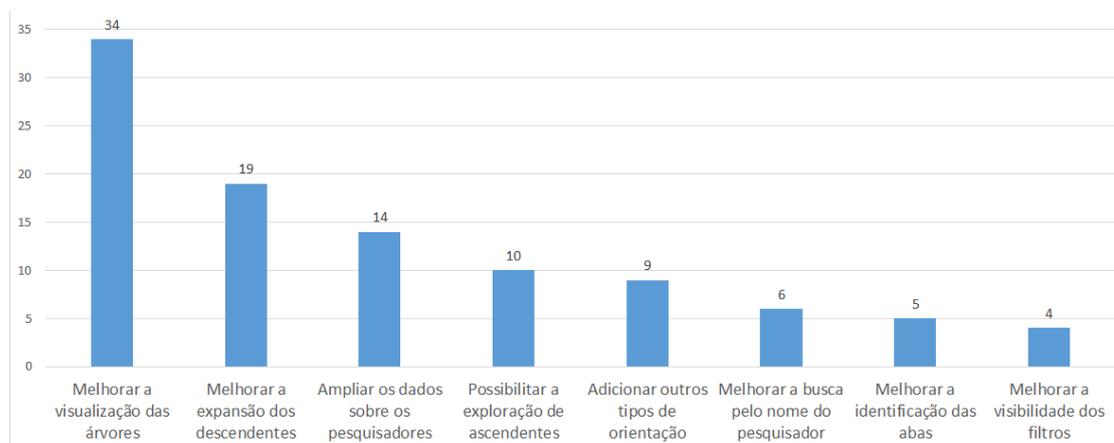


Figura 4.16. Sugestões propostas para melhoria da versão atual da plataforma Árvore da Ciência.

Dentre essas melhorias, a mais sugerida foi relacionada à visualização das árvores representando 34 dos 132 respondentes, como pode ser visto na Figura 4.16. Também foi relatado pelos usuários que a organização e as cores das arestas, bem como a disposição dos nós poderiam ser melhor apresentados como sugerido, por exemplo, no seguinte comentário: *“a árvore não tem visualização fácil para quem tem mais do que 10 orientados; [...]”*. A segunda melhoria mais solicitada foi relacionada à expansão dos nós dos descendentes nas árvores, totalizando 19 respondentes, conforme mostrado na Figura 4.16 e exemplificado pelo comentário *“Não fecha depois de expandir uma seleção, ou seja [–] não funcionou. [...]”*. Ambos os problemas descritos neste parágrafo estão relacionados à biblioteca *Vis.js*, utilizada para apresentação e interação com as árvores, e deverão ser sanados em versões futuras da plataforma.

Em menor quantidade, porém, não menos importante, 14 respondentes relataram a necessidade de se fornecer mais dados referentes aos pesquisadores como, por exemplo, a instituição à qual estão vinculados e palavras-chave relacionadas às suas pesquisas, relacionamento com outros pesquisadores e nuvens de palavras contendo as palavras-chave mais utilizadas em suas publicações. Exemplos dessas sugestões são evidenciados pelos comentários *“Esperava maior cruzamento de informações.”* e *“A árvore é montada de forma simplificada, utilizando apenas as orientações diretas. Seria interessante tentar incluir ramificações mais complexas. Por exemplo, quando um estudante foi orientado por 2 pessoas (ou doutorado sanduíche), ambas poderiam aparecer [...]”*. O

segundo comentário, especificamente, denota uma certa dificuldade de alguns usuários em perceber como se dá a visualização das árvores, uma vez que elas apresentam não só as orientações plenas, mas também as coorientações do pesquisador selecionado.

Outras sugestões indicadas na Figura 4.16 incluem funcionalidades específicas para exploração dos ascendentes, totalizando 10 respondentes, e a inclusão de outros tipos de orientação, como pós-doutoramento ou iniciação científica, além de dados específicos sobre o perfil dos egressos (por exemplo, a instituição em que foi concluída a graduação e a pós-graduação), totalizando nove respondentes. Outra sugestão de melhoria na plataforma está relacionada ao aperfeiçoamento da busca pelo nome dos pesquisadores, pois seis respondentes relataram dificuldades com esse tipo de busca. Uma das propostas de solução sugeridas foi realizar a busca considerando partes dos nomes dos pesquisadores ou mesmo nomes similares, conforme evidenciado pelo comentário *“O reconhecimento de nomes deve incluir buscas com o nome mais parecido pois às vezes não se conhece o nome completo dos pesquisadores [...]”*. Entretanto, a busca atual da plataforma *Árvore da Ciência* já contempla algumas dessas facilidades, como a busca com base em nomes parciais. Assim, pretendemos rever se a busca é de fato efetiva, ou se há algum problema na implementação desta ferramenta.

Outro ponto relatado foi a duplicação de nomes ao realizar buscas na plataforma (6 respondentes). Esse é um problema conhecido e que envolve pesquisadores que tiveram seu nome alterado por diversos motivos, entre eles casamento e adoção de nome científico. Há também pesquisadores com nomes semelhantes, mas que trabalham em áreas ou instituições distintas. Assim, a busca baseada exclusivamente nos nomes retornará todos os pesquisadores com o nome digitado. O comentário a seguir reflete a dificuldade desse tipo de busca, indicando a demanda por técnicas específicas para a desambiguação de nomes de autores [Ferreira et al., 2020]: *“Parabéns pela ideia! Excelente contribuição com a ciência brasileira. O único problema que encontrei foi que quando digitei meu nome ****, apareceram 3 deles e cada um com informação diferente. [...]”*.

Alguns respondentes não perceberam a existência das abas Dados do Currículo Lattes e Métricas, e conseqüentemente não exploraram a navegação por meio delas. Em razão disso, foi sugerido por cinco respondentes uma melhoria que tornasse essas abas mais visíveis e, conseqüentemente, mais fáceis de serem utilizadas. Por fim, foi sugerido, por quatro respondentes, melhorar a visibilidade dos filtros de busca disponíveis na plataforma. Esta sugestão está relacionada à dificuldade dos respondentes em encontrar tais filtros, conforme indicado pelo seguinte comentário de um dos respondentes: *“quando o usuário seleciona um pesquisador, o sistema atualiza o painel abaixo da árvore e o usuário nem percebe. [...]”*.

Também foram apresentadas varias sugestões para ampliar o escopo da plataforma *Árvore da Ciência*, as quais estão sumarizadas em quatro categorias conforme a Figura 4.17. Tais sugestões podem ser entendidas como estratégias para uma melhor caracterização da produção científica dos pesquisadores, que poderiam ser consideradas para análise futura da produção desses pesquisadores utilizando a plataforma *Árvore da Ciência*. Como pode ser observado, as duas primeiras sugestões, “Adicionar dados sobre publicações” e “Melhorar a frequência de atualização dos dados”, foram indicadas por sete respondentes para cada uma. Entretanto a primeira foge ao escopo atual da plataforma que visa particularmente registrar a genealogia acadêmica dos pesquisadores brasileiros, enquanto que a segunda será contemplada em atualizações regulares a partir de novas coletas na plataforma Lattes.

A terceira sugestão solicitada por quatro respondentes consiste no cruzamento de dados com outras plataformas como, por exemplo, a *PubMed* e a *Academic Tree*, conforme mencionado no seguinte comentário: “*Esperava maior cruzamento de informações.*”. Entretanto, tais plataformas possuem um escopo mais abrangente que foge aos objetivos da plataforma *Árvore da Ciência*, que tem como foco a comunidade acadêmica brasileira.

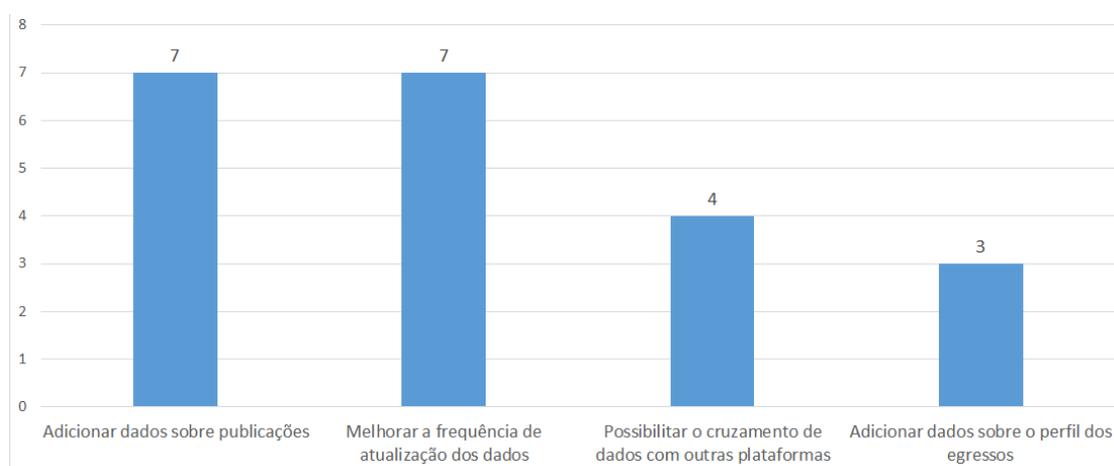


Figura 4.17. Sugestões propostas para ampliação do escopo da plataforma *Árvore da Ciência*.

Foi sugerido, também, o monitoramento dos egressos dos programas de pós-graduação de modo a melhor caracterizar a trajetória acadêmica dos pesquisadores brasileiros, conforme evidenciado pelo seguinte comentário: “*Sugiro divulgar junto à CAPES e coordenadores de programas de pós-graduação pois permite acompanhar de maneira fácil a trajetória acadêmica dos egressos da pós-graduação, nova métrica solicitada pela CAPES aos programas. [...]*”.

Finalmente, vale ressaltar que várias sugestões dos respondentes para ampliação do escopo da plataforma focaram nas publicações dos pesquisadores e no cruzamento de dados da plataforma, representando um esforço adicional que não seria relevante considerando o contexto da genealogia acadêmica. Assim os comentários referentes à ampliação do escopo da plataforma revelam que o foco na orientação de pós-graduação talvez tenha sido estreito e que esta deficiência poderia ser contornada no futuro coletando dados adicionais dos próprios currículos Lattes e oferecendo opções para exploração da ascendência acadêmica. Entretanto, vale ressaltar que a expansão dessa funcionalidade deverá considerar as características específicas de cada tipo de orientação, de modo a não prejudicar a visualização e o caminhar pelas árvores.

4.1.5.3 Problemas Relatados

Nesta categoria estão os comentários que apontam problemas de interação enfrentados pelos usuários durante a utilização da plataforma. Nela estão incluídas dificuldades relacionadas à utilização de funções específicas da plataforma, bem como problemas enfrentados com o navegador utilizado e na expansão das árvores genealógicas acadêmicas.

Em relação a problemas com diferentes dispositivos, foi solicitado que a navegação seja melhorada em vários deles, como *laptops* que utilizam *touchpad*, principalmente os da marca Apple devido à sua sensibilidade, conforme comentou um dos respondentes (“*No notebook apple tive problemas para aumentar e entrar nos boxes. [...]*”). Foi sugerido também melhorar a visualização da plataforma em *smartphones* e *tablets*. Como a plataforma *Árvore da Ciência* não foi originalmente desenvolvida para esses dispositivos, alguns respondentes relataram dificuldades para a sua utilização nesses casos, como relatado pelo seguinte comentário: “*Visualizacao por smart phone e ruim. A arvore mostrada e bastante limitada e pouco informativa (sic)*”.

Também foram reportadas algumas árvores geradas com defeitos. Por exemplo, em alguns desses casos as árvores deveriam mostrar orientações que não foram adicionadas a elas. Outras estavam com a sua visualização incorreta, resultando na solicitação da melhoria do algoritmo de geração das árvores. Com relação a alguns desses casos, os comentários mencionavam explicitamente o nome dos pesquisadores, de modo que a partir deles foi possível analisar caso-a-caso. Assim, para alguns casos analisados o problema não era especificamente na visualização das árvores, mas nos dados lançados nos respectivos currículos Lattes.

Em relação aos problemas relatados, pode observar que eles foram reportados pelos participantes, conforme ilustrado no gráfico da Figura 4.18. De modo que 28

participantes informaram sobre árvores com visualização incorreta, 10 participantes relataram a dificuldade em utilizar a plataforma efetuando o uso de dispositivos diversos como *smartphones* e 4 relataram a dificuldade de utilizar a plataforma em diversos navegadores de *internet*. Vale ressaltar que melhorias na geração das árvores estão sendo realizadas para evitar os problemas de visualização, minimizar os problemas na utilização de outros dispositivos e aumentar a compatibilidade com outros navegadores de *internet*.

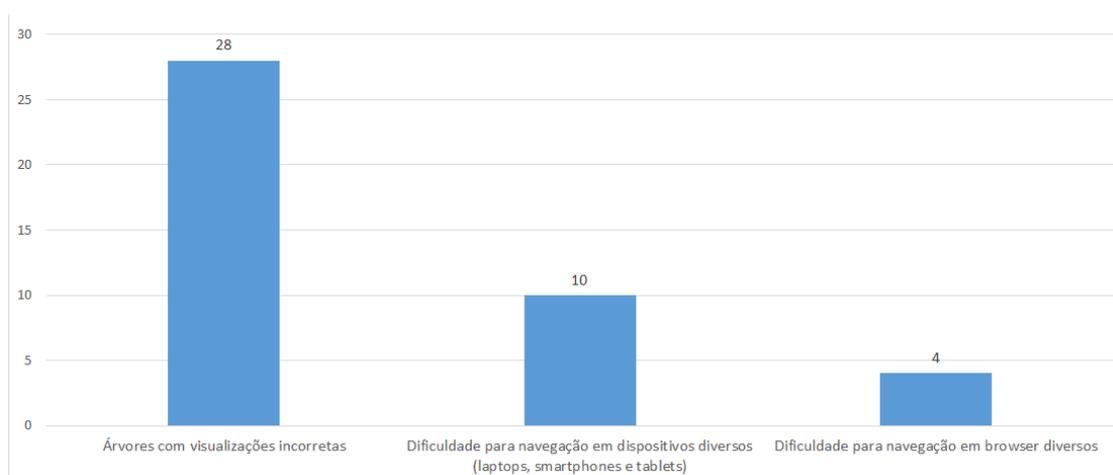


Figura 4.18. Problemas relatados sobre a plataforma Árvore da Ciência.

4.1.5.4 Comentários Adicionais

Por fim, alguns respondentes deixaram comentários adicionais que foram agrupados em duas categorias, considerações sobre a utilidade da plataforma e comentários gerais, conforme caracterizados no gráfico da Figura 4.19. Em relação às considerações sobre a utilidade da plataforma, oito respondentes a consideraram de pouca utilidade, enquanto cinco a consideraram útil. Em relação aos comentários gerais, 23 respondentes parabenizaram a iniciativa da plataforma, 13 mencionaram algum tipo de dúvida e apenas dois fizeram alguma referência a outro sistema ou plataforma similar.

Vale ainda ressaltar que poucos respondentes expressaram algum tipo de dificuldade ao utilizar a plataforma Árvore da Ciência, embora alguns possam eventualmente não ter compreendido a sua finalidade, tendo em vista o questionamento de um respondente específico (*“Qual é a utilidade de ver árvores genealógicas de pesquisadores brasileiros? A história é recente, a quantidade de mestrados é exagerada, e várias vezes nem tem pesquisa envolvida. [...]”*). Por outro lado, vários respondentes consideraram a iniciativa bastante atraente e útil, como o que deixou o seguinte comentário sobre

a plataforma: “é muito útil para nós pesquisadores sabermos os desdobramentos dos nossos orientados. [...]”.

Entre os comentários gerais feitos sobre a plataforma, foram relatadas algumas dúvidas específicas que não se encaixavam em nenhum dos tipos de comentário já citados anteriormente. Houve também referências a iniciativas semelhantes e parabenizações por parte dos respondentes que consideraram o desenvolvimento da plataforma *Árvore da Ciência* uma iniciativa positiva, como é o caso, por exemplo, do comentário “*Fiquei sensibilizada ao ver minha trajetória acadêmica sintetizada de forma tão clara, inédita, estética e concisa. [...]*”.

