



IDENTIFICAÇÃO DO PERFIL DOS PESQUISADORES DE PRODUTIVIDADE EM PESQUISA DO CNPq UTILIZANDO DADOS DA PLATAFORMA LATTES

Leandro Roberto Ferreira dos Santos

Thiago Magela Rodrigues Dias

Tales Moreira

Gray Farias Moita

leandroroberto@comp.ufla.br

thiogomagela@gmail.com

tales.info@gmail.com

gray@dppg.cefetmg.br

CEFET-MG

Rua Amazonas, 7675, Nova Gameleira, 30510-000, Minas Gerais, Brasil

Abstract. *The research productivity scholarship from CNPq is aimed at researchers with at least three years of PhD (or equivalent scientific profile) that stand out among their peers in conducting research in scientific and technological areas. Thus, CNPq provides a constant stimulus to the segments of excellence research in the country, focusing on the quality, the depth and the possible application of the studies. Considering the growth of research in all areas and, consequently, the number of researchers, the proportion of research productivity scholarship from CNPq tends to represent an increasingly smaller portion of researchers set. Thus the pressure for these research grant tends to increase considerably, leading to the adoption of increasingly restrictive criteria of selection, for example, increasing the completion time for the researcher to become eligible to compete for a research scholarship. Therefore, this paper will present the preliminary results of profile identification of the researchers with research productivity scholarship from CNPq, examining their scientific production throughout his academic life.*

Keywords: *Plataforma Lattes, Bolsa de Produtividade em Pesquisa, Análise de Dados Científicos, Extração e Recuperação de Informações*

1 INTRODUÇÃO

Nos últimos anos, tem se pesquisado bastante nos tópicos relacionados a redes acadêmicas. A extração de dados de produção científica e modelagem e visualização efetiva de redes de coautoria são assuntos de pesquisa atuais. Esse aumento da produção científica pode ser explicado pela maior disseminação de materiais de cunho científico na *internet*. Serviços como bibliotecas digitais e repositórios bibliográficos tem sido cada vez mais utilizados por pesquisadores para armazenar sua trajetória acadêmica e científica (DIAS; MOITA, 2015). Nesse âmbito, tem sido dado especial interesse a prospecção de dados em repositórios de produção científica (banco de dados de produções bibliográficas, de orientação acadêmica, de projetos de pesquisa e de diretórios de grupos de pesquisa) (MENA-CHALCO; JUNIOR, 2011).

Para este projeto utilizou-se a base de currículos acadêmicos do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) denominada Plataforma Lattes. Os currículos presentes nessa plataforma são denominados Currículos Lattes e são considerados um padrão nacional para análise do mérito e competência dos pesquisadores cadastrados, pois reúne toda a trajetória acadêmica do mesmo (AMORIN, 2003).

A Plataforma Lattes é a principal fonte de informações sobre os pesquisadores brasileiros, possuindo um banco de dados completo, altamente qualificado e com um elevado potencial para a extração de informação (LANE, 2010). Além disso, os currículos são monitorados pela própria comunidade científica, o que fortalece a qualidade e a fidelidade das informações contidas em cada currículo.

Pode-se extrair diversos tipos de informações utilizando a Plataforma Lattes, como dados bibliométricos e cientométricos, bem como analisar grupos específicos de pesquisadores, como os bolsistas de produtividade em pesquisa do CNPq. As bolsas de produtividade em pesquisa do CNPq são concedidas aos pesquisadores que se destacam entre seus pares. A extração desses dados é trabalhosa, por esse motivo é necessário utilizar mecanismos automáticos que realizem essa tarefa.

Baseado na Plataforma Lattes e especificamente nos Currículos Lattes, este artigo visa extrair os currículos de maneira automática e traçar o perfil dos pesquisadores que possuem a bolsa de produtividade em pesquisa do CNPq. Para isso é preciso obter atributos como modalidade da bolsa, instituição em que o pesquisador atua, tempo de doutorado, produção científica e orientações concluídas de mestrado e doutorado.

2 TRABALHOS RELACIONADOS

A plataforma lattes tem sido foco de muitos estudos nos últimos anos, nas mais diversas áreas do conhecimento. Diversos trabalhos fazem uma análise do perfil dos pesquisadores contemplados com a Bolsa de Produtividade em Pesquisa do CNPq em áreas distintas como Saúde, Ciência da Computação, Química e Engenharia de Produção, visando caracterizar o perfil dos bolsistas, já que representam os pesquisadores de excelência em cada uma destas áreas.

Martelli-Junior et al. (2010) analisaram os pesquisadores com bolsa de produtividade em pesquisa do CNPq na área de Medicina no triênio 2006-2008. As variáveis estudadas foram: sexo, área de atuação, artigos publicados, e orientação de alunos de graduação, mestres e doutores. Houve prevalência do sexo masculino (68%) e de bolsistas na categoria 2 (55,7%).

Foi constatado que 90% dos pesquisadores se concentram em quatro estados (SP,RJ,RS e MG). Do total, 80% dos pesquisadores estão vinculados a apenas oito instituições, com destaque para a USP (30,7%) e UNIFESP (17%).

Santos et al. (2010) analisaram o perfil dos pesquisadores com bolsas de produtividade em pesquisa do CNPq na área de química. Os dados dos pesquisadores foram extraídos da Plataforma Lattes em Março de 2009. Foi constatado que 67,2% dos pesquisadores são do sexo masculino, 59,4% atuam em instituições federais e 63,7 % estão localizados na região sudeste. Os autores observaram que não há homogeneidade no perfil dos pesquisadores dentro de uma mesma categoria/nível, evidenciando que o processo de distribuição das bolsas ainda não é totalmente transparente.

Alves et al. (2011) apresenta o perfil dos pesquisadores bolsistas de produtividade em pesquisa do CNPq nas áreas de Engenharia de Produção e de Transportes atuantes na subárea de Pesquisa Operacional. Os dados foram extraídos da Plataforma Lattes de forma automática, utilizando a linguagem LattesMiner, no período de Abril a Maio de 2011. Foram analisados atributos como: gênero, categoria/nível da bolsa, instituição, tempo de doutorado, produção científica e orientações concluídas de mestrado e doutorado. Do total de 186 pesquisadores, 91 (48,92%) pertencem à subárea de Pesquisa Operacional. A maioria (75,83%) desse pesquisadores são bolsistas categoria 2 e 1D. O tempo médio de doutorado é de 15,19 anos. O gênero masculino predomina com 71,43% e 78,02% atuam na região sudeste.

Picinin et al. (2013) analisam a produção científica do triênio 2007-2009 dos bolsistas de produtividade em pesquisa do CNPq da área de engenharia de produção, quanto aos critérios gerais para a concessão da bolsa. A fonte de informação utilizada foi a Plataforma Lattes. O trabalho analisou atributos relativos à liderança em grupos de pesquisa, publicações em periódicos, produções técnicas e orientações concluídas. Os autores observaram que há correlação entre a categoria/nível da bolsa e o volume de publicações em periódicos. Entretanto, na área pesquisada, o nível 1C apresentou produção científica mais consistente que o nível 1A. Os autores também destacaram fatores qualitativos, como formação de recursos humanos e liderança em grupos de pesquisa, como critérios que definem a concessão ou não das bolsas.

Alves et al. (2014) mostram um estudo recente dos bolsistas de produtividade em pesquisa do CNPq na área de química, comparando os resultados obtidos com Santos et al. (2010). Os dados foram extraídos da Plataforma Lattes de forma automática, utilizando a linguagem LattesMiner, entre 2002 e 2011. Os autores confirmaram que a maioria dos pesquisadores são do sexo masculino (67,9%), com bolsa categoria 2 (63,2%) e atuando na região sudeste (60,7%).

Mais recentemente, (LIMA et al., 2015) avaliaram a pesquisa de excelência na área de Ciência da Computação. Foram analisados dados quantitativos e qualitativos. Assim como os demais trabalhos, o conjunto de dados foi obtido através da Plataforma Lattes. Foram avaliados três dimensões principais: tempo de carreira, número de estudantes orientados e volume de publicações e citações. O artigo mostra que é preciso ir além da produção científica para avaliar a qualidade dos pesquisadores. Portanto, se faz necessário também avaliar aspectos como tempo de carreira e orientações.

Este artigo se diferencia dos demais trabalhos apresentados nesta seção por algumas razões. Utiliza-se toda a base dos currículos Lattes, portanto há uma dificuldade em processar esses dados. Todas as áreas são analisadas concomitantemente. Dessa forma é obtido o panorama da pesquisa de excelência brasileira com uma visão mais ampla.

3 A BOLSA DE PRODUTIVIDADE EM PESQUISA DO CNPq

A bolsa de produtividade em pesquisa do CNPq(PQ) é destinada aos pesquisadores que se destacam entres seus pares e segue critérios normativos gerais, estabelecidos pelo CNPq, e específicos, pelos Comitês de Assessoramento (CAs) do CNPq. Dessa forma, o CNPq premia os pesquisadores segundo sua produção científica e formação de recursos humanos com uma mensalidade que varia de acordo com a categoria/nível do mesmo.

A avaliação em ciência e, em particular, dos pesquisadores, leva em conta muitas dimensões, objetivas e subjetivas. Dentre as dimensões objetivas (quantitativas), estão a produção científica e o impacto da mesma. As medidas de produção científica de um pesquisador incluem o número total de artigos publicados em toda a carreira, podendo ser ponderadas por fatores que indicam a “qualidade” do veículo de publicação. Esses fatores são atribuídos à revistas indexadas e são denominados de Qualis. Medidas de impacto avaliam o quanto a produção do pesquisador teve consequências para a sua área de pesquisa, ou na sociedade em geral. A medida mais tradicional de impacto é o número de citações. Há outras medidas que relacionam a produção com o número de citações, tais como citações médias por artigo ou índice h (HIRSCH, 2005). Outras medidas qualitativas são a formação de recursos humanos, na forma de orientações de mestrado e doutorado, bem como organização de grupos de pesquisa e coordenação de projetos.

A avaliação dos pesquisadores é tradicionalmente feita por pares. As concessões das Bolsas de Produtividade em Pesquisa do CNPq, financiamento de pesquisa, banca de concursos, prêmios acadêmicos, entre outros, são destinados aos pesquisadores que se destacam entre seus pares. Esse tipo de avaliação baseia-se em critérios quali-quantitativos. Dessa forma, os avaliadores observam aspectos além da produção e fatores de impacto. Os critérios subjetivos incluem o potencial da pesquisa proposta, entre outros.

O CNPq é organizado em comitês de assessoramento para cada subárea de atuação e conhecimento. Os CAs são compostos por mais de 300 pesquisadores reconhecidos de cada subárea, escolhidos periodicamente por um conselho deliberativo, que seleciona os membros com base em consultas feitas à comunidade científico-tecnológica nacional. Uma das atribuições do CA é julgar as propostas de pesquisa e atribuir níveis da bolsa de produtividade. Os CAs são organizados em grandes áreas e cada subárea tem um critério de julgamento diferente.

A Bolsa de Produtividade em Pesquisa é organizada em duas modalidades, 1 e 2, sendo que a categoria 1 possui quatro níveis, em ordem crescente: 1D, 1C, 1B e 1A. Cada nível provê uma mensalidade crescente, havendo um salto maior entre a categoria 2 e o nível 1D. Além da complementação salarial, há outras vantagens em ser pesquisador categoria 1. Por exemplo, há um adicional de bancada para despesas relacionadas ao projeto de pesquisa. Além disso, apenas pesquisadores categoria 1 podem ser membros dos CAs e apenas eles participam das consultas do conselho deliberativo à comunidade científica.

As Bolsas de Produtividade em Pesquisa tem duração de 3 a 5 anos. Ao final da vigência da bolsa, o pesquisador deve fazer uma nova submissão. Essa submissão é avaliada e, sendo aprovada, o pesquisador recebe o nível da avaliação e a bolsa correspondente. Normalmente a avaliação é uma das seguintes alternativas:

- O pesquisador perde a bolsa;
- O pesquisador passa a ser classificado em um nível inferior;

- O pesquisador mantém o mesmo nível;
- Ou o pesquisador sobe um nível.

Há alguns requisitos e critérios mínimos exigidos pelo CNPq para a concessão das bolsas, por categoria e por nível. O requisito mínimo por categoria é o tempo de doutoramento para o pesquisador se tornar elegível para receber a bolsa, sendo necessários, pelo menos, 3 anos para a categoria 2 e 8 anos para a categoria 1. Para a categoria 2, onde não há especificação de nível, é avaliada a produtividade do pesquisador, com ênfase nos trabalhos publicados e orientações, ambos referentes aos últimos 5 anos.

Para a categoria 1, onde há 4 níveis (A,B,C e D), o pesquisador é submetido a uma comparação entre os seus pares e nos dados dos últimos 10 anos de pesquisa, tendo como critério geral a capacidade de formação contínua de recursos humanos. A diferenciação entre os níveis A, B, C e D é baseado nos critérios relacionados a seguir, sendo que os 6 últimos devem ser considerados quando pertinentes: mérito científico do projeto; relevância, originalidade e repercussão da produção científica do candidato; formação de recursos humanos em nível de Pós-Graduação; contribuição científica, tecnológica e de inovação, incluindo patentes; coordenação ou participação em projetos e/ou redes de pesquisa; inserção internacional do proponente; participação como editor científico; participação em atividades de gestão científica e acadêmica; foco nos grandes problemas nacionais; abordagens multi e transdisciplinares; impacto social; comunicação com a sociedade; interação com o parque produtivo e conservação ambiental.

Esses critérios tem maior peso na análise da pesquisa, mas cada CA tem a autonomia de incluir outros itens que julgar importante para a área de pesquisa. A recomendação do CNPq é que cada CA deve privilegiar a qualidade e o conjunto da obra do pesquisador.

Especificamente para os níveis C e B, além dos critérios supracitados, é avaliada a contribuição na organização de grupos de pesquisa e programas de graduação e pós-graduação. O CNPq também espera que os pesquisadores desses níveis participem de forma significativa em atividades de pesquisa nas instituições em que atuam. Lembrando que no caso dos pesquisadores de nível B, é esperado a participação em órgãos de fomento à pesquisa.

O nível A é reservado a candidatos que tenham mostrado excelência continuada na produção científica e na formação de recursos humanos, e que liderem grupos de pesquisa consolidados. O perfil deste nível de pesquisador deve, na maior parte dos casos, extrapolar os aspectos unicamente de produtividade para incluir aspectos adicionais que mostrem uma significativa liderança dentro da sua área de pesquisa no Brasil e capacidade de explorar novas fronteiras científicas em projetos de risco.

4 Plataforma Lattes

Para a análise do perfil dos Bolsistas de Produtividade do CNPq foi utilizado os dados da Plataforma Lattes como principal fonte de informação. A Plataforma Lattes, criada em 1997 pelo CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) com a participação do Grupo Stela do PPGE/UFSC, tornou-se grande conhecida da comunidade científica e tecnológica tanto no Brasil quanto em outros países americanos e europeus. Essa plataforma foi criada com o objetivo de integrar os sistemas de informação das agências federais, concentrando o processo de gestão de Ciência e Tecnologia (C&T), tanto do ponto de

vista do usuário quanto das agências de fomento e das instituições de ensino e pesquisa (CNPQ, 2014). O principal componente da plataforma é o Currículo Lattes.

Atualmente o sistema de Currículo Lattes armazena cerca de 4.300.000 de currículos de docentes, estudantes, pesquisadores e profissionais de diversas áreas do conhecimento com atuação em ciência, tecnologia e inovação. Esse sistema identifica cada currículo com um ID único, chamado de ID Lattes, composto por 16 dígitos. Com esse número é possível acessar determinado currículo utilizando mecanismos automáticos.

5 Identificação e Extração dos Currículos Lattes

Todo o processo de extração é realizado pelo *framework* proposto por Dias et al. (2013). Esse *framework* é responsável por realizar a extração de toda a base de Currículos Lattes ou de grupos específicos, sendo que, na última opção é necessário informar o conjunto de IDs a serem extraídos. O *framework* proposto é dividido em 3 partes principais, como segue:

Extração: O processo de aquisição inicia-se com uma requisição do extrator que faz a extração da lista dos identificadores dos usuários que possuem currículos cadastrados na plataforma. A lista gerada é armazenada em arquivo para posterior utilização do extrator de currículos.

É importante ressaltar que a requisição retorna uma listagem com os identificadores dos currículos por ordem de atualização, sendo os primeiros registros aqueles com maior tempo de atualização e os últimos são os currículos mais recentemente atualizados e cadastrados.

De posse do arquivo com a listagem atualizada contendo todos os identificadores dos currículos, o extrator de currículos permite selecionar uma quantidade de currículos a serem extraídos, extrair os currículos a partir de uma determinada data de atualização, ou todo o repositório.

Todos os currículos são armazenados em disco um a um até o fim da listagem enviada ao extrator. Para o armazenamento, alguns elementos não necessários para a análise dos currículos como foto e imagens como da bandeira do país de origem dos integrantes são ignoradas. A extração dos currículos é eficiente e permite por exemplo, que o processo seja interrompido e retomado de qualquer parte da lista de identificadores.

Transformação: A etapa de transformação é caracterizada por uma série de algoritmos que transformam os dados extraídos dos currículos lattes em formato estruturado que visa facilitar o uso destes em outras etapas da arquitetura.

Ainda no módulo de transformação, algoritmos são responsáveis por gerar consultas e análises dos currículos. É possível visualizar todos os currículos como um grande grafo ou gerar redes específicas a partir de termos que são informados pelos usuários e localizados nos currículos.

Para este trabalho podem ser utilizados vários elementos contidos nesta etapa de transformação, como trabalho publicados em anais de congressos, artigos publicados em revistas e periódicos, orientações, organização de eventos, local e instituição em que o pesquisador atua, entre outros.

Esta etapa também é responsável por fazer cálculos de algumas métricas e gerar redes específicas de pesquisadores. Ainda são gerados arquivos padronizados que podem ser

importados por ferramentas de visualização de grafos. Todos os resultados são enviados para a etapa de visualização, que é responsável pela interação entre toda a arquitetura e o usuário final.

Visualização: Esta etapa é caracterizada por apresentar aos usuários os resultados das análises e possibilitar a interação entre usuário e sistema.

Com o auxílio de uma interface gráfica o usuário pode escolher quais dados quer extrair, quais métricas calcular, pode selecionar toda a rede de currículos lattes ou somente um grupo específico de pesquisadores, como a grande área em que atua.

Após a geração da rede, é possível aplicar métricas de análise de redes, bem como visualizar relatórios bibliométricos.

Dessa forma, o *framework* foi utilizado para realizar a extração de todos os currículos da Plataforma Lattes e posteriormente identificar aqueles que possuem a bolsa de produtividade em pesquisa do CNPq. Após a identificação dos pesquisadores bolsistas, foram extraídas as informações sobre a produção científica para análise.

6 Metodologia

No presente artigo, é avaliado um conjunto bem definido de pesquisadores de acordo com diferentes aspectos. Para isso, foi necessário a criação de um conjunto de dados com diversas informações sobre cada pesquisador. As etapas realizadas ao construir o conjunto de dados que será utilizado para análise na Seção 7 são as seguintes:

Etapa 1: *Levantamento de informações sobre as bolsas e os bolsistas de produtividade em pesquisa do CNPq.* Atualmente existem cerca de 4.300.000 de currículos cadastrados na Plataforma Lattes, divididos em 46 áreas do conhecimento (CNPQ, 2014). Considerando todos os potenciais pesquisadores contidos na Plataforma Lattes, foi definido que o presente artigo trataria o subconjuntos dos pesquisadores que possuem bolsa de produtividade em pesquisa do CNPq. Tais bolsas, divididas em 5 níveis, são concedidas através de critérios avaliados por 46 Comitês de Assessoramento.

Existem 14.154 pesquisadores de todas as áreas distribuídos entre os 5 níveis da bolsa, como é mostrado na Figura 1. Note de 46% dos pesquisadores estão nas modalidades 1A-1D e o restante, 54%, estão na modalidade 2. É importante salientar que a modalidade 2 é a porta de entrada para a bolsa, o qual se concentra a maior parte dos bolsistas.

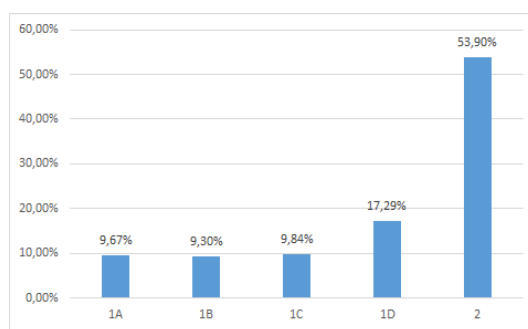


Figure 1: Distribuição dos bolsistas por modalidade

Etapa 2: *Organização dos dados dos bolsistas de produtividade em pesquisa do CNPq.* Os pesquisadores devem preencher um currículo na Plataforma Lattes para entrar no programa de bolsas. Nessa etapa foi construído o perfil de cada um dos 14.154 pesquisadores extraído da Plataforma Lattes. Sendo que o perfil traçado é composto pela formação acadêmica, produção científica e orientação de estudantes. A Figura 2 mostra a distribuição dos pesquisadores bolsistas por grande área do conhecimento.

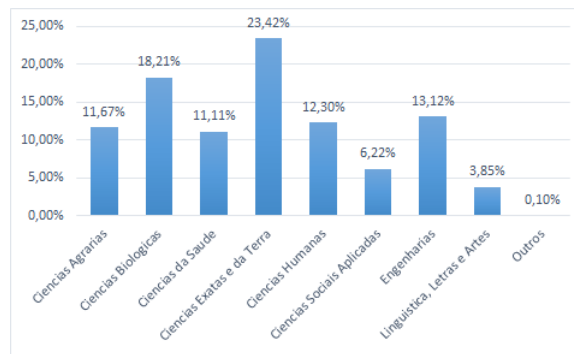


Figure 2: Distribuição por grande área

Etapa 3: *Tratamento do perfil dos pesquisadores.* Os dados referentes à produção científica precisam ser tratados antes de serem analisados. As publicações de cada autor podem não ser agrupadas devido a alguma ambiguação de nomes: o mesmo autor pode aparecer com nomes diferentes (FERREIRA et al., 2012), ou autores diferentes podem ter nomes similares, causando divisão ou mistura de citações (LEE et al., 2007). Para lidar com ambiguidade de nomes presentes no conjunto, foi utilizado um método estado-da-arte desenvolvido por Dias e Moita (2015), chamado ISColl.

7 RESULTADOS

A fim de caracterizar uma visão geral do bolsistas PQ, são avaliados diversos aspectos relativos à carreira dos pesquisadores. Especificamente, duas variáveis são importantes e fáceis de definir: tempo de carreira e estudantes orientados. Primeiramente foi analisado como o tempo de carreira impacta na produtividade do pesquisador. Consideremos que o tempo de carreira inicia-se com o ano em que o pesquisador terminou seu doutorado.

A Figura 3 mostra a distribuição de tempo desde que o pesquisador obteve o grau de doutor, agrupados pela modalidade da bolsa. Observa-se que o valor da mediana aumenta gradativamente da modalidade 2 em direção à modalidade 1A, formando um caminho de evolução. Consequentemente, pode-se dizer que os pesquisadores mais experientes pertencem às modalidades mais altas. É importante salientar a existência de *outliers* na modalidade 2, evidenciando que o tempo de carreira por si só não é suficiente para alcançar níveis mais altos.

Observa-se que o intervalo entre as medianas das modalidades 2 até 1B são bem próximos. A diferença maior está entre a modalidade 1B e 1A. Para obter o nível 1A, o pesquisador deve ter contribuições nacionais e internacionais para a comunidade acadêmica, o que geralmente demora mais tempo e esforço.

A Figura 4 mostra o número de mestres e doutores orientados pelos pesquisadores, distribuídos entre as cinco modalidades da bolsa. Verifica-se que o valor da mediana diminui ao

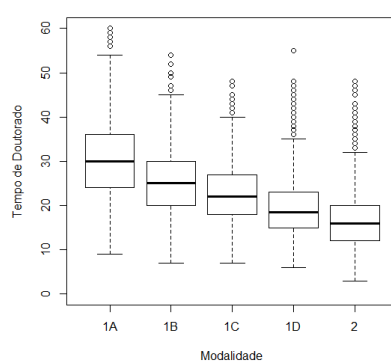
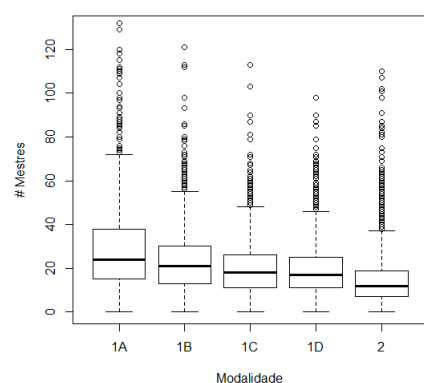
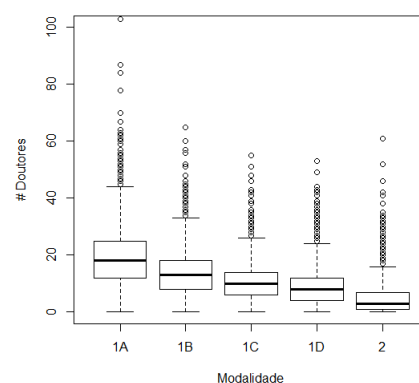


Figure 3: Tempo de Doutorado

longo das modalidades, partindo da modalidade 1A até a 2, em ambos os gráficos. Na Figura 4a é possível observar que as medianas estão próximas em todas as modalidades. Já na Figura 4b, há uma diferença significativa entre as medianas das modalidades 1A e 1B. Essa diferença pode ser explicada pelo fato de que os pesquisadores pertencentes ao nível 1A geralmente ocupam cargos administrativos (direção, coordenação, entre outros). Além disso, a orientação de estudantes não é a única prioridade para os pesquisadores nos níveis mais altos.



(a) Número de Mestres Formados



(b) Número de Doutores Formados

Figure 4: Quantidade de Mestres e Doutores formados

Outro aspecto importante para se caracterizar o perfil dos pesquisadores bolsistas é o volume da produção científica, que é composta por artigos em anais de congressos e em periódicos, livros e capítulos de livros. A Figura 5 mostra a o volume da produção de artigos em anais de congressos, na qual observa-se que o valor das medianas estão bem próximas entre as modalidades 1A-1D, diminuindo consideravelmente na categoria 2.

A presença de *outliers*, embora poucos, em todas as modalidades da bolsa mostra que existem alguns pesquisadores com produção acima da média para a modalidade. Os Comitês de Assessoramento consideram outros fatores, além do volume de publicação, como critérios de avaliação. Dentre esses critérios estão formação de recursos humanos e impacto das publicações

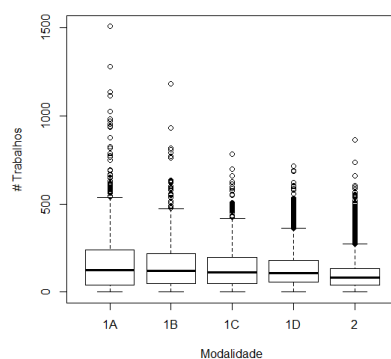


Figure 5: Artigos em anais de congressos

na sociedade em geral.

A diferença de produção da modalidade 1A para as demais fica ainda mais evidenciada ao observar a produção de artigos em periódicos, mostrada na Figura 6. Nota-se que o valor da mediana diminui gradativamente da modalidade 1A até a modalidade 2. Sendo que a diferença é mais expressiva entre as modalidades 1A e 1B. Também é importante salientar a presença de *outliers* em todas as modalidades.

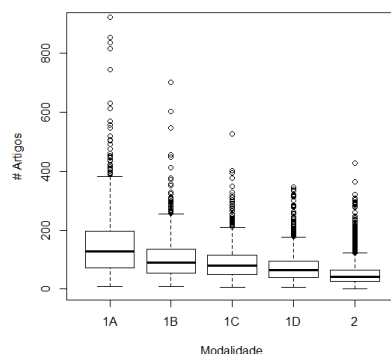


Figure 6: Artigos em periódicos

Quanto mais alto o nível da bolsa, maiores são os critérios para se manter naquele nível. Dessa forma, a tendência natural é que os pesquisadores com níveis mais altos tenham uma média de publicações maior que os demais. Além disso, pesquisadores de nível 1A, por exemplo, tendem a orientar mais alunos de mestrado e doutorado, o que eleva a produção científica do pesquisador.

É importante considerarmos a produção de capítulos e livros completos para a caracterização do perfil dos pesquisadores. Dessa forma, a Figura 7 mostra a produção de capítulos de livros. Podemos notar que o valor da mediana é bem próximo entre todas as modalidades. Sendo que a modalidade 1A possui a maior mediana e esse valor decai conforme diminui o nível da bolsa.

Por fim, foi analisado a produção referente aos livros completos. De maneira geral, como pode ser visto na Figura 8, a produção de livros é pequena em relação aos artigos. Mas vale

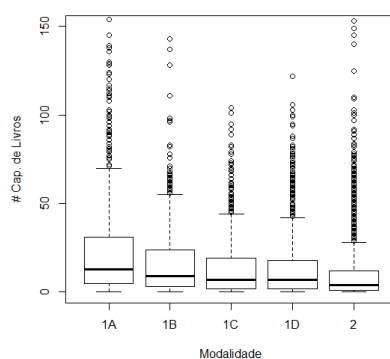


Figure 7: Capítulos de livros

lembrar que a publicação de um livro exige mais esforço e tempo dos autores. Novamente a modalidade 1A se destaca, sendo que pelo menos 75% dos pesquisadores publicaram mais de 10 livros durante a carreira. Em contrapartida, na modalidade 2 esse valor cai para 3.

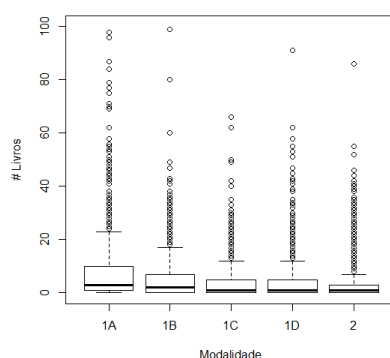


Figure 8: Livros

8 CONSIDERAÇÕES FINAIS

Este artigo apresentou uma visão geral do perfil dos bolsistas de produtividade em pesquisa do CNPq e o volume da produção científica em cada nível da bolsa. Foi observado que a maioria dos pesquisadores são bolsistas na modalidade 2. Tal modalidade é a porta de entrada dos pesquisadores no programa de bolsas fomