



Um método orientado ao propósito aplicado ao Currículo Lattes para fins de concessão de fomento a pesquisadores em grupos colaborativos

Resumo

Mensurar a qualidade de pesquisadores tem sido desafiante para gestores de pesquisa científica. A busca por soluções, criou métodos objetivos que legitimam a qualidade e prestígio de pesquisadores, a partir de produções científicas. Contudo, este desafio ainda requer métodos que considerem diferentes propósitos. Por exemplo, na concessão de fomento em diferentes modalidades: auxílio, formação, e colaboração. Este artigo apresenta um novo método para mensurar a qualidade de pesquisadores orientado ao propósito da avaliação. Nós ilustramos o método, simulando um processo de concessão de fomento a grupos de pesquisadores, usando uma amostra de 200 mil currículos da base *Lattes*. Nós apresentamos um instrumento que seleciona currículos que descrevem pesquisadores com o mesmo potencial de qualidade de pesquisadores do grupo sob consideração.

Palavras-chave: Colaboração científica. Currículo *Lattes*. Fomento. Mensuração da qualidade do pesquisador. Raciocínio baseado em casos.

Abstract

Assessing researcher quality has been challenging for managers of scientific research. The search for solutions, has provided objective methods, based on their scientific productions, which legitimize the quality and prestige of researchers. However, this challenge still requires methods that consider different purposes. For example, in funding decisions for: research project, scholarship, and collaboration. This article presents a new method to assess researcher quality oriented to the purpose of the evaluation. We illustrate the method by simulating a process of funding to researcher groups using a sample of 200,000 curriculums from the *Lattes* database. We present an instrument that selects *curriculum vitae* describing researchers with the same quality researcher of the group under consideration.

Keywords: Case-based reasoning. Curriculum *Lattes*. Research collaboration. Researcher quality assessment. Funding.



1. INTRODUÇÃO

O desafio de mensurar a qualidade de pesquisadores, tem sido um constante desafio, para gestores que conduzem políticas de pesquisa científica (GARFIELD; MALIN, 1968; LANE, 2015). O instrumento mais antigo para este propósito é a revisão por pares, o qual é baseado em critérios subjetivos (HAUSTEIN; LARIVIÈRE, 2015). No entanto, estudos em Bibliometria tem avançando significativamente, propondo métricas objetivas que mensuram a qualidade de pesquisadores, por meio de seus trabalhos, legitimando assim, sua autoridade e prestígio (GARFIELD; MALIN, 1968; VAN NOORDEN, 2010).

Apesar da evolução da Bibliometria, críticas sobre sua aplicação tem levado gestores de ciência e tecnologia a um debate sobre, o que precisa ser mensurado, porque precisa, e qual o propósito da mensuração (VAN NOORDEN, 2010). Um ponto crítico desta questão, diz respeito ao uso de indicadores sem levar em conta a necessidade dos diferentes contextos de avaliação (STI, 2016).

Em resposta a tal debate, um grupo de pesquisadores propôs um conjunto de princípios denominado *Leiden Manifesto* (HICKS et al., 2015). Tais princípios recomendam o uso de métricas objetivas para o apoio a julgamentos qualitativos, e sugerem que tais métricas sejam alinhadas à missão e objetivos das instituições. Eles enfatizam a necessidade de reconhecimento das pesquisas locais, declaram que as métricas devem ser transparentes, e ressaltam a importância da qualidade dos dados e de sua atualização. Além disso, apontam a necessidade de se rever o portfólio dos pesquisadores e alertam sobre os riscos de resultados injustos quando usadas erroneamente.

Considerando ainda, as tendências da pesquisa científica, é preciso observar a emergência de uma nova cultura de conscientização colaborativa entre pesquisadores, motivada por uma ciência aberta que inclui práticas colaborativas (OCDE, 2016). Porém, não são muitas as métricas voltadas a colaboração científica, sendo a coautoria de artigos científicos o método mais utilizado, o que leva alguns autores, a sugerir a busca por melhores formas de sua mensuração (BOZEMAN, FAY E SLADE, 2013).

Neste cenário, podemos citar as agências de fomento, nas quais o foco da concessão do fomento geralmente está no que é fomentado (a proposta de pesquisa) e não em quem recebe o fomento (pesquisadores), que são os responsáveis pela descoberta e



transmissão da descoberta do conhecimento (LANE, 2015). Além disso, o processo de avaliação é comumente realizado por meio de revisão por pares, seja o propósito, o auxílio à pesquisa, a formação do pesquisador, ou o incentivo à colaboração científica (SARTORI, 2011; MOLAS-GALLART, 2012).

Neste artigo, considerando o contexto descrito acima, e perguntamos a seguinte questão de pesquisa: "*Como mensurar a qualidade do pesquisador, especificamente em propósitos de concessão de fomento a grupos colaborativos?*"

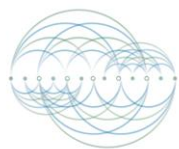
Nossa proposta ao responder tal questão parte do princípio que nós buscamos qualidade, e entendemos que qualidade é "*fitness-for-purpose*" (JURAN; GODFREY, 1999), um conceito que considera qualidade como dependente de perspectivas, necessidades e prioridades de usuários.

Desta forma, propomos um método para mensurar a qualidade de pesquisadores, orientado ao propósito do processo de avaliação, o qual segue os princípios do *Leiden Manifesto* (HICKS et al., 2015), é fundamentado em uma técnica de inteligência artificial chamada Raciocínio Baseado em Casos (RICHTER e WEBER, 2013), e usa dados de currículos de pesquisadores, nos quais estão registradas as produções científicas realizadas ao longo de suas trajetórias de carreiras.

Nós ilustramos a solução, aplicando o método proposto na base de dados do *Currículo Lattes* (lattes.cnpq.br) simulando a seleção de pesquisadores para trabalhar em colaboração com grupos de pesquisa do *Diretório dos Grupos de Pesquisa no Brasil* (lattes.cnpq.br/web/dgp), para fins de concessão de fomento a estes pesquisadores.

A solução provê aos gestores de ciência e tecnologia, em especial, os gestores de agências de fomento, um instrumento que seleciona apenas os currículos que descrevem pesquisadores candidatos, com o mesmo potencial de qualidade de pesquisadores do grupo colaborativo em questão, considerados bem-sucedidos pelo comitê de avaliação.

Na próxima seção nós fornecemos a fundamentação teórica para este trabalho. Em seguida, na Seção 3 apresentamos os procedimentos metodológicos. A Seção 4, introduz o método orientado a propósitos para mensurar a qualidade do pesquisador. O estudo de caso ilustrando a aplicação do método é apresentado na Seção 5. Concluindo o artigo, a Seção 6 apresenta as considerações finais e trabalhos futuro.



2. REFERENCIAL TEÓRICO

Nessa seção nós introduzimos alguns conceitos e técnicas usadas no método proposto. Nós iniciamos com um breve relato sobre o *curriculum vitae*, seu uso em estudos bibliométricos, e citamos como exemplo, o *Currículo Lattes* (lattes.cnpq.br). Passamos então à apresentação de algumas técnicas baseadas em Inteligência Artificial, que foram adotadas na implementação de nosso proposto método, tais como, Aprendizagem de Pesos, e Raciocínio Baseado em Casos (RBC).

2.1. CURRICULUM VITAE (CV)

O *curriculum vitae* do pesquisador é um documento histórico declarado pelo próprio pesquisador, o qual registra sua trajetória de vida acadêmica, isto é, sua formação educacional, atuação profissional, produções científicas, colaborações na pesquisa, e fomentos recebidos em projetos de pesquisa, dentre outros (DIETZ et al., 2000).

A riqueza de conhecimento contida no currículo, torna-o um atrativo e potencial fonte de conhecimento para estudos em Bibliometria, principalmente em pesquisas sobre mobilidade de pesquisadores e colaboração internacional (GAUGHAN, 2009; SANDSTRÖM, 2009). Além disso, os currículos têm contribuído como valiosas fontes de informação para indicadores de C&T em universidades, institutos de pesquisa, e agências de fomento (CGEE, 2016; FURTADO et al., 2015).

Uma das maiores referências repositórios de dados de currículos, é a base de dados do *Currículo Lattes* (lattes.cnpq.br), a qual foi criada em 1999, pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), e constitui o núcleo da Plataforma *Lattes* (PACHECO, 2003). A base de dados *Lattes* foi reconhecida internacionalmente por sua alta qualidade de dados (LANE, 2010), e atualmente conta com mais de cinco milhões¹ de currículos (CVs). A Plataforma contempla quatro unidades de informação (currículo, grupos de pesquisa, projetos de pesquisa, e instituições), sendo duas delas fontes de dados para este trabalho: o *Currículo Lattes* (lattes.cnpq.br), e o *Diretório de Grupos de Pesquisa do Brasil* (lattes.cnpq.br/web/dgp).

¹ Mais informações podem ser obtidas no sítio: < <http://estatico.cnpq.br/painelLattes/mapa>>



2.2. APRENDIZAGEM DE PESOS

A aprendizagem de pesos é uma técnica de classificação, baseada em inteligência artificial, que consiste em identificar características comuns em um conjunto de dados, atribuindo pesos aos atributos, de forma a refletir a relevância de cada atributo (HALL, 1999). *Feature weighting* é exemplo de método de aprendizagem de pesos, que enfatiza a importância dos atributos, produzindo uma lista ordenada de atributos, pelo peso atribuído (WETTSCHERECK; AHA; MOHRI, 1997). Diversas soluções de *feature weighting* tem sido propostas, dentre elas: *Correlation Based Feature Selection* (CFS) (HALL, 1999), e *ReliefF* (KNONENKO; ROBNIK-SIKONJA; POMPE, 1996).

2.3. RACIOCÍNIO BASEADO EM CASOS

Raciocínio baseado em casos (RBC) é uma técnica resultante de teorias da ciência cognitiva e ciência da computação, para um modelo computacional de inteligência artificial, semelhante ao raciocínio humano (RICHTER; WEBER, 2013). RBC tem como princípio fundamental, aprender com problemas passados que podem ser adaptados para ajudar a resolver problemas futuros (MÁNTARAS et al., 2005).

Neste processo de aprender por analogia, o conhecimento sobre um fato é recuperado da memória, depois apropriadamente transformado, aplicado à nova situação e, então, armazenado para uso futuro (CARBONELL; MICHALSKI; MITCHELL, 1983). Um ciclo de RBC foi proposto por Aamodt e Plaza (1994), o qual consiste em quatro etapas básicas: *Retrieve*, *Reuse*, *Revise* e *Retain*. Este ciclo representa o núcleo dos métodos que adotam RBC como uma metodologia capaz de combinar, em cada etapa de seu ciclo, diferentes técnicas e abordagens de aprendizado de máquina (AHMED; BEGUM; FUNK, 2012; WATSON, 1999).

O Ciclo do RBC (AAMODT; PLAZA, 1994) é estabelecido da seguinte forma: Dado um problema (novo caso), o primeiro passo, *Retrieve*, tem a missão de recuperar um caso candidato de uma base de casos, de forma que seja tão similar ao novo caso, que suas soluções possam ser trocadas. Neste passo, o grau de similaridade, entre o novo caso e os casos armazenados, é calculado por meio de métricas de similaridade. No segundo passo, *Reuse*, a solução recuperada é reusada em um novo problema (novo



caso). No passo seguinte, *Revise*, fontes externas são usadas para verificar se a solução proposta é válida, e assim fazer os ajustes necessários. No passo final, *Retain*, a solução é confirmada e retida na base de conhecimento para futuro reúso.

Várias aplicações baseadas em RBC tem sido desenvolvidas para diferentes áreas de conhecimento, em tarefas tanto analíticas quanto sintéticas, como por exemplo, no planejamento de processos de negócios (AGORGIANITIS et al., 2016), e em sistemas de recomendação (GUNAWARDENA, 2013).

3. METODOLOGIA

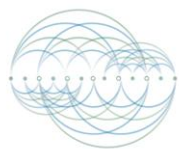
Nesta seção, primeiro descrevemos a fonte de dados, atributos, e os conjuntos de dados usados no cenário do estudo de caso. Em seguida, apresentamos o método orientado a propósitos, o qual leva em consideração o propósito da avaliação, e experiências, ou produções científicas dos pesquisadores, as quais estão mapeadas em seus currículos.

3.1. DADOS

Neste estudo utilizamos a base de dados *Lattes* (lattes.cnpq.br) como fonte de informações de currículos. Para o estudo de caso, foi extraída uma amostra de currículos constituída por 208.150 CVs de pesquisadores com doutorado concluído até o ano de 2015, e por produções científicas no período de 1950 a 2015, em Julho de 2016.

Para cada tipo de produção científica relacionada no Quadro 1, a produtividade (P) do pesquisador foi calculada como o total de produções alcançadas ao longo de sua trajetória de carreira.

A fim de enfatizar o contexto colaborativo na mensuração da qualidade do pesquisador, os 10 primeiros atributos relacionados no Quadro 1 foram analisados com base em literatura, considerando a coautoria, o número de coautores, e a respectiva posição na coautoria. Após isto, cada atributo foi subdividido em outros seis, por exemplo: artigos solo-autorado, artigos com um coautor, artigos com dois ou mais coautores, artigos como primeiro autor, artigos como segundo autor, e artigos como terceiro autor ou outro. Para as produções que não configuram coautorias diretamente, o aspecto colaborativo foi também justificado em literatura, conforme (PRICE e BEAVER, 1966;



NETO e DA COSTA, 2016), porém não foram subdivididos em novos atributos. Ao final, uma lista com 70 tipos de atributos de produções científicas foram identificados e categorizados em atributos colaborativos ou atributos não colaborativos.

Quadro 1. Exemplos de tipos de produções científicas identificadas na base *Lattes*.

#	Produções Científicas
1	Artigo complete publicado em periódico
2	Trabalhos publicados em Anais de eventos
3	Trabalhos publicados em Anais de eventos
4	Livro publicado e organização de obra publicada
5	Capítulo de livro
6	Patentes
7	Projetos de pesquisa – Auxílio financeiro
8	Projetos de pesquisa – Bolsa
9	Projetos de pesquisa – Cooperação
10	Projetos de pesquisa – Outros
11	Orientação e supervisão concluída – Dissertação de mestrado (Orientador principal)
12	Orientação e supervisão concluída – Dissertação de mestrado (Coorientador)
13	Orientação e supervisão concluída – Tese de doutorado (Orientador principal)
14	Orientação e supervisão concluída – Tese de doutorado (Coorientador)
15	Orientação e supervisão em andamento – Dissertação de mestrado (Orientador principal)
16	Orientação e supervisão em andamento – Dissertação de mestrado (Coorientador)
17	Orientação e supervisão em andamento – Tese de doutorado (Orientador principal)
18	Orientação e supervisão em andamento – Tese de doutorado (Coorientador)
19	Participação em bancas de trabalhos de conclusão – Mestrado
20	Participação em bancas de trabalhos de conclusão – Doutorado

Fonte: Elaborada pelos autores

Um cenário de aplicação foi criado a partir da amostra de 208.150 CVs, com dois conjuntos de dados. O primeiro, chamado de *CVs exemplos*, inclui os CVs do grupo de pesquisa registrado no *Diretório dos Grupos de Pesquisa no Brasil* (lattes.cnpq.br/web/dgp). O segundo conjunto de dados inclui os CVs de pesquisadores de todas as cinco regiões do Brasil, das mesmas áreas de conhecimento *CVs exemplos*. Este segundo conjunto de dados é chamado *CVs candidatos*.



3.2. O MÉTODO ORIENTADO AO PROPÓSITO PARA MENSURAR A QUALIDADE DO PESQUISADOR

O método proposto é uma solução implementada por meio de técnicas computacionais baseadas em aprendizagem de máquina, apoiadas em RBC. Especificamente, ele é uma implementação dos estágios *Retrieve* e *Reuse*, do Ciclo de RBC. A essência da solução é um classificador, o qual é uma estrutura de dados que pode classificar, por exemplo, um currículo de pesquisador como *fit* (apto) ou *unfit* (inapto) a um específico propósito de avaliação. Este classificador, que chamamos de *MOP-CV* é resultado de trabalhos anteriores destes autores, é composto de quatro estágios, conforme Figura 1.

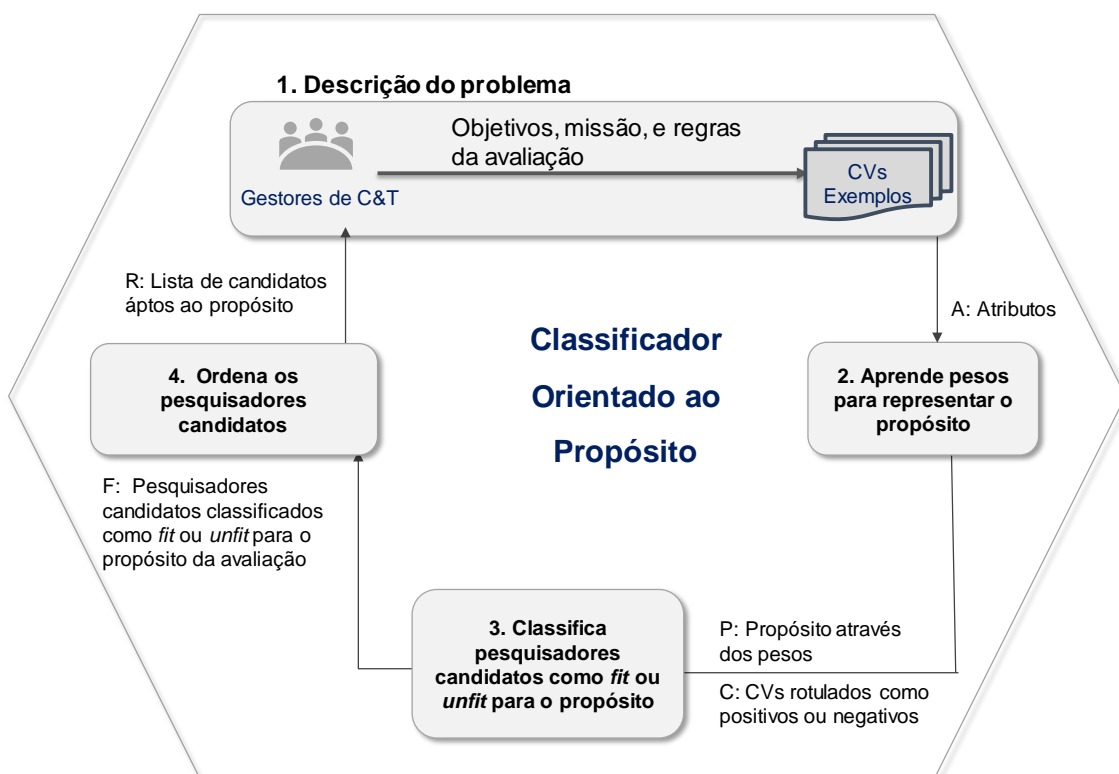


Figura 1. Classificador orientado a propósitos para mensurar a qualidade de pesquisadores baseado em seus CVs

Fonte: Elaborada pelos autores

No *Estágio 1* os objetivos do processo de avaliação são estabelecidos pelos gestores de C&T através de CVs exemplos. Por exemplo, supomos um processo de concessão de fomento à colaboração internacional de pesquisadores, para trabalharem com grupos de pesquisa em instituições de um determinado país. A solução adotada pressupõe que se



as produções científicas dos pesquisadores destes grupos estiverem registradas em CVs, métodos de aprendizagem de máquina poderão ser aplicados para caracterizar automaticamente o contexto de tais grupos, e conseqüentemente, identificar os requisitos de qualidade de tais grupos, traduzindo o propósito da seleção.

No *Estágio 2*, o propósito da avaliação é representado por meio de pesos relacionados aos atributos das produções científicas dos pesquisadores. Este estágio é uma etapa de pré-processamento que recebe como parâmetro os atributos do conjunto de CVs exemplos, e avalia a relevância relativa de cada atributo por meio de soluções de *Feature weighting* (WETTSCHERECK; AHA; MOHRI, 1997). Como resultado, o propósito dos CVs exemplos será refletida por meio da relevância de tais atributos.

No *Estágio 3*, os pesquisadores candidatos são classificados como *fit* (aptos) ou *unfit* (inaptos) ao propósito da avaliação, calculando-se o escore de similaridade (medida de *fitness*) entre um novo caso (um candidato a pesquisador) e cada caso (exemplar CV). Ao final deste processo, um conjunto de casos suficientemente semelhantes ao novo caso são retornados.

No *Estágio 4*, currículos de pesquisadores candidatos mais aptos ao propósito da avaliação são finalmente ordenados em uma lista classificatória. Esta lista resultante é então entregue aos gestores de C&T, e pode ser reutilizada e adaptada em um novo ciclo da implementação do classificador *MOP-CV*.

4. RESULTADOS

Nesta seção descrevemos os resultados da aplicação do classificador *MOP-CV* em um estudo de caso que simula um processo de seleção de fomento a pesquisadores candidatos a trabalhar em colaboração com membros de grupos Brasileiros de reconhecida excelência. Assim, nosso primeiro desafio foi encontrar um grupo de pesquisa ideal para este estudo de caso. O Diretório dos Grupos de Pesquisa do Brasil (lattes.cnpq.br/web/dgp), foi então investigado, e encontramos no *Grupo de Pesquisa da Epidemia de Microcefalia (MERG)*² uma adequada referência, indiscutivelmente de alta qualidade, a qual adotamos como exemplo de absoluto sucesso.

² Mais informações podem ser obtidas no sítio: <<http://www.cpqam.fiocruz.br/merg/>>



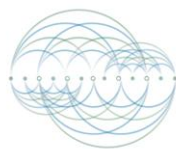
O próximo desafio foi representar o propósito da avaliação, caracterizando o contexto do grupo de pesquisa. Este processo foi realizado aplicando o algoritmo CFS (HALL, 1999), uma solução de *Feature weighting* (WETTSCHERECK; AHA; MOHRI, 1997). Ao final, a relativa relevância dos atributos dos CVs dos membros do grupo foi obtida em uma lista ordenada, do atributo de maior peso para o menor. A Figura 2 apresenta o resultado desta fase, mostrando os 20 atributos mais relevantes, e que representam o contexto do grupo de pesquisa estudado.



Figura 2. Contexto do grupo MERG, representado por seus 20 atributos mais relevantes

Fonte: Elaborada pelos autores

Ao observar os atributos listados na Figura 2, é possível destacar algumas tendências do contexto do grupo de pesquisa analisado. Por exemplo, o desenvolvimento de atividades voltadas à formação de pesquisadores, tais como, orientações a estudantes de mestrado e doutorado. Outra tendência está relacionada à produção de atividades científicas, tais como a publicação de artigos e capítulos de livros. Conseqüentemente, o propósito da avaliação é selecionar pesquisadores dentro deste mesmo contexto.



O passo seguinte, foi classificar os 15266 CVs de pesquisadores candidatos como *fit* (apto) ou *unfit* (inapto) ao contexto identificado no passo anterior. Assim uma medida de *fitness*, ou *score* de similaridade, foi calculada pela comparação de cada atributo do currículo de um pesquisador candidato, com os atributos dos currículos dos membros do grupo de pesquisa. O processo de classificação resultou em 3465 pesquisadores candidatos classificados como *fit* ao propósito do grupo de pesquisa, e 11335 pesquisadores candidatos classificados como *unfit* a tal propósito. A Figura 3 (a) apresenta o resultado do processo de seleção, mostrando os 3465 pesquisadores *fit*, distribuídos por áreas de conhecimento. A Figura 3 (b) apresenta os 3465 pesquisadores *fit*, distribuídos geograficamente, pelas cinco regiões do Brasil.

Neste resultado, destacamos a área de Ciências da Saúde, e as regiões sudeste e centro-oeste, com maior número de candidatos *fit*, a trabalharem em colaboração com os membros do grupo de pesquisa, e conseqüentemente receberem fomento para este propósito. Desta forma, o gestor do processo de avaliação do fomento, poderia querer selecionar pesquisadores, atendendo especificidades de determinadas áreas de conhecimento, ou mesmo, características regionais no Brasil.

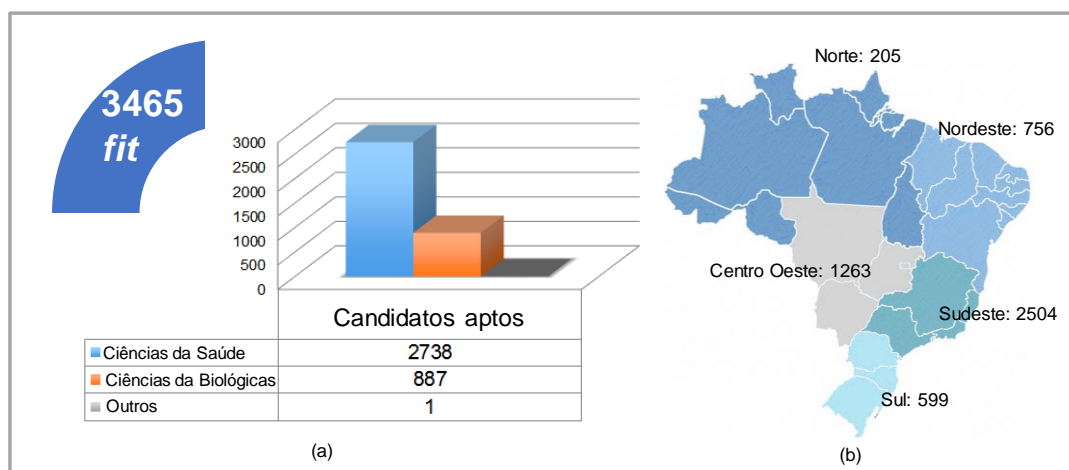


Figura 3. Candidatos *fit*, aptos a trabalharem em colaboração com o grupo de pesquisa MERG, distribuídos por área de conhecimento e regiões do Brasil.

Fonte: Elaborada pelos autores

Na última etapa da aplicação do classificador, a similaridade entre os membros do grupo de pesquisa em estudo, também foi investigada. Por exemplo, quão similar seria um colaborador a outro, dentro do grupo de pesquisa. Como consequência, três distintos



subgrupos emergiram. Por exemplo, o primeiro subgrupo mostrou que os colaboradores nr.3 e nr.16 são similares entre si. O segundo subgrupo, com 5 colaboradores, tem os pesquisadores nr.2 e nr.11 com forte similaridade, formando o núcleo principal. O terceiro subgrupo, e também o maior com 8 colaboradores, apresenta os pesquisadores nr.8 e nr.12 formando o núcleo principal. Concluindo o processo de simulação, estes três subgrupos foram utilizados para ordenar a lista final de candidatos aptos ao propósito de trabalhar em colaboração com o grupo de pesquisa em estudo.

5. CONCLUSÕES

Este artigo apresentou um novo método capaz de mensurar a qualidade de pesquisadores. O método que é orientado ao propósito da avaliação, e foi ilustrado por um estudo de caso que usou dados do Currículo *Lattes* (lattes.cnpq.br), para simular um processo de concessão de fomento a pesquisadores candidatos a colaborar em um grupo de pesquisa do *Diretório dos Grupos de Pesquisa no Brasil*. Os resultados do presente estudo levam a pelo menos três significantes contribuições do método proposto.

A primeira contribuição vai além de um método, entregando aos gestores de CT&I, uma metodologia que descreve em passos, como avaliar pesquisadores científicos, por meio de seus currículos, no caso, utilizando o *currículo Lattes* (lattes.cnpq.br).

A segunda contribuição, diz respeito à capacidade da solução de representar os critérios da avaliação, por meio da relativa relevância dos atributos dos CVs dos colaboradores do grupo de pesquisa, permitindo compreender o contexto deste grupo.

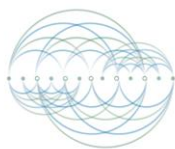
A terceira, e uma das maiores contribuições, é que o método funciona como um instrumento que, automaticamente seleciona apenas os currículos que descrevem pesquisadores com o mesmo potencial de qualidade de pesquisadores considerados bem-sucedidos pelo comitê de avaliação.

Quanto às limitações, este artigo não detalha como o contexto do grupo de pesquisa é caracterizado, e nem como as métricas de similaridade são aplicadas para mensurar o score de similaridade, que chamamos de medida de *fitness* do pesquisador. Além disso, não discutimos como produzir a lista final ordenada de pesquisadores aptos ao propósito do grupo. Estes são tópicos que serão descritos em trabalhos futuros.



REFERÊNCIAS

- AAMODT, A.; PLAZA, E. Case-Based reasoning: Foundational issues, methodological variations, and system approaches. **AI Communications**, v. 7, n. 1, p. 39–59, 1994.
- ADAMS J.; ARVANITIS, R.; MAHROUM, S.; SALAZAR, M. Global networks, internationalization and local research agendas: indicators for benchmarking or context specific? In: 21st International Conference on Science and Technology Indicators (STI2016): Peripheries, frontiers and beyond. **Anais ...**, 2016, p.44-45.
- AGORGIANITIS, I.; PETRIDIS, M., KAPETANAKIS, S.; FISH, A. Evaluating distributed methods for CBR systems for monitoring business process workflows. In: **ICCBR Workshops**, p. 122–131, 2016.
- AHMED, M. U.; BEGUM, S.; FUNK, P. Case Studies on the Clinical Applications using Case-Based Reasoning. In: **Computer Science and Information Systems (FedCSIS)**, 2012 Federated Conference on. IEEE, p.3–10, 2012.
- BEAVER, D. DEB.; ROSEN, R. Studies in scientific collaboration: Part I. The Professional Origins of Scientific Co-Authorship. **Scientometrics**, v. 1, n. 1, p. 65–84, 1978.
- BOZEMAN, B.; FAY, D.; SLADE, C. P. Research collaboration in universities and academic entrepreneurship: The-state-of-the-art. **Journal of Technology Transfer**, v. 38, n. 1, p. 1–67, 2013.
- CARBONELL, J. G.; MICHALSKI, R. S.; MITCHELL, T. M. Machine Learning: A Historical and Methodological Analysis. **AI Magazine**, v. 4, n. 3, p. 69, 1983.
- DIETZ, J. S. et al. Using the curriculum vita to study the career paths of scientists and engineers: An exploratory assessment. **Scientometrics**, v. 49, n. 3, p. 419–442, 2000.
- FURTADO, C. A. et al. A spatiotemporal analysis of Brazilian science from the perspective of researchers' career trajectories. **PLoS ONE**, v. 10, n. 10, p. 1–28, 2015.
- GAUGHAN, M. Using the curriculum vitae for policy research: an evaluation of National Institutes of Health center and training support on career trajectories. **Research Evaluation**, v. 18, n. June, p. 117–124, 2009.
- GARFIELD, E. “Science Citation Index” - A New Dimension in Indexing. **Science**, v. 144, n. 3619, p. 649–654, 1964.
- Garfield, E., & Malin, M. (1968). Can Nobel Prize winners be predicted. In: **135th Annual Meeting, American Association for the Advancement of Science** (pp. 1–8). Dallas, Texas.
- GUNAWARDENA, S. **Recommending Research Profiles for Multidisciplinary Academic Collaboration**. PhD Thesis. Drexel University, 2013.
- HALL, M. A. **Correlation-based Feature Selection for Machine Learning**. PhD thesis. University of Waikato, 1999.
- HAUSTEIN, S.; LARIVIÈRE, V. The use of bibliometrics for assessing research: Possibilities, limitations and adverse effects. Incentives and Performance: **Governance of Research Organizations**, p. 1–14, 2015.
- HICKS, D. et al. The Leiden Manifesto for research metrics. **Nature**, v. 520, n. 7548, p. 9–11, 2015.
- JURAN, J. M.; GODFREY, A. B. Juran's Quality Handbook. 5th edition, **McGraw-Hill**, 1999.



- KONONENKO, I.; ROBNIK-SIKONJA, M.; POMPE, U. ReliefF for estimation and discretization of attributes in classification, regression, and ILP problems. **Artificial intelligence: methodology, systems, applications**, p. 31-40, 1996.
- LANE, J. Let's make science metrics more scientific. **Nature**, v. 464, n. 7288, p. 488–9, 25 mar. 2010.
- LANE, J. I.; OWEN-SMITH, J.; ROSEN, R. F.; WEINBERG, B. A. New linked data on research investments: Scientific workforce, productivity, and public value. **Research Policy**, v. 44, n. 9, p. 1659–1671, 2015.
- NETO, João Estevão Barbosa; DA CUNHA, Jacqueline Veneroso Alves. Masters Committees and Academic Cooperation in Graduate Studies in Accounting. **Contabilidade, Gestão e Governança**, v. 19, n. 1, p. 126-145, 2016.
- MÁNTARAS, L. M. et al. Retrieval, reuse, revision, and retention in case-based reasoning. In: The Knowledge Engineering Review. **Cambridge University Press**, 2005. v. 0, p. 1–24.
- MOLAS-GALLART, J. Research Governance and the Role of Evaluation: A Comparative Study. **American Journal of Evaluation**, p. 1–16, 2012.
- ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT – OCDE. Science, Technology and Innovation Outlook 2016, Paris: OECD Publishing, 2016.
- PACHECO, R. C. S. Uma metodologia de desenvolvimento de plataformas de governo para geração e divulgação de informações e de conhecimento. **Grupo Stela**. 2003.
- PERLIN, M. S. et al. The Brazilian scientific output published in journals: A study based on a large CV database. **Journal of Informetrics**, v. 11, n. 1, p. 18–31, 2017.
- PRICE, D. J.; BEAVER, D. D. Collaboration in an invisible college. **The American psychologist**, v. 21, n. 11, p. 1011–1018, 1966.
- PRICE, D. J. DE S. Little science, big science. New York: **Columbia University Press**, 1963.
- RICHTER, M. M.; WEBER, R. O. Case-Based Reasoning: A Textbook. Berlin: **Springer-Verlag**, 2013.
- SANDSTRÖM, U. Combining curriculum vitae and bibliometric analysis: mobility, gender and research performance. **Research Evaluation**, v. 18, n. 2, p. 135–142, 2009.
- SARTORI, Rejane. **Governança em agentes de fomento dos sistemas regionais de CT&I**. 2011. 238f. Tese (Doutorado) – Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, Florianópolis, 2011.
- VAN NOORDEN, R. A profusion of measures. **Nature**, v. 465, n. June, p. 864–866, 2010.
- WETTSCHERECK, D.; AHA, D. W.; MOHRI, T. A Review and Empirical Evaluation of Feature Weighting Methods for a Class of Lazy Learning Algorithms. **Artificial Intelligence Review**, v. 11, p. 273–314, 1997.
- WATSON, I. Case-based reasoning is a methodology not a technology. **Knowledge-Based Systems**, v. 12, n. 5–6, p. 303–308, 1999.