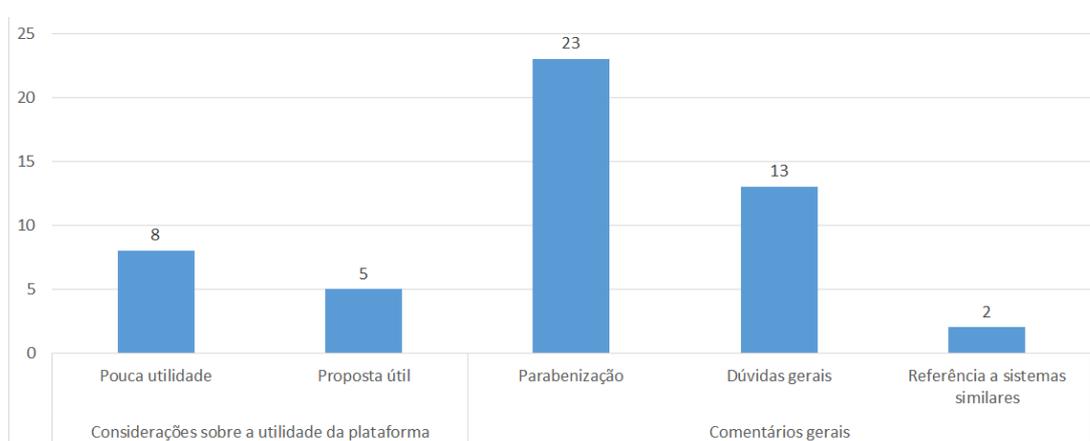


Figura 4.19. Comentários adicionais sobre a plataforma *Árvore da Ciência*.

4.2 Avaliação por Entrevistas

Para esta segunda etapa de avaliação, foram selecionados participantes que possuíssem algum envolvimento com pesquisa e orientação na pós-graduação. Assim, foram enviados convites para 27 pesquisadores de diversas áreas do conhecimento e diferentes níveis da carreira acadêmica (recém contratados, seniores e titulares). O *e-mail* enviado continha uma breve explicação do processo de avaliação, bem como uma pequena descrição da plataforma *Árvore da Ciência*. Deste total de 27 *e-mails* enviados, foram obtidas sete respostas positivas para avaliar a plataforma por meio de uma entrevista. Em seguida, foi agendado o local e o horário para a realização da entrevista, durante a qual eles teriam de executar algumas tarefas utilizando a plataforma.

O material utilizado na avaliação incluiu um *laptop*, o termo de compromisso a ser assinado pelo participante, a lista de tarefas a serem executadas e a lista de questões que seriam colocadas para os participantes. As entrevistas ocorreram no período de

Fevereiro a Março de 2020 nas instalações da Universidade Federal de Minas Gerais (UFMG) e para avaliação da plataforma foi utilizado um computador com capacidade para gravar a tela e o áudio dos respondentes. O tempo de duração das entrevistas solicitado no *e-mail* foi de uma hora.

Como parte do protocolo adotado nesse tipo de avaliação, foi elaborado e entregue a cada entrevistado um termo de consentimento (Apêndice A) com o objetivo de resguardar o seu anonimato e também para obter o seu consentimento para que os dados produzidos durante a entrevista pudessem ser incluídos na avaliação de usabilidade da plataforma *Árvore da Ciência*. O termo continha também a informação de que todo o processo seria gravado por meio de áudio e que os movimentos realizados na tela da plataforma seriam capturados. A assinatura do termo foi indispensável para a realização da entrevista.

4.2.1 Perfil dos Participantes

Os sete participantes desta avaliação eram professores, com idades entre 30 e 60 anos, de duas instituições públicas de ensino superior, a Universidade Federal de Minas Gerais (UFMG) e a Universidade Federal de Ouro Preto (UFOP), todos eles com envolvimento em atividades de pesquisa. Desses participantes, seis são professores titulares e possuem em média 20 orientações concluídas em programas de pós-graduação. As especialidades dos participantes estão relacionadas às áreas de engenharia elétrica, matemática, física, sociologia, biologia e comunicação. O tempo médio das entrevistas foi de 37 minutos. Dois dos participantes relataram terem acessado a plataforma *Árvore da Ciência* antes da entrevista. Esta foi uma situação interessante, pois permitiu não só avaliar mais amplamente a usabilidade da plataforma sem a presença do desenvolvedor, mas também sanar dúvidas relativas tarefas que outros participantes não conseguiram realizar sozinhos. Durante a entrevista, apenas um participante encontrou dificuldade em realizar as tarefas propostas. A análise das entrevistas com esses sete participantes será feita de forma anônima, de modo que eles serão identificados por P1, P2, P3, P4, P5, P6 e P7.

4.2.2 Tarefas e Perguntas

Com o objetivo de verificar a usabilidade da plataforma *Árvore da Ciência*, foi preparado um roteiro contendo as diretrizes e temas a serem explorados junto com os entrevistados. Além disso, foram elaboradas quatro tarefas envolvendo o uso da plata-

forma. O roteiro iniciava com a apresentação do projeto de pesquisa e da plataforma, e informava o objetivo desta etapa de avaliação. Outro ponto contido no roteiro foi a proposta de um diálogo constante durante toda a entrevista, no qual o participante poderia fazer qualquer pergunta sobre a plataforma e a qualquer momento.

Durante a entrevista, foram preenchidas duas vias do termo de consentimento, tendo uma sido entregue ao participante e a outra anexada à documentação necessária para a realização desse tipo de experimento. As tarefas propostas foram entregues uma a uma aos participantes para evitar que eles ficassem ansiosos quanto ao tempo do processo de avaliação e à quantidade de tarefas. Os pontos de maior relevância, na perspectiva do entrevistador, foram anotados e utilizados para embasar a discussão do processo avaliativo.

Após a execução das tarefas, os participantes foram submetidos a um questionário semi-estruturado, diferente do utilizado anteriormente, com o intuito de avaliar a sua perspectiva em relação à usabilidade e apresentação da plataforma. A seguir são detalhadas as tarefas e as perguntas feitas durante a entrevista.

Tarefa 1

Você está participando do comitê do Congresso da Associação Brasileira de Engenharia Agrícola (SBEA) responsável pela homenagem aos pesquisadores que receberão prêmios durante o evento. Este ano o premiado será o Professor Wagner Bettiol da Embrapa Meio Ambiente. Em uma conversa recente, um colega lhe falou sobre a plataforma *Árvore da Ciência* que permite a visualização gráfica de dados do Currículo Lattes de um pesquisador. Assim, você resolveu utilizar essa plataforma para obter algumas informações sobre o professor que o ajudem a preparar a homenagem. Além disso, você gostaria de chamar para participar da homenagem ao Professor Wagner Bettiol o seu primeiro orientado de doutorado e também o mais recente. Assim, registre no formulário que lhe foi entregue: (1) O nome desses dois orientados (um antes de 2000 e o outro após 2010); (2) Até três fatos específicos que você tenha obtido na plataforma e que seria interessante divulgar sobre o professor durante a homenagem.

Esta primeira tarefa tinha como objetivo verificar se o participante conseguiria realizar buscas na plataforma e quais estratégias utilizaria. A tarefa de encontrar o nome de dois orientados entre os anos de 2000 e 2010 pode ser realizada ao passar com o cursor pelos nós da árvore, ou pela lista de orientações na aba de Dados do

Currículo Lattes. Um outro objetivo da tarefa era descobrir quais seriam os dados mais interessantes disponibilizados pela plataforma do ponto de vista do participante.

Tarefa 2

Você está participando do comitê de avaliação dos pesquisadores do CNPq e foi feita uma distribuição de tarefas entre os seus membros, ficando você incumbido de realizar uma análise sobre as orientações de cada pesquisador nos últimos cinco anos. Agora que já conhece a plataforma Árvore da Ciência, você resolveu usá-la como fonte de dados para analisar tais pesquisadores. Assim, para que possa elaborar o parecer sobre o pesquisador Dermeval Saviani da PUC-SP, registre a seguir os seguintes dados sobre ele: (1) Quantidade de orientações de mestrado ou doutorado por ele concluídas (de 1995 a 2000); (2) Número médio de orientações de pós-graduação por ano; (3) Total de seus orientados de doutorado que já orientaram pelo menos um aluno de mestrado ou doutorado.

A segunda tarefa teve como objetivo avaliar a capacidade do participante em explorar os diversos recursos da plataforma, em particular as métricas relativas às árvores genealógicas acadêmicas. Dessa maneira, o participante deveria acessar os últimos cinco anos do “gráfico de orientações por ano” localizado na aba Métricas e calcular a quantidade de orientações concluídas no período de 1995 e 2000. Além disso, o usuário deveria identificar a métrica que expressa a quantidade média de orientações concluídas por ano na pós graduação e informar quais orientandos do pesquisador em questão já possui algum descendente acadêmico. Para obter esse resultado, o participante deveria verificar se a métrica de fertilidade do pesquisador era diferente de zero e se algum dos nós de sua árvore estava marcado com o símbolo de expansão ([+]).

Tarefa 3

Você esteve presente à palestra da professora Sabrina que comentou sobre o trabalho de um aluno dela. No entanto, como você chegou atrasado à palestra, não sabe o sobrenome da professora ou qual é a sua instituição, nem o nome do aluno. Entretanto, você sabe que a professora Sabrina foi orientada pelo Prof. Wagner Meira Júnior da UFMG e que o trabalho do aluno é denominado GREMLIN. Assim, você acessa a plataforma Árvore da Ciência e busca pela árvore do Prof. Wagner Meira Júnior para ver se identifica o trabalho em questão. Registre o nome do autor do trabalho

Esta terceira tarefa teve por objetivo observar a navegação de um usuário por uma árvore genealógica acadêmica. Assim, foi proposto que a partir de um pesquisador inicial, o usuário encontrasse um de seus “netos” acadêmicos. Esta informação é obtida ao se expandir o nó do pesquisador em questão, de modo a obter a sua sub-árvore. Outra maneira de se obter a mesma informação é clicar no nó da orientação e descobrir qual é o orientado pela aba de “Dados do Currículo Lattes”.

Tarefa Livre

Explore a sua própria árvore ou a de alguém que você conheça bem e analise os pontos que achar interessantes sobre a árvore em questão (algo livre mesmo). Lembre-se que através da engrenagem de configuração você pode configurar alguns aspectos da árvore. Por favor, explique o que fez e porque à medida que for explorando a plataforma.

Por último, na tarefa de exploração livre foi solicitado ao participante que, se possível, verbalizasse a ação prevista antes dela ser realizada [Barbosa & Silva, 2010]. Os pontos de maior interesse dos usuários para execução dessa tarefa variaram entre a exploração das árvores e a visualização das métricas.

Ao final das tarefas os participantes foram entrevistados sobre a sua percepção e experiência com a plataforma *Árvore da Ciência*. Para isso, foi utilizado um roteiro com sete perguntas que buscaram: (a) determinar a intuitividade da plataforma para o usuário, (b) entender quais são os pontos da plataforma que precisam ser melhorados, (c) verificar se os textos de ajuda presentes na plataforma são úteis, (d) obter sugestões para melhoria da plataforma e (e) verificar a utilidade da plataforma sob a ótica do usuário. As perguntas formuladas aos entrevistados foram:

1. Qual a sua opinião geral sobre a plataforma? Por quê? Quais são os seus pontos positivos e negativos?
2. Você teve dificuldade em realizar alguma tarefa? Qual? Por quê?
3. Sobre as métricas escolhidas na Tarefa 1, quais as que você achou interessantes? Por que você as achou interessantes? Que outras métricas você achou úteis ou não? Por quê?
4. Sobre a visualização, o que achou? Justifique. O que achou do *zoom*, da opção de expansão da sub-árvore, das linhas e cores, e dos dados sobre os nós (*tooltip*)?
5. Sobre a configuração das opções de visualização, o que achou?

6. Acredita que usaria a plataforma? Em que situações? Por quê? (ou Por que não?) Quando usaria a plataforma Árvore da Ciência e não a plataforma Lattes?
7. Existe alguma outra sugestão de melhoria para a plataforma?

4.2.3 Resultados da Avaliação

As tarefas realizadas pelos participantes e as respostas dadas por eles às sete perguntas finais foram posteriormente analisadas de maneira qualitativa, isto é, elas foram transcritas e analisadas com base no roteiro proposto. Por meio dos comentários colhidos durante a entrevista e da observação da forma como a plataforma foi utilizada pelos participantes, foi possível identificar alguns pontos que precisam ser melhorados para que a plataforma possa ser ainda mais útil no dia-a-dia dos pesquisadores.

Em relação à primeira pergunta, referente à opinião geral dos usuários sobre a plataforma, foi obtido um retorno positivo, em que todos os entrevistados consideraram a iniciativa interessante. Quanto aos pontos positivos, os participantes P1, P2, P3, P4 e P6 afirmaram ter gostado da plataforma, enquanto o participante P1 fez comentários específicos demonstrando a sua admiração com a visualização das árvores e das métricas. O participante P7 mencionou que, embora possa haver alguma resistência por se tratar de uma novidade, ele considerou a plataforma Árvore da Ciência atrativa, possuindo recursos visuais interessantes de serem explorados. As métricas consideradas mais relevantes pelos participantes foram a *fertilidade* (quantidade de orientados que possuem orientações concluídas), o *caminho para a orientação mais profunda* e o *número de descendentes diretos e indiretos*, além daquelas apresentadas no gráfico de orientações por ano. Sobre os pontos negativos, o participante P1 destacou a dificuldade em perceber as abas com dados adicionais (DADOS DO CURRÍCULO LATTES e MÉTRICAS), enquanto que os participantes P2, P3, P4 e P5 destacaram que a visualização das árvores poderia ser melhorada, principalmente em relação à sua geometria, uma vez que ela não favorece a visualização de grandes árvores, pois as arestas que indicam os relacionamentos ficam sobrepostas. Outro ponto negativo levantado foi a ausência de instruções sobre os comandos possíveis de serem utilizados na plataforma, apesar de existir uma aba específica onde esses comandos são descritos que, aparentemente, não foi percebida por alguns usuários.

Em relação à segunda pergunta, que mede a dificuldade ao realizar uma tarefa na plataforma, foi observado que os participantes P1 e P6 consideraram a segunda tarefa mais complicada, pois não perceberam a aba de métricas. Por outro lado, o participante P5 encontrou dificuldade na Tarefa 1, pois optou por encontrar os orientados navegando pelos nós da árvore. No entanto, ao perceber que a mesma informação

poderia ter sido obtida a partir da aba de DADOS DO CURRÍCULO LATTES, o participante P5 considerou a tarefa fácil de realizar. Os demais participantes não apontaram dificuldades em realizar as quatro tarefas propostas.

A terceira pergunta avaliou o motivo da escolha das métricas na Tarefa 1 e também a utilidade de outras métricas do ponto de vista do participante. O participante P1 relatou não ter encontrado as métricas sozinho, sendo necessária a interferência do entrevistador para mostrar a aba de métricas. Ao encontrá-las, o participante P1 apontou como interessantes o número médio de orientações concluídas, os títulos das teses e dissertações orientadas, e o nível de formação dos pesquisadores, embora esses dois últimos não sejam efetivamente métricas. O participante P2 escolheu para esta tarefa a métrica profundidade das árvores que mostra todas as relações de descendência. A escolha de P2 foi justificada por mostrar o impacto na formação direta ou indireta de novos pesquisadores. Já o participante P3 considerou muito relevante mostrar a árvore de um pesquisador com dados de suas orientações de pós-graduação. Os participantes P4, P5 e P6 sugeriram a implementação de métricas relacionadas às colaborações e publicações dos pesquisadores, embora tais métricas fujam do escopo proposto para a plataforma Árvore da Ciência. Por fim, as métricas melhor avaliadas pelo participante P7 foram a fertilidade e a quantidade total de descendentes diretos e indiretos. Esta terceira pergunta mostrou que a métrica fertilidade foi a mais indicada pelos participantes ouvidos e que métricas relacionadas à colaboração em trabalhos acadêmicos deveriam ser implementadas a fim de enriquecer a plataforma, embora tais métricas não sejam diretamente associadas à genealogia acadêmica.

Em relação à quarta pergunta que tinha o objetivo de analisar outros aspectos relacionados às facilidades de visualização da plataforma, como *zoom*, expansão das árvores, cores das linhas que representam as orientações e o uso do *tooltip* (balão com textos que aparece ao passar o cursor em cima dos nós dos pesquisadores), o participante P1 considerou tais facilidades interessantes, com as cores bem explicativas, e sugeriu aumentar o tamanho do texto da fonte dos *tooltips*. O participante P2 considerou que as cores são legíveis e achou que o *tooltip* ajudou a obter outros detalhes sem a necessidade de utilizar a aba DADOS DO CURRÍCULO LATTES, embora tenha relatado ter demorado um pouco a perceber esse recurso. Os participantes P3 e P7 consideraram que a visualização e a expansão de grandes árvores precisam ser melhoradas, enquanto o participante P4 solicitou instruções para detalhar melhores maneiras de interagir com as árvores, por exemplo, como utilizar o *scroll* do *mouse* para fazer o *zoom* nas árvores. Em relação ao tamanho dos *tooltips*, foi mencionado que a fonte dos respectivos textos era pequena, de modo que ela foi aumentada para melhorar a sua visualização.

A quinta pergunta estava relacionada às configurações que podem ser definidas para exibição das árvores. O participante P1 relatou que, embora essas configurações provocassem alguns erros na exibição das árvores, como a perda da ordenação dos nodos, elas foram úteis, principalmente na realização da Tarefa 3. Entretanto, outros participantes relataram que a aba de configuração das opções de visualização não era facilmente visível, o que teria dificultado a realização de algumas tarefas propostas. Os participantes P3 e P4 sugeriram ainda que o padrão de ordenação dos nós fosse cronológico. Entretanto, vale ressaltar que essa aba foi um recurso implementado de maneira experimental, de modo que o seu uso resultou em alguns problemas que precisam ser sanados. Além disso, o local onde essa aba foi incluída mostrou-se inadequado, por não ser facilmente visível. Assim, esta é uma funcionalidade que carece de uma reavaliação para ser implementada de maneira definitiva na plataforma *Árvore da Ciência*.

A sexta pergunta se referia à intenção e situações em que o usuário utilizaria a plataforma. Além disso, foi investigado o quanto a plataforma *Árvore da Ciência* é útil em relação à plataforma Lattes. Os participantes P1, P2, P3, P5 e P6 consideraram a plataforma interessante e, a partir do convite para participar de sua avaliação, a têm utilizando com frequência, principalmente em tarefas administrativas. Os participantes P4 e P7 elogiaram a maior eficiência da busca de pesquisadores quando comparada à da plataforma Lattes. Eles mencionaram também que caso a plataforma seja atualizada com frequência e venha a incluir também dados de publicações, ela poderia ser utilizada em substituição a outras já existentes.

A sétima pergunta teve por objetivo obter sugestões de melhoria ou de novas funcionalidades para a plataforma. Todos os participantes solicitaram a inclusão de dados sobre pesquisas em andamento e publicações dos pesquisadores. Alguns, como por exemplo o participante P5, sugeriram que fossem implementados outros tipos de ajuste, como um controle de *zoom* para uma melhor visualização as árvores. O participante P6 sugeriu ainda a exportação de dados, como a lista de orientações de um pesquisador, para serem utilizados por outras aplicações. Por último, o participante P7 sugeriu a implementação de formas diferentes de visualização das árvores, ou pelo menos que o formato das árvores pudesse ser alterado de forma a possibilitar ao usuário a escolha daquele que lhe parecesse mais apropriado. O participante P7 sugeriu também a inclusão explícita na plataforma de dados sobre a formação dos pesquisadores, como por exemplo, as instituições em que foram realizados os seus cursos de pós-graduação. Vale ressaltar que todas as sugestões apresentadas pelos participantes são pertinentes e poderão ser consideradas em futuras versões da plataforma *Árvore da Ciência*.

4.3 Considerações Finais

Os resultados apresentados neste capítulo revelam uma convergência entre a avaliação por questionário e a avaliação por entrevista. Em relação às limitações da plataforma *Árvore da Ciência*, foi destacado em ambas avaliações que a visualização das árvores genealógicas acadêmicas precisa ser melhorada. Este fato foi evidenciado nos comentários feitos pelos entrevistados e também nos comentários deixados nos questionários pelos respondentes, relatando dificuldades para a expansão das árvores com grande número de orientações. Outras limitações destacadas foram a ausência de um texto de ajuda detalhado sobre as métricas e a pouca visibilidade dos filtros que acabaram quase não utilizados. Já o ponto positivo de maior destaque foi a facilidade de utilização da plataforma, reconhecida por 85,2% dos respondentes do questionário e por 5 dos 7 entrevistados. Outro ponto de destaque foi a apresentação dos dados, com uma taxa de aprovação de 65,3% dos respondentes do questionário e de 4 dos 7 entrevistados. Além disso, não foram observadas maiores divergências entre os resultados oriundos da avaliação por questionário e aqueles providos pela avaliação por entrevista.

Com o objetivo de melhor caracterizar os respondentes e relacionar os dados coletados pelo questionário, foi observado que a maioria deles (80,2%) possui doutorado e exerce a função de professor universitário ou pesquisador em alguma instituição pública de ensino superior. Em relação à formação acadêmica, todos os livre-docentes são professores universitários ou pesquisadores. Quanto aos mestres, 50,0% deles estão cursando doutorado e os outros 50,0% são professores universitários. Já em relação à área de formação, 54,2% dos que atuam na área de biologia possuem doutorado e 14,4% realizaram pós-doutorado. Por outro lado, 90,8% dos que possuem formação em engenharia são doutores. Sobre os entrevistados, quatro são homens e três são mulheres, desses seis são doutores e apenas um único participante é mestre, estando cursando o doutorado. Em relação aos entrevistados, todos eram professores atuantes em pesquisa e filiados a universidades federais. Vale ainda ressaltar que as duas avaliações ocorreram simultaneamente e, por esse motivo, não houve tempo de uma influenciar a outra.

A conclusão obtida a partir dessas duas avaliações é que, embora possua pontos a serem melhorados, a plataforma *Árvore da Ciência* é de fácil utilização e se apresenta como uma ferramenta relevante para analisar a formação de novos pesquisadores no Brasil. É importante salientar ainda que algumas limitações levantadas pelos usuários já foram resolvidas, como a inclusão de uma página de explicação das métricas e o aumento do tamanho da fonte dos textos de ajuda. Entretanto, ainda há alguns aspectos específicos da plataforma que precisam ser melhorados como, por exemplo,

a visualização de árvores com muitos níveis e grande número de nodos (orientados), e o posicionamento de determinados elementos na tela, como os filtros utilizados em determinados tipos de consulta.

Total		Usuários		Duração média da sessão	
		627		00:04:57	
Estado	Quantidade de Usuários	Duração média de acesso	Estado	Quantidade de Usuários	Duração média de acesso
São Paulo	178 (28,39%)	00:02:34	Paraíba	10 (1,59%)	00:03:07
Minas Gerais	138 (22,01%)	00:08:28	Goiás	9 (1,44%)	00:02:14
Rio de Janeiro	93 (14,83%)	00:04:04	Rio Grande do Norte	8 (1,28%)	00:09:13
Rio Grande do Sul	40 (6,38%)	00:03:38	Mato Grosso	6 (0,96%)	00:05:51
Paraná	27 (4,31%)	00:01:38	Pará	6 (0,96%)	00:00:18
Santa Catarina	20 (3,19%)	00:04:42	Alagoas	4 (0,64%)	00:03:59
Distrito Federal	18 (2,87%)	00:02:30	Mato Grosso do Sul	4 (0,64%)	00:03:35
Bahia	15 (2,39%)	00:04:59	Sergipe	4 (0,64%)	00:01:28
Ceará	15 (2,39%)	00:04:27	Rondônia	3 (0,48%)	00:03:54
Pernambuco	14 (2,23%)	00:03:04	Piauí	1 (0,16%)	00:00:26
Amazonas	13 (2,07%)	00:01:47	Roraima	1 (0,16%)	00:03:40

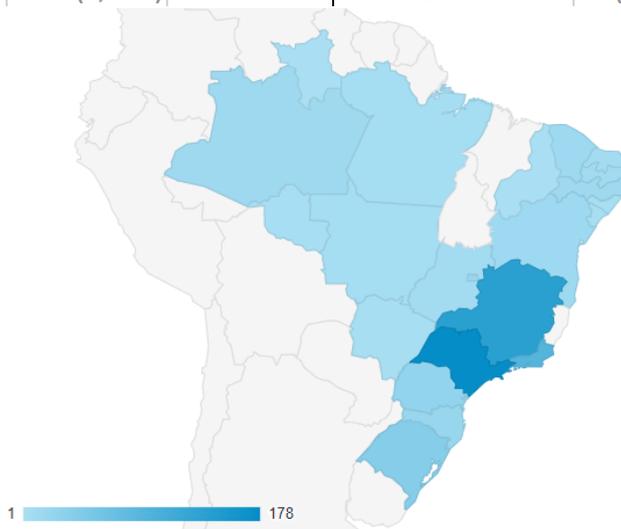


Figura 4.20. Acessos por estado no período de 1 de junho de 2019 a 1 de novembro de 2020. Dados coletados do serviço de estatísticas de tráfego do *Google Analytics*.

