

Genealogia acadêmica: um novo olhar sobre impacto acadêmico de pesquisadores

Luciano Rossi¹, Rafael J. P. Damaceno² e Jesús P. Mena-Chalco³

Resumo

A genealogia acadêmica (GA) considera as relações de orientação, em diferentes níveis, para a avaliação e classificação de grupos e indivíduos atuantes na academia. Os resultados obtidos por meio da aplicação da GA podem ser considerados complemento às tradicionais análises sobre produção bibliográfica. A GA é relevante para a Ciência, Tecnologia e Inovação (CT&I) porque possibilita a descoberta de novas informações que ajudem as agências de fomento a documentar, identificar e analisar o impacto individual e coletivo de pesquisadores. O conhecimento obtido a partir da GA está relacionado à formação de recursos humanos, à estruturação de comunidades acadêmicas e a seus respectivos desdobramentos.

Abstract

The academic genealogy considers the academic mentoring relationships, at different levels, for the evaluation and classification of groups and individuals acting in the universities. The results obtained through the application of the academic genealogy can be considered in a way that complements the traditional analysis of scholarly output. The relevance of academic genealogy lies in the possibility of discovering new information that will assist Science, Technology and Innovation (ST&I) development agencies to document, identify, and analyze the academic impact of a researcher or a group of researchers. The knowledge obtained from academic genealogy is related with the propagation of human resources, with the structuring of academic communities and with their unfolding.

1 Mestre em Ciência da Computação pelo Centro de Matemática, Computação e Cognição (CMCC) da Universidade Federal do ABC (UFABC).

2 Mestre em Ciência da Computação pelo CMCC/UFABC.

3 Engenheiro de Sistemas pela Universidade Nacional de San Agustín; mestre, doutor e PhD em Ciência da Computação pelo Instituto de Matemática e Estatística da Universidade de São Paulo (USP). Professor do CMCC/UFABC.

Palavras-chave: Genealogia. Indicadores de Ciência e Tecnologia. Cientometria.

Keywords: *Genealogy. Science and Technology Indicators. Scientometrics.*

1. Introdução

A busca da humanidade por suas origens é tema recorrente e pode ser observada desde os registros bíblicos no Livro de Gênesis até as iniciativas contemporâneas de documentar a história de comunidades acadêmicas por meio de suas relações. A genealogia acadêmica (GA) busca, no registro de acadêmicos em diversas áreas e em seus relacionamentos de orientação, documentar, de forma perene, a história de comunidades acadêmico-científicas e, assim, oferecer um melhor entendimento sobre como o passado impactou na formação do presente e qual a tendência para o futuro desses grupos (MALMGREN *et al.*, 2010).

A genealogia é uma ciência auxiliar da história que estuda a origem, evolução e disseminação de grupos inter-relacionados por laços familiares ou outro tipo de ligação qualquer. Os objetos dessa área do saber são os ascendentes e os descendentes de um indivíduo. O processo de pesquisa abrange a identificação de parentesco formal ou informal, por meio de registros históricos como certidões de nascimento, casamento, óbito, registro de propriedades e outros documentos (CHAO *et al.*, 1992).

A finalidade é obter conhecimento que permita uma possível compreensão a respeito do futuro com base no passado, este, por vezes, desconhecido. Esse estudo possibilita traçar a memória de diferentes grupos familiares (GARGIULO *et al.*, 2016).

O compartilhamento amplo e assertivo de dados referentes aos indivíduos e seus relacionamentos é ferramenta fundamental de trabalho para pesquisadores da genealogia. Por essa razão, o desenvolvimento da internet, de novas tecnologias e de plataformas de informação mudou radicalmente a forma pela qual os estudiosos dessa área conduzem suas pesquisas, acessam e compartilham dados.

A atividade de orientação, supervisão ou mentoria acadêmica é um tipo de relacionamento que promove a evolução do orientado, da instituição, da ciência e da sociedade em termos gerais. Atualmente, diversas iniciativas de grupos específicos documentam, analisam e classificam estruturas que facilitam a GA. Analisar os relacionamentos de orientação, sob a forma de

uma estrutura genealógica (e.g., grafo ou árvore), permite um maior entendimento sobre a comunidade científica, a caracterização do acadêmico por meio de seus relacionamentos e a identificação do impacto gerado por esses atores na constituição de seus respectivos grupos.

Segundo Lane (2010), a Plataforma Lattes, projetada e desenvolvida pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), apresenta-se como o maior e mais relevante repositório curricular de acadêmicos relacionados com a ciência brasileira. A partir desse conjunto de dados, é possível explorar a produção científica nacional de todo indivíduo/grupo atuante e as relações formais de orientação nos níveis de mestrado e doutorado, dentre outros. No contexto de genealogia, a propagação ou projeção acadêmica pode ser mensurada objetivamente usando os currículos registrados na Plataforma Lattes, um valioso conjunto de dados pouco explorado sob essa perspectiva.

O objetivo deste trabalho é apresentar a definição do conceito de genealogia acadêmica, os bancos de dados existentes e os principais benefícios para a mensuração quantitativa e qualitativa da ciência por meio de aplicações. Adicionalmente, é apresentado um novo sistema de consulta sobre a genealogia acadêmica do Brasil, denominado Plataforma Acácia. Acredita-se que este novo olhar permitirá, de forma complementar a outras análises cientométricas, evidenciar importantes atores em Ciência, Tecnologia e Inovação (CT&I) no País.

2. Genealogia acadêmica

Segundo Cronin e Sugimoto (2014), a genealogia acadêmica (GA) é definida como o estudo quantitativo da herança intelectual perpetuada por relacionamentos de orientação entre professores (orientadores/supervisores) e seus alunos (orientados). Trata-se da análise da formação de recursos humanos, no âmbito científico, o qual considera (i) a identificação dos atores envolvidos nesse processo e (ii) dos relacionamentos que os conectam; (iii) a estruturação dessas informações e (iv) a extração do conhecimento implícito nessas estruturas.

Os relacionamentos de mentoria acadêmica promovem a propagação de conhecimento científico por meio da relação entre orientador, que possui diferentes desempenhos nessa atividade, e seus orientados, que, por sua vez, comumente são influenciados pelas características de seus orientadores. Neste contexto, a GA fornece meios para mensurar e analisar essas interações formais de transmissão de conhecimento (DAVID; HAYDEN, 2012).

O grafo de GA é o principal objeto de estudo acadêmico-genealógico e é definido como uma estrutura matemática utilizada para representar coleções de objetos que possuam algum tipo de conexão entre si. Em outras palavras, um grafo é um conjunto de vértices que são ligadas aos pares, por meio de arestas. Um grafo de GA representa os acadêmicos por meio de vértices e os relacionamentos de orientação observados entre eles são representados por arestas. A Figura 1 apresenta um exemplo de grafo sem contextualização semântica.

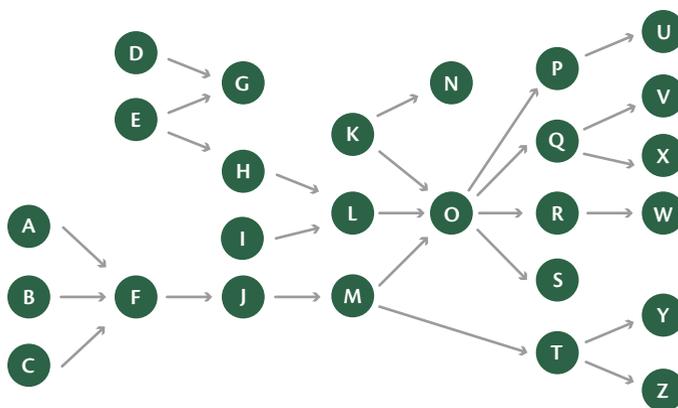


Figura 1. Exemplo de grafo. As circunferências e os segmentos de reta representam, respectivamente, os vértices e as arestas. As setas representam a orientação acadêmica de orientador para o orientado

Fonte: Elaborada pelos autores.

2.1. Conceitos preliminares

Um grafo de GA apresenta uma série de informações que podem ser utilizadas para a caracterização de sua topologia e, conseqüentemente, dos elementos que o compõem. Os conceitos apresentados a seguir consideram o grafo da Figura 1 como exemplo, de modo a auxiliar o entendimento sobre cada um deles:

- *Relação de orientação acadêmica:* comumente envolve um orientador e um orientado, representados, como no exemplo da Figura 1, pelos vértices A e F, respectivamente;
- *Ascendência direta:* um orientado (F) pode apresentar *mais de um orientador, ou coorientador* (A, B e C), denominados como sua *ascendência direta*;

- *Origem*: denominação de um acadêmico sem orientadores. No grafo exemplificado na Figura 1, os vértices que representam as origens são A, B, C, D, E, I e K;
- *Ascendência*: corresponde aos ancestrais acadêmicos de um indivíduo, isto é, todos aqueles que influenciaram sua formação, direta ou indiretamente. A *ascendência* do vértice L é formada pelos vértices E, H e I;
- *Descendência direta ou fecundidade*: formada pelos orientados (alunos) de um acadêmico. Os vértices P, Q, R, S e T formam a descendência direta do vértice O;
- *Descendência*: inclui, além da *descendência direta*, todos os influenciados, direta ou indiretamente, por um acadêmico. No caso do vértice O, a *descendência* inclui U, V, X, W, Y e Z;
- *Vértices não fecundos*: acadêmicos que não possuem descendência, como os representados por G, N, S, U, V, X, W, Y e Z;
- *Geração*: hierarquia entre os vértices, em função de seus relacionamentos. O grafo exemplo descreve sete *gerações* de acadêmicos: a primeira é formada pelos vértices A, B e C, a segunda, por D, E e F e, assim, sucessivamente;
- *Famílias*: relações entre acadêmicos nas quais todos os membros são conectados.

O conceito de *gerações de acadêmicos* é particularmente útil quando há famílias, visto que as gerações não são diretamente correlatas ao fator tempo e sim à distância observada entre os vértices iniciais (aqueles que não possuem orientadores) e seus descendentes.

2.2. Tipologia

A GA vem despertando um interesse crescente na comunidade científica, motivando o desenvolvimento de trabalhos com finalidades distintas. Os objetivos de sua aplicação podem ser parâmetros para uma classificação geral, porém não definitiva, em cinco tipos de genealogia, a saber (CRONIN; SUGIMOTO, 2014):

- *Honorífica*: estudo da descendência de um indivíduo ou grupo de interesse (ou grupo de indivíduos), com o objetivo de homenageá-lo em razão de sua linhagem, evidenciando

a relevância de sua contribuição na formação da comunidade em função de suas orientações acadêmicas;

- *Egocentrista*: uso da ascendência de um indivíduo de interesse a fim de identificar ancestrais ilustres e comprovar uma ligação, ainda que remota, entre ambos. Comumente, a motivação para esse tipo de estudo está relacionada com curiosidade ou autopromoção;
- *Histórica*: ferramenta complementar a estudos que visam à identificação de personagens memoráveis para determinadas áreas do conhecimento. Permite contextualizar as relações entre os indivíduos;
- *Paradigmática*: considerado um dos tipos de GA mais inseridos no contexto científico. Utiliza os relacionamentos orientador-orientado para estudar a extensão do conhecimento transmitido entre essas relações; e
- *Analítica*: ferramenta para a avaliação e, por vezes, predição de padrões entre os membros de comunidades acadêmicas que são objeto de estudo. Surgiu devido ao crescimento do número de bases de dados genealógicos e ao desenvolvimento de análises estatísticas. Comumente, os resultados desse tipo de estudo são possibilitados pela aplicação de métricas topológicas (ROSSI *et al.*, 2018).

2.3. Dados genealógicos

Atualmente, o interesse pela GA é crescente e se reflete nos esforços empreendidos, por diferentes comunidades, em desenvolver projetos que busquem a captação de dados de genealogia (HEINISCH; BUENSTORF, 2018). As iniciativas identificadas geralmente disponibilizam um conjunto de dados via *web* de forma semiestruturada.

Exemplos de grandes conjuntos de dados internacionais, associados a projetos no nível de doutorado, são: Mathematics Genealogy Project; The Neurotree Project; The Academic Family Tree; e RePEc Genealogy. Além desses, pode-se contar com os currículos acadêmicos disponíveis na Plataforma Lattes e com o catálogo de teses e dissertações da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes).

Para todos os acadêmicos registrados na Plataforma Lattes, são encontrados dados correspondentes à produção científica, técnica e artística. A informação sobre orientação/supervisão está disponível para aqueles que atuam na formação de alunos. Conforme observado

no processo de prospecção dos currículos, o registro desses elementos não mantém um padrão e, em muitos casos, é incompleto e sujeito a erros (e.g., registros com nomes abreviados, tanto do orientador quanto do orientado e sem associação do número identificador, isto é, o ID Lattes). Assim, obter uma estrutura hierárquica correta, que representa o parentesco acadêmico, torna-se um problema não trivial.

É importante destacar que a própria coleta de dados, sua estruturação e análise representam atuais oportunidades de pesquisa e aplicação que envolvem as áreas de Ciência da Computação, Ciência da Informação e Política Científica. O crescente interesse de comunidades acadêmicas em registrar, de forma perene, sua história, por meio dos relacionamentos de seus membros, além de possibilitar o avanço das pesquisas na área de GA, contribui para a manutenção da memória científica de diferentes áreas do conhecimento (PIERRO, 2017).

2.4. Aplicações

As pesquisas desenvolvidas no âmbito da GA podem ter objetivos orientados pela tipologia descrita anteriormente. Comumente, essas pesquisas baseiam-se em um grafo que se caracteriza como principal objeto de estudos acadêmicos-genealógicos. Um ponto em comum na maioria dos trabalhos é a possibilidade de registrar, de forma duradoura, o contexto histórico de indivíduos e sua importância na formação da comunidade acadêmica na qual estão inseridos (BENNETT *et al.*, 2005; DORES *et al.*, 2016).

Na Figura 2, é apresentado um exemplo de grafo de GA onde são representados matemáticos ilustres e seus relacionamentos de orientação acadêmica no nível de doutorado. Os vértices que ilustram os respectivos países de titulação estão agrupados em famílias acadêmicas. Esse tipo de estruturação é útil, por exemplo, para o estudo da formação dessas famílias, sob o ponto de vista geopolítico, por meio da observação da influência de estrangeiros nesse processo de estruturação em um determinado país.

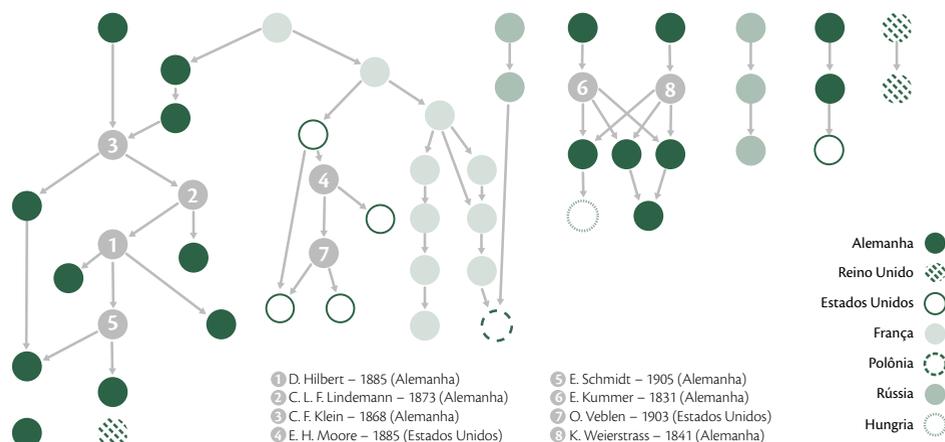


Figura 2. Exemplo de grafo de GA que representa os relacionamentos de orientação, em nível de doutorado, entre matemáticos ilustres. Cada vértice representa um pesquisador. Cada ligação representa a orientação formal no nível de doutorado. O conjunto de dados considerado corresponde ao Projeto de Genealogia dos Matemáticos. O eixo vertical representa a hierarquia observada de acordo com as relações de orientação

Fonte: Elaborado pelos autores.

Além do contexto histórico, é comum utilizar a GA para prestar homenagem a acadêmicos ilustres e evidenciar sua importância para o desenvolvimento da área na qual estão inseridos. Na Figura 3, é apresentada a descendência acadêmica do professor Cesare Mansueto Giulio Lattes, mais conhecido como César Lattes, renomado físico brasileiro que empresta seu nome ao maior repositório acadêmico-curricular do País. No grafo, os vértices rotulados em verde representam os orientados diretos do professor Lattes, ou seja, seus filhos acadêmicos. Os rotulados em verde claro completam a sua descendência que, no total, apresenta 425 acadêmicos (coleta de novembro de 2017).

Identificar a descendência de um acadêmico pode ser útil para outras descobertas. É possível observar, por exemplo, a divergência entre áreas de atuação do orientador e do orientado, um padrão especialmente interessante, pois contraria o senso comum da perpetuação dos acadêmicos de uma mesma linha na mesma área de atuação (evidência de multidisciplinaridade).

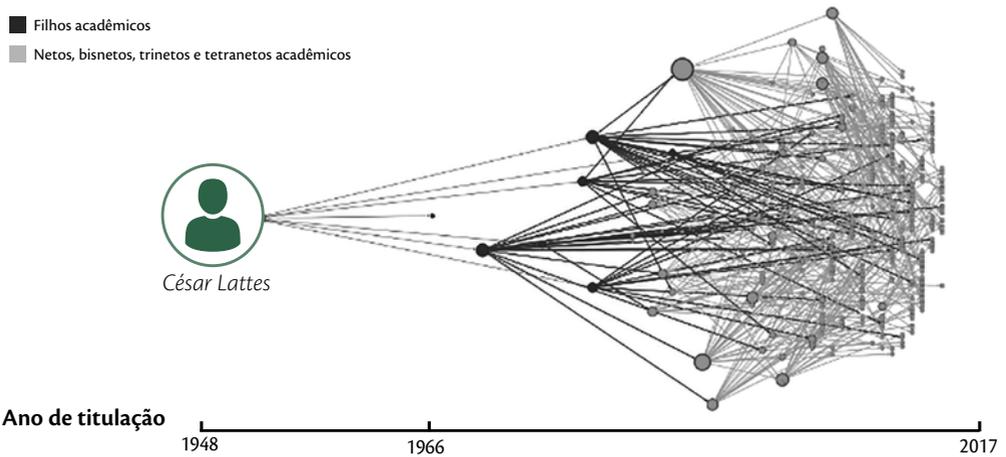


Figura 3. Descendência acadêmica do físico brasileiro César Lattes (em destaque). Os orientados diretos e indiretos são representados nos tons preto e cinza, respectivamente. O eixo horizontal representa o menor ano de titulação (Mestrado e/ou Doutorado) dos acadêmicos. O tamanho do vértice representa a quantidade de orientações realizadas, exceto aquele que representa o físico César Lattes

Fonte: Elaborado pelos autores.

As pesquisas alinhadas com a GA do tipo analítica buscam prover atributos quantitativos aos grafos de genealogia. Estes atributos podem descrever características sobre a configuração topológica do grafo, de modo que seja possível, por exemplo, classificar os acadêmicos com base em suas métricas topológicas (*i.e.*, medidas que apenas observem as relações de orientação). Nesse contexto, são descritas a seguir algumas possíveis métricas topológicas que podem ser consideradas para a caracterização de acadêmicos e tratadas como indicadores quantitativos (ROSSI *et al.*, 2018).

O desenvolvimento de medidas quantitativas tem por objetivo descrever o grafo de GA ou um vértice específico, de acordo com a sua estrutura topológica. Assim, é possível caracterizar estas estruturas e classificá-las de acordo com padrões observados (CHARIKER *et al.*, 2017; LIU *et al.*, 2018). Quatro métricas topológicas básicas, fundamentadas nas definições disponíveis na Seção 2.1, podem ser citadas como exemplo:

- Fecundidade: representa o número de orientados diretos de um acadêmico de interesse.
- Descendência: representa o número de orientados indiretos. Essas duas primeiras métricas são úteis para verificar a produtividade/projeção dos acadêmicos e evidenciar sua contribuição na formação de recursos humanos e da comunidade na qual estão inseridos.
- Geração: indica o maior número de gerações influenciadas por um determinado acadêmico.
- Folha⁴: relacionada com o número de vértices não fecundos (acadêmicos que ainda não orientaram alunos).

As métricas topológicas são importantes para evidenciar características específicas dos grafos de GA (ROSSI *et al.*, 2017). Estas características podem ser utilizadas para identificar padrões de desenvolvimento dessas estruturas, relacionando as influências recebidas (medidas ascendentes) e exercidas (medidas descendentes), respectivamente. Neste contexto, acredita-se ser possível prever (e comparar) o impacto que um determinado acadêmico exercerá em sua comunidade, com base na observação de suas influências prévias. Isso é importante, por exemplo, para a definição de políticas de fomento à atividade acadêmico-científica.

Os grafos de GA podem, ainda, ser utilizados para a representação estruturada de outros atributos associados aos acadêmicos que os compõem, como, por exemplo, sua área de atuação. A identificação das áreas de atuação (tópicos) dos acadêmicos e a quantificação dos relacionamentos observados entre as áreas resultam na estruturação de um meta grafo. As áreas e as influências existentes entre elas são representadas, respectivamente, por vértices e arestas ponderadas.

Este tipo de estrutura é útil, por exemplo, para o estudo do fluxo de conhecimento científico (DING *et al.*, 2017; AN *et al.*, 2017) e seus padrões evolutivos. Assim, é possível compreender o ciclo de vida de disciplinas específicas, bem como observar seus desdobramentos devidos à especialização do saber. A área da Ciência da Computação, por exemplo, é um desdobramento da Engenharia Elétrica e ambas apresentam influências mútuas entre si (DAMACENO *et al.*, 2018).

4 O nome Folha é uma analogia relacionada com as árvores e está ligado à ideia de ser o último elemento em um ramo, no caso dos grafos, é o último vértice em um caminho.

3. A Plataforma Acácia

A história da Ciência no Brasil é muito recente, se comparada a de outros países com tradições seculares na formação de novos acadêmicos que promovem a expansão das fronteiras do conhecimento pelo mundo. Por outro lado, a Plataforma Lattes é uma iniciativa ímpar no registro das atividades de milhões de acadêmicos, por disponibilizar, por meio de currículos, dados sobre formação, produção científica e participação em eventos, dentre outros registros relevantes.

Porém, os dados de GA são, em parte, incompletos, principalmente nos campos referentes ao registro de informações acadêmico-genealógicas, onde são descritos os ascendentes e descendentes. Isso impossibilita a identificação de vínculos que existem, mas não estão preenchidos na plataforma. A organização da estrutura de dados é voltada para a exibição de currículos e não para a construção de análises genealógicas.

Nesse contexto, a Plataforma Acácia⁵, iniciativa de um grupo de pesquisadores brasileiros, foi desenvolvida com o objetivo de fornecer informações acadêmico-genealógicas sobre a comunidade acadêmico-científica no País. Seus dados são resultado da mineração de currículos oriundos da Plataforma Lattes e submetidos a processos computacionais que visam a corrigir inconsistências e incompletudes, tornar as informações mais assertivas e privilegiar a hierarquia acadêmica como estrutura para a disponibilização das informações. O procedimento computacional nesta plataforma foi inicialmente considerado por DAMACENO *et al.* (2017).

Além da estrutura hierárquica, baseada nos relacionamentos de orientação acadêmica nos níveis de mestrado e doutorado e apresentada na forma de listas de orientadores e orientados, a Plataforma Acácia disponibiliza informações sobre: (i) Grande Área; (ii) Área e (iii) instituição de atuação do acadêmico; (iv) a maior titulação obtida pelo(a) pesquisador(a); e (v) o ano no qual a orientação ocorreu. Adicionalmente, a plataforma disponibiliza o resultado de três métricas topológicas sobre a GA: (i) Descendência, (ii) Fecundidade e (iii) Índice genealógico⁶.

Para exemplificar, na Figura 4a, é apresentado o número de mestres e doutores por geração dos acadêmicos identificados na Plataforma Acácia. Existem, ao todo, 13 gerações, isto é, 12 níveis separam o acadêmico mais antigo identificado na plataforma do mais recente. A maior parte dos

5 O termo Acácia é inspirado na árvore Acácia, uma espécie nativa do sudeste australiano. O formato da copa desta espécie assemelha-se aos grafos de genealogia acadêmica identificados no contexto brasileiro, ou seja, são compactos em termos de altura, indicando que, no Brasil, a ciência é jovem (possui poucas gerações de doutores e mestres), mas largos, em termos de comprimento.

6 O índice genealógico de um acadêmico é o maior número g de descendentes que possuem, no mínimo, g descendentes cada um.

acadêmicos está entre as gerações três e cinco, que concentram em torno de 650 mil mestres e doutores e representam mais de 58% do total. Esses dados mostram que a Ciência no Brasil ainda é jovem, dado que mais da metade dos cientistas está em gerações relativamente recentes.

Outra análise que pode ser feita diz respeito a grupos específicos de pesquisadores. Um exemplo são os bolsistas de produtividade em pesquisa do CNPq (com vigência em 2018) nos níveis 2, 1D, 1C, 1B, 1A e SR. A Figura 4b-d apresenta a distribuição desses grupos por descendência, fecundidade e índice genealógico. Os acadêmicos do grupo SR possuem maior descendência e índice genealógico. Já a métrica fecundidade, associada ao número de orientações realizadas, é maior para o grupo 1A.

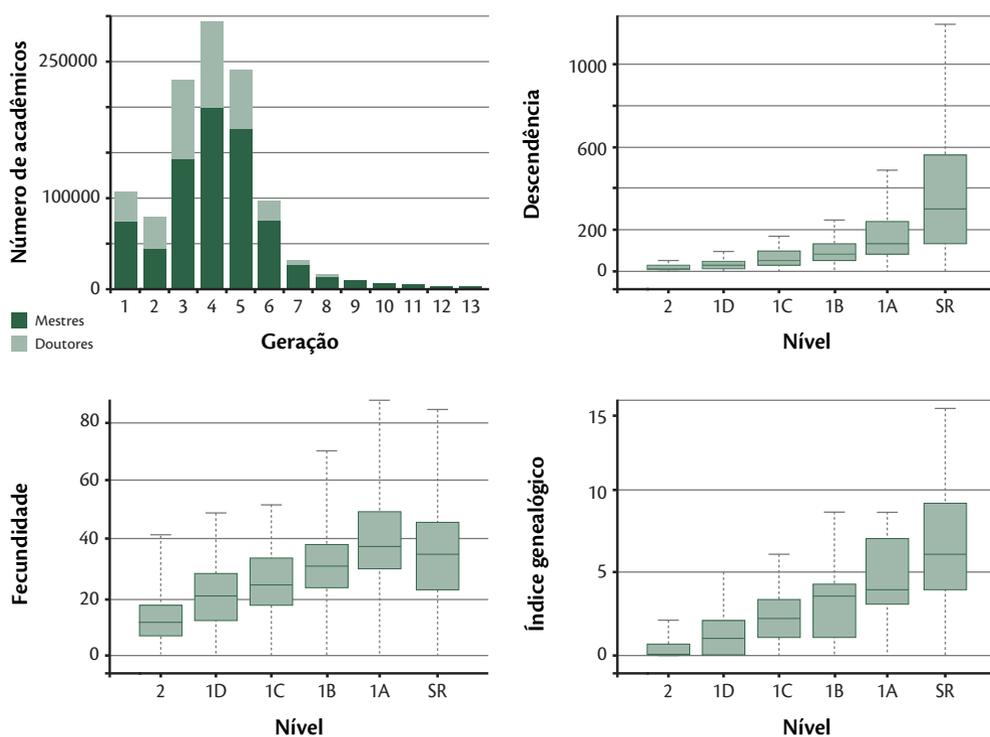


Figura 4. Número de acadêmicos (mestres e doutores) por (a) geração acadêmica. Distribuição dos 14.723 bolsistas de produtividade em pesquisa do CNPq/2018 (níveis 2, 1D, 1C, 1B, 1A e SR) conforme as métricas (b) descendência, (c) fecundidade, e (d) índice genealógico

Fonte: Elaborado pelos autores a partir dos dados da Plataforma Acácia.

A Plataforma Acácia apresenta-se como um projeto em desenvolvimento que, além de disponibilizar as informações genealógicas de forma mais completa e assertiva, possibilitará novos olhares sobre a comunidade acadêmica no Brasil, sob a ótica do processo de desenvolvimento de recursos humanos. As principais atualizações que serão implementadas na plataforma e seus respectivos objetivos são:

- Disponibilizar um conjunto mais amplo de métricas topológicas que auxilie na caracterização dos acadêmicos, com base em seu desempenho na formação de recursos humanos; e
- Fornecer ferramentas que possibilitem observar o grau de parentesco entre acadêmicos e disponibilizar relatórios personalizados que descrevam as informações acadêmico-genealógicas de interesse do usuário.

4. Considerações finais

A orientação acadêmica é parte importante das atividades que integram um ambiente de pesquisa, de modo a garantir a perpetuação da atuação acadêmica. A documentação histórica dessas atividades possibilita a criação de métodos para compreender as relações de interação na comunidade que constitui o ambiente: os professores orientadores, os alunos orientados, as instituições onde se formaram os alunos, as áreas mais proeminentes/carentes na formação, dentre outras características.

Analisar as relações de orientações, por meio de genealogia acadêmica, pode revelar informações relevantes a respeito da evolução da produção científica de um país. Conhecer o cenário atual e o caminho percorrido pela área de pesquisa acadêmico-científica poderá resultar em contribuições para a tomada de decisões mais assertivas de futuro para a comunidade.

O aprofundamento na pesquisa de modelos alternativos (*i.e.*, novos olhares à ciência) contribuirá para a criação de insumos importantes para responder a questões pouco exploradas atualmente e que permeiam a CT&I.

Referências

AN, Y.; HAN, M.; PARK, Y. Identifying dynamic knowledge flow patterns of business method patents with a hidden markov model. **Scientometrics**, v. 113, n. 2, p. 783-802, 2017.

BENNETT, A.F.; LOWE, C. The academic genealogy of George A. Bartholomew. **Integrative and Comparative Biology**, v. 45, n. 2, p. 231-233, 2005.

CHAO, G.T.; WALZ, P.; GARDNER, P.D. Formal and informal mentorships: A comparison on mentoring functions and contrast with nonmentored counterparts. **Personnel Psychology**, v. 45, n. 3, p. 619-636, 1992.

CHARIKER, J.H.; ZHANG, Y.; PANI, J.R.; ROUCHKA, E.C. Identification of successful mentoring communities using network-based analysis of mentor-mentee relationships across Nobel laureates. **Scientometrics**, v. 111, n. 3, p. 1733-1749, 2017.

CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO. **Plataforma Lattes**. Disponível em: <http://lattes.cnpq.br>. Acesso em: 1 jun. 2018.

COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR – CAPES. Catálogo de teses e dissertações. Disponível em: <http://catalogodeteses.capes.gov.br>. Acesso em: 1 jun. 2018.

CRONIN, B.; SUGIMOTO, C.R. **Beyond bibliometrics**: Harnessing multidimensional indicators of scholarly impact. MIT Press, 2014.

DAMACENO, R.J.P.; ROSSI, L.; MENA-CHALCO, J.P. Identificação do grafo de genealogia acadêmica de pesquisadores: Uma abordagem baseada na Plataforma Lattes. In: BRAZILIAN SYMPOSIUM ON DATABASES, 32. 2017. **Proceedings...** p. 76-87, 2017.

DAMACENO, R.J.P.; ROSSI, L.; MENA-CHALCO, J.P. Identification of the Brazilian academic roots through mining advisor-advisee relationships. In: SYMPOSIUM ON KNOWLEDGE DISCOVERY, MINING AND LEARNING (KDMiLe). **Proceedings...** p. 1-12, 2018.

DAVID, S.V.; HAYDEN, B.Y. Neurotree: A collaborative, graphical database of the academic genealogy of neuroscience. **PLoS One**, v. 7, n. 10, p. e46608, 2012.

DING, C.G.; HUNG, W.C.; LEE, M.C.; WANG, H.J. Exploring paper characteristics that facilitate the knowledge flow from science to technology. **Journal of Informetrics**, v. 11, n. 1, p. 244-256, 2017.

DORES, W.; BENEVENUTO, F.; LAENDER, A.H. Extracting academic genealogy trees from the networked digital library of theses and dissertations. In: ACM/IEEE-CS ON JOINT CONFERENCE ON DIGITAL LIBRARIES, 16., 2016. **Proceedings...** ACM: 2016. p. 163-166.

GARGIULO, F.; CAEN, A.; LAMBIOTTE, R.; CARLETTI, T. The classical origin of modern mathematics. **EPJ Data Science**, v. 5; n. 1, p. 26, 2016.

HEINISCH, D.P.; BUENSTORF, G. The next generation (plus one): an analysis of doctoral students' academic fecundity based on a novel approach to advisor identification. **Scientometrics**, p. 1-30, 2018.

LANE, J. Let's make science metrics more scientific. **Nature**, v. 464, n. 7288, p. 488-489, 2010.

LIU, J.; TANG, T.; KONG, X.; TOLBA, A.; AL-MAKHADMEH, Z.; XIA, F. UNDERSTANDING the advisor-advisee relationship via scholarly data analysis. **Scientometrics**, p. 1-20, 2018.

MALMGREN, R.D.; OTTINO, J.M.; AMARAL, L.A.N. The role of mentorship in protégé performance. **Nature**, v. 465, n. 7298, p. 622-626, 2010.

MATHEMATICS GENEALOGY PROJECT. Site. North Dakota State University. Disponível em: <https://www.genealogy.math.ndsu.nodak.edu/>

PIERRO, B. Galhos e raízes da árvore da ciência. **Revista Pesquisa FAPESP**, n. 249, nov. 2016.

PLATAFORMA ACÁCIA. Genealogia acadêmica do Brasil. Site. Universidade Federal do ABC - UFABC. Disponível em: <http://plataforma-acacia.org/quem>. Acesso em: 10 jun. 2018.

REPEC GENEALOGY. Site. Disponível em: <https://genealogy.repec.org/>

ROSSI, L.; DAMACENO, R.J.P.; FREIRE, I.L.; BECHARA, E.J.H.; MENA-CHALCO, J.P. Topological metrics in academic genealogy graphs. **Journal of Informetrics**, v. 12, n. 4, p. 1042-1058, 2018.

ROSSI, L.; FREIRE, I.L.; MENA-CHALCO, J.P. Genealogical index: A metric to analyze advisor-advisee relationships. **Journal of Informetrics**, v. 11, n. 2, p. 564-582, 2017.

THE ACADEMIC FAMILY TREE. *Building a single, interdisciplinary academic genealogy*. Site. Disponível em: <https://academictree.org/>

THE NEUROTREE PROJECT. Neuroscience Academic Family Tree. Disponível em: neurotree.org/neurotree