Com o objetivo de verificar como tem sido o acesso à plataforma *Árvore da Ciência* nas diferentes regiões do país, foi criada por meio da ferramenta *Google Analytics*⁶ uma página que provê estatísticas atualizadas sobre a sua utilização. A Figura 4.20 mostra a quantidade de acessos à plataforma em cada estado do Brasil entre junho de 2019

⁶<http://analytics.google.com/>, Acesso em Novembro de 2020.

e novembro de 2020, período em que a plataforma está disponível publicamente, mas sem qualquer publicidade. Foram constatados 627 usuários distintos que acessaram a plataforma com uma média de uso de quatro minutos e 27 segundos, indicando que há usuários navegando pela plataforma. Os três estados com maior número de acessos até o momento foram São Paulo, com 178 (28,4%) acessos, Minas Gerais com 138 (22,0%) e Rio de Janeiro com 93 (14,8%). Por outro lado, não foram registrados quaisquer acessos nos estados do Acre, Espírito Santo, Maranhão e Tocantins.

Capítulo 5

Conclusões e Trabalhos Futuros

5.1 Conclusões

Nos últimos anos, tem havido um grande esforço visando o desenvolvimento de plataformas e aplicações específicas voltadas para a construção e manutenção de árvores genealógicas acadêmicas de pesquisadores das mais diversas áreas do conhecimento [Damasceno et al., 2019; David & Hayden, 2012; Dores et al., 2017; Hart & Cossuth, 2013; Jackson, 2007; Rossi et al., 2019]. Em sua maioria, essas iniciativas adotam uma estratégia de colaboração coletiva, em que tais plataformas ou aplicações crescem com dados inseridos por diversos colaboradores.

Nesta dissertação, foi realizado um esforço no sentido de desenvolver uma plataforma para visualização e exploração da genealogia acadêmica dos pesquisadores brasileiros. Para isso, foram utilizados dados de orientações, extraídos dos currículos dos pesquisadores com título de doutor, coletados da plataforma Lattes no formato XML. Diferentemente de outras iniciativas similares, como o *Mathematics Genealogy Project*, que se limitam apenas a uma área do conhecimento, a abordagem utilizada nesta dissertação permitiu construir árvores genealógicas acadêmicas de pesquisadores de diversas áreas do conhecimento.

A plataforma *Árvore da Ciência*, entretanto, não se limita apenas a exibir as árvores geradas para os pesquisadores, mas também possibilita caminhar por elas à medida que são expandidas com os nós de seus descendentes acadêmicos. A plataforma provê ainda métricas específicas relacionadas às árvores geradas e às orientações correspondentes, além de exibir dados de interesse extraídos dos currículos Lattes dos respectivos pesquisadores.

Assim, por meio da plataforma *Árvore da Ciência* é possível visualizar o papel dos nossos pesquisadores na formação acadêmica brasileira, bem como dos seus ori-

entados a partir de novas orientações. Muitos desses pesquisadores são pioneiros em suas áreas, como o Prof. André Dreyfus, considerado o pai da genética no Brasil e um dos fundadores da USP, cuja descendência possui nove gerações que contabilizam 6328 descendentes diretos e indiretos. Um outro exemplo importante é o da Profa. Annita Castilho, um dos grandes nomes da área de Psicologia no país e criadora do curso de Psicologia da USP, cuja descendência possui sete gerações de descendentes acadêmicos, incluindo 8917 descendentes diretos e indiretos. Joel Martins, também um dos grandes nomes da área de Educação do país, possui 11 gerações de descendentes acadêmicos e 15091 descendentes diretos e indiretos. Finalmente, o Prof. Crodowaldo Pavan, cuja árvore foi abordada na Seção 3.4, possui oito gerações de descendentes acadêmicos que somam 6327 descendentes diretos e indiretos.

A utilização do *framework* *.Net Core* e das bibliotecas visuais *ReactJS* e *VisJS*, em conjunto com o SGBD *MariaDB*, possibilitaram um melhor desempenho da plataforma ora desenvolvida em relação à sua versão inicial implementada a partir do trabalho de Dores [2017] e que apenas permitia a visualização individual das árvores geradas. Além disso, a nova interface também teve um impacto positivo na experiência de interação dos usuários, principalmente devido ao emprego de modernas técnicas de desenvolvimento de aplicações *web*, o que possibilitou um ganho de desempenho e uma melhor visualização dos dados por parte dos usuários.

Neste contexto, a plataforma *Árvore da Ciência*, disponibilizada por meio de um portal *online*¹, permite a visualização, navegação e exploração das árvores genealógicas acadêmicas dos pesquisadores brasileiros, além de prover métricas que quantificam o impacto desses pesquisadores como orientadores de pós-graduação. Além disso, vale ressaltar a sua avaliação realizada com dois grupos de usuários, o primeiro deles composto por 286 pesquisadores que responderam a um questionário contendo 19 questões de múltipla escolha e uma questão aberta, e o segundo, composto por 7 pesquisadores com experiência em pesquisa e orientação de pós-graduação, que aceitaram participar de uma avaliação presencial por meio de uma entrevista, durante a qual eles tiveram de executar algumas tarefas utilizando a plataforma. Os resultados dessas duas avaliações com usuários típicos possibilitaram não só a validação das principais facilidades oferecidas pela plataforma, como também a identificação de novos recursos que poderiam ser adicionados a ela.

¹<https://sciencetree.net/>, Acesso em Dezembro de 2020.

5.2 Trabalhos Futuros

Apesar de a plataforma *Árvore da Ciência* já estar disponível para ampla utilização, tendo sido acessada de forma espontânea por usuários de diversos estados do país, como mostram os dados da Figura 4.20, a avaliação realizada com usuários de diferentes perfis levantou alguns pontos importantes que devem ser levados em consideração para futuros melhoramentos. Como exemplo, podemos citar a necessidade de adicionar textos complementares e mais detalhados que contenham explicações acerca das métricas implementadas. Outra perspectiva futura é a implementação de métricas relacionadas não apenas às árvores dos pesquisadores, mas também às suas orientações, como o cálculo e a exibição dos caminhos mais longos de suas respectivas árvores. A identificação da instituição e das respectivas áreas de atuação dos pesquisadores também seria relevante para melhor caracterizá-los. A visualização das árvores também é um ponto que carece de melhoramento para proporcionar uma exploração mais intuitiva dos seus vários níveis, tanto descendentes quanto ascendentes.

Ainda relacionado à arquitetura da plataforma, faz-se necessária a atualização de forma automática do seu repositório de dados a partir de coletas periódicas realizadas não só na plataforma Lattes, mas também de outros repositórios que possam ser considerados no futuro, tornando assim a plataforma ainda mais atrativa para a comunidade científica brasileira. Outro ponto a ser mencionado é a melhoria do algoritmo de geração das árvores [Dores et al., 2017] a fim de tornar mais efetiva a desambiguação de pesquisadores com nomes semelhantes ou que tenham os seus nomes registrados na plataforma Lattes com diferentes grafias, o que pode ser realizado usando alguns dos métodos desenvolvidos no próprio Laboratório de Bancos de Dados da UFMG [Ferreira et al., 2020]. Finalmente, vale ressaltar a importância de se ampliar a divulgação da plataforma a fim de aumentar a sua utilização e a demanda por novas funcionalidades.

Referências Bibliográficas

- Abramo, G.; D'Angelo, C. A. & Di Costa, F. (2018). The effect of multidisciplinary collaborations on research diversification. *Scientometrics*, 116(1):423–433.
- Adams, J. (2012). Collaborations: The rise of research networks. *Nature*, 490(7420):335–336.
- Albaum, G. (1997). The likert scale revisited. *Market Research Society. Journal.*, 39(2):1–21.
- Almeida, E. C. E. & Guimarães, J. A. (2017). *A pós-graduação e a evolução da produção científica brasileira*. Senac, São Paulo.
- Anil, S.; Kurian, A.; Dey, S. R.; Saha, S. & Sinha, A. (2018). Genealogy tree: Understanding academic lineage of authors via algorithmic and visual analysis. *Journal of Scientometric Research*, 7(2):120–124.
- Araújo, E. B.; Moreira, A. A.; Furtado, V.; Pequeno, T. H. & Andrade Jr, J. S. (2014). Collaboration networks from a large CV database: dynamics, topology and bonus impact. *PloS One*, 9(3):e90537.
- Barabási, A.-L.; Jeong, H.; Néda, Z.; Ravasz, E.; Schubert, A. & Vicsek, T. (2002). Evolution of the social network of scientific collaborations. *Physica A: Statistical mechanics and its applications*, 311(3-4):590–614.
- Barbosa, S. & Silva, B. (2010). *Interação Humano-Computador*. Elsevier Editora Ltda, Rio de Janeiro, RJ, Brasil.
- Benevenuto, F.; Laender, A. H. F. & Alves, B. L. (2015). How connected are the ACM SIG communities? *SIGMOD Rec.*, 44(4):57–63.
- Bevan, N. (2009). What is the Difference Between the Purpose of Usability and User Experience Evaluation Methods. In *Proceedings of the Workshop UXEM*, pp. 1–4.

- Brown, S. S. & Robinson, P. (2001). A World Wide Web Mediator for Users with Low Vision. In *Proceedings of the Conference on Human Factors in Computing Systems (Workshop No. 14)*.
- Chang, S. (2003). Academic genealogy of american physicists. *AAPPS Bulletin*, 13(6):6–41.
- Coccia, M. & Wang, L. (2016). Evolution and convergence of the patterns of international scientific collaboration. *PNAS*, 113(8):2057–2061.
- Cunningham, S. J. (2001). The birth of a field: an analysis of the 1994-2000 ACM Digital Libraries Conferences. In *Proceedings of the 8th International Conference on Scientometrics and Informetrics*, pp. 139–146, Sydney, Australia.
- Dalmero, M. & Vieira, K. M. (2014). Dilemas na construção de escalas Tipo Likert: o número de itens e a disposição influenciam nos resultados? *Revista Gestão Organizacional*, 6(3).
- Damaceno, R. J.; Rossi, L.; Mugnaini, R. & Mena-Chalco, J. P. (2019). The Brazilian academic genealogy: evidence of advisor-advisee relationships through quantitative analysis. *Scientometrics*, 119(1):303–333.
- Damaceno, R. J. P.; Rossi, L. & Mena-Chalco, J. P. (2017). Identificação do Grafo de Genealogia Acadêmica de Pesquisadores: Uma Abordagem Baseada na Plataforma Lattes. In Hara, C. S., editor, *XXXII Simpósio Brasileiro de Bancos de Dados, Uberlandia, MG, Brazil, October 4-7, 2017*, pp. 76–87. SBC.
- David, S. V. & Hayden, B. Y. (2012). Neurotree: A collaborative, graphical database of the academic genealogy of neuroscience. *PloS One*, 7(10).
- Dawson, S.; Gašević, D.; Siemens, G. & Joksimovic, S. (2014). Current state and future trends: A citation network analysis of the learning analytics field. In *Proceedings of the Fourth International Conference on Learning Analytics and Knowledge*, pp. 231–240. ACM.
- Delgado-Garcia, J. F.; Laender, A. H. F. & Meira, W. (2014). Analyzing the Co-authorship Networks of Latin American Computer Science Research Groups. In *2014 9th Latin American Web Congress*, pp. 77–81.
- Dhar, V. (2013). Data science and prediction. *Communications of the ACM*, 56(12):64–73.

- Dias, T. M. R. (2016). *Um Estudo Sobre a Produção Científica Brasileira a partir de Dados da Plataforma Lattes*. Tese de Doutorado, Centro Federal de Educação Tecnológica de Minas Gerais, Belo Horizonte, Minas Gerais.
- Digiampietri, L.; Mugnaini, R.; Trucolo, C.; Delgado, K.; Mena-Chalco, J. & Köhler, A. (2019). Geographic and disciplinary distribution of the Brazilian's PhD community. *Brazilian Journal of Information Science: Research Trends*, 13(4):113–131.
- Digiampietri, L. A.; Mena-Chalco, J. P.; Silva, G. S.; Oliveira, L. B.; Lima, J. J. S.; Malheiro, A. P. & Meira, D. (2015). Análise da Evolução das Relações de Coautoria nos Programas de Pós-Graduação em Computação no Brasil. *Revista Eletrônica de Sistemas de Informação*, 14(1):1–23.
- Digiampietri, L. A.; Mugnaini, R.; Pérez-Alcázar, J. J.; Delgado, K. V.; Tuesta, E. F. & Mena-Chalco, J. P. (2017). Análise da evolução, impacto e formação de redes nos cinco anos do BraSNAM. In *Anais do VI Brazilian Workshop on Social Network Analysis and Mining*, pp. 492–503. SBC.
- Dores, W.; Benevenuto, F. & Laender, A. H. F. (2016). Extracting Academic Genealogy Trees from the Networked Digital Library of Theses and Dissertations. In *Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraries*, pp. 163–166, Newark, NJ, USA.
- Dores, W.; Soares, E.; Benevenuto, F. & Laender, A. H. F. (2017). Building the Brazilian Academic Genealogy Tree. In *Research and Advanced Technology for Digital Libraries - 21st International Conference on Theory and Practice of Digital Libraries, TPDL 2017, Thessaloniki, Greece, September 18-21, 2017, Proceedings*, pp. 537–543.
- Dores, W. J. (2017). *Um Estudo sobre a Genealogia Acadêmica Brasileira*. Dissertação de Mestrado, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.
- Dumas, J. S.; Dumas, J. S. & Redish, J. (1999). *A practical guide to usability testing*. Intellect books, Bristol, UK.
- Easley, D. & Kleinberg, J. (2010). *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. Cambridge University Press, Cambridge, UK.
- Ezzy, D. (2013). *Qualitative analysis*. Routledge, Abingdon, UK.
- Ferreira, A. A.; Gonçalves, M. A. & Laender, A. H. F. (2020). *Automatic Disambiguation of Author Names in Bibliographic Repositories*. Synthesis Lectures on

- Information Concepts, Retrieval, and Services. Morgan & Claypool Publishers, San Rafael, California.
- Freeman, A. (2017). *Pro ASP.NET Core MVC 2*. Apress, New York, NY.
- Hart, R. E. & Cossuth, J. H. (2013). A family tree of tropical meteorology's academic community and its proposed expansion. *Bulletin of the American Meteorological Society*, 94(12):1837–1848.
- Heinisch, D. P. & Buenstorf, G. (2018). The next generation (plus one): an analysis of doctoral students' academic fecundity based on a novel approach to advisor identification. *Scientometrics*, 117(1):351–380.
- Hirshman, B. R.; Tang, J. A.; Jones, L. A.; Proudfoot, J. A.; Carley, K. M.; Marshall, L.; Carter, B. S. & Chen, C. C. (2016). Impact of medical academic genealogy on publication patterns: an analysis of the literature for surgical resection in brain tumor patients. *Annals of Neurology*, 79(2):169–177.
- Iglič, H.; Doreian, P.; Kronegger, L. & Ferligoj, A. (2017). With whom do researchers collaborate and why? *Scientometrics*, 112(1):153–174.
- Jackson, A. (2007). A Labor of Love: The Mathematics Genealogy Project. *Notices of the AMS*, 54(8):1002–1003.
- Joshi, A.; Kale, S.; Chandel, S. & Pal, D. K. (2015). Likert scale: Explored and explained. *British Journal of Applied Science & Technology*, 7(4):396.
- Lane, J. (2010). Let's make science metrics more scientific. *Nature*, 464(7288):488–489.
- Lazar, J.; Feng, J. H. & Hochheiser, H. (2017). *Research methods in human-computer interaction*. Morgan Kaufmann.
- Lima, H.; Silva, T. H. P.; Moro, M. M.; Santos, R. L. T.; Jr., W. M. & Laender, A. H. F. (2015). Assessing the profile of top Brazilian computer science researchers. *Scientometrics*, 103(3):879–896.
- Liu, J.; Tang, T.; Kong, X.; Tolba, A.; Zafer, A.-M. & Xia, F. (2018). Understanding the advisor-advisee relationship via scholarly data analysis. *Scientometrics*, 116(1):161–180.
- Liu, X.; Bollen, J.; Nelson, M. L. & Van de Sompel, H. (2005). Co-authorship networks in the digital library research community. *Information Processing & Management*, 41(6):1462–1480.

- Liu, Y.; Dai, S.; Wang, C.; Zhou, Z. & Qu, H. (2017). Genealogyvis: A system for visual analysis of multidimensional genealogical data. *IEEE Transactions on Human-Machine Systems*, 47(6):873–885.
- Lopes, G. R.; da Silva, R.; Moro, M. M. & de Oliveira, J. P. M. (2012). Scientific collaboration in research networks: A quantification method by using gini coefficient. *Int. J. Comput. Sci. Appl.*, 9(2):15–31.
- Madeira, G.; Borges, E. N.; Lucca, G.; Santos, H. & Dimuro, G. (2020). A Tool for Analyzing Academic Genealogy. In Filipe, J.; Śmiałek, M.; Brodsky, A. & Ham-moudi, S., editores, *Enterprise Information Systems*, pp. 443–456, Cham. Springer International Publishing.
- Malmgren, R. D.; Ottino, J. M. & Amaral, L. A. N. (2010). The role of mentorship in protégé performance. *Nature*, 465(7298):622.
- Mena-Chalco, J. P.; Digiampietri, L. A.; Lopes, F. M. & Cesar, R. M. (2014). Brazilian bibliometric coauthorship networks. *Journal of the Association for Information Science and Technology*, 65(7):1424–1445.
- Montoya, F. G.; Baños, R.; Alcayde, A.; Montoya, M. G. & Manzano-Agugliaro, F. (2018). Power quality: Scientific collaboration networks and research trends. *Energies*, 11(8):2067.
- Newman, M. E. (2001). Scientific collaboration networks. ii. shortest paths, weighted networks, and centrality. *Physical review E*, 64(1):016132.
- Newman, M. E. (2004). Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences*, 101(suppl 1):5200–5205.
- Pessoa Junior, G. J.; Dias, T. M. R.; Silva, T. H. P. & Laender, A. H. F. (2020). On interdisciplinary collaborations in scientific coauthorship networks: the case of the Brazilian community. *Scientometrics*, 124(3):2341–2360.
- Prates, R. O. & Barbosa, S. D. J. (2003). Avaliação de Interfaces de Usuários - Conceitos e Métodos. In *Jornada de Atualização em Informática do Congresso da Sociedade Brasileira de Computação*, capítulo 6. SBC, Rio de Janeiro.
- Pressman, R. S. & Maxim, B. R. (2016). *Engenharia de Software - Uma Abordagem Profissional*. McGraw Hill Education, São Paulo, Brasil, 8a. edição.

- Rossi, L.; Damaceno, R. J.; Freire, I. L.; Bechara, E. J. & Mena-Chalco, J. P. (2018). Topological metrics in academic genealogy graphs. *Journal of Informetrics*, 12(4):1042–1058.
- Rossi, L.; Damaceno, R. J. & Mena-Chalco, J. P. (2019). Genealogia acadêmica: um novo olhar sobre impacto acadêmico de pesquisadores. *Parcerias Estratégicas*, 23(47):197–212.
- Rossi, L.; Freire, I. L. & Mena-Chalco, J. P. (2017). Genealogical index: A metric to analyze advisor-advisee relationships. *Journal of Informetrics*, 11(2):564–582.
- Rossi, L. & Mena-Chalco, J. P. (2014). Caracterização de Árvores de Genealogia Acadêmica por Meio de Métricas em Grafos. In *Anais do III Brazilian Workshop on Social Network Analysis and Mining*, pp. 21–32, São Paulo, SP.
- Rubin, J. & Chisnell, D. (2008). *Handbook of usability testing: how to plan, design and conduct effective tests*. John Wiley & Sons, Hoboken, New Jersey, USA.
- Sheng, S.; Zhou, P. & Wu, X. (2019). CEPV: A Tree Structure Information Extraction and Visualization Tool for Big Knowledge Graph. In *Proceedings of the 2019 IEEE International Conference on Big Knowledge (ICBK)*, pp. 221–228, Beijing, China.
- Silva, T. H. P.; Laender, A. H. F.; Davis Jr., C. A.; da Silva, A. P. C. & Moro, M. M. (2016). The Impact of Academic Mobility on the Quality of Graduate Programs. *D-Lib Magazine*, 22(9/10).
- Simpson, K. (2015). *You Don't Know JS - Async & Performance*. O'Reilly Media, Inc., Sebastopol, Califórnia.
- Sonnenwald, D. H. (2007). Scientific collaboration. *Annual review of information science and technology*, 41(1):643–681.
- Strehl, L.; Calabró, L.; Souza, D. O. & Amaral, L. (2016). Brazilian Science between National and Foreign Journals: Methodology for Analyzing the Production and Impact in Emerging Scientific Communities. *PloS One*, 11(5):e0155148.
- Su, C.; Tong, J.; Zhu, Y.; Cui, P. & Wang, F. (2018). Network embedding in biomedical data science. *Briefings in Bioinformatics*, 21(1):182–197. ISSN 1477-4054.
- Sugimoto, C. R.; Ni, C.; Russell, T. G. & Bychowski, B. (2011). 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 and Technology*, 62(9):1808–1828.

- Tuesta, E. F.; Delgado, K. V.; Mugnaini, R.; Digiampietri, L. A.; Mena-Chalco, J. P. & Pérez-Alcázar, J. J. (2015). Analysis of an advisor-advisee relationship: An exploratory study of the area of exact and earth sciences in Brazil. *PloS one*, 10(5):e0129065.
- van der Aalst, W. (2016). *Data Science in Action*, pp. 3–23. Springer Berlin Heidelberg, Berlin, Heidelberg.

Apêndice A

Termo de Consentimento da Avaliação por Entrevistas

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Título: Entendendo os usuários e suas necessidades sobre genealogia acadêmica

Período: Fev/2020 **Instituição:** DCC / UFMG

Pesquisadores Responsáveis: João Marcos Martins da Costa Cota (joaocota@dcc.ufmg.br)
Alberto H.F. Laender (laender@dcc.ufmg.br)
Raquel O. Prates (rprates@dcc.ufmg.br)

Introdução: Este Termo de Consentimento contém informações sobre a pesquisa indicada acima. Para assegurar que você esteja informado sobre a sua participação nesta pesquisa, pedimos que leia este Termo de Consentimento Livre e Esclarecido. Caso tenha alguma dúvida, não hesite em perguntar ao pesquisador responsável. Você também deverá assinar o termo do qual receberá uma cópia.

Objetivo da avaliação: O objetivo desta entrevista é entender as necessidades dos participantes, pesquisadores ou que compõe a academia de modo geral, em relação a como uma ferramenta poderia melhorar as pesquisas realizadas para obter informações sobre outros pesquisadores..

Informação geral sobre a pesquisa: Você será solicitado a realizar algumas ações no sistema proposto, Árvore da Ciência, e após responderá algumas perguntas simples sobre a utilização do sistema e sobre as atividades propostas. A realização dessa entrevista pode ser gravada (áudio) para posterior análise pelos investigadores.

Utilização dos dados coletados: Os dados coletados durante a entrevista serão utilizados para entender as necessidades dos participantes com relação à sua interação com plataformas de visualização e pesquisa da genealogia acadêmica. Quaisquer dados utilizados para publicação serão apresentados de forma a garantir o anonimato dos participantes da avaliação.

Privacidade: Informações que possam identificar os participantes da pesquisa não serão divulgadas. O seu nome não aparecerá em nenhum relatório. Caso deseje, poderá solicitar uma cópia dos dados gerados por você.

Se você decidir não participar na pesquisa: Você é livre para decidir, a qualquer momento, se quer participar ou não nesta pesquisa. Sua decisão não afetará sua vida estudantil/profissional e nem qualquer relacionamento com os avaliadores ou a instituição por trás desta.

Compensação: A participação nesta pesquisa é voluntária, e não será oferecida nenhuma remuneração aos seus participantes.

Se tiver algum problema ou se tiver outras perguntas: Se você tiver algum problema que pensa que pode estar relacionado com sua participação nesta pesquisa, ou se tiver qualquer pergunta sobre a pesquisa, poderá entrar em contato com os pesquisadores a qualquer momento pelos e-mails indicados acima.

Novas condições: Caso deseje, você pode especificar novas condições que devem ser atendidas para que você participe desta avaliação.

Consentimento Livre e Esclarecido (Acordo Voluntário) O documento mencionado acima descrevendo os benefícios, riscos e procedimentos da pesquisa foi lido e explicado. Eu tive a oportunidade de fazer perguntas sobre a pesquisa, que foram respondidas satisfatoriamente. Eu estou de acordo em participar como voluntário.

Assinatura do participante: _____

Nome do participante: _____

Assinatura do pesquisador: _____

Nome do pesquisador: _____

Data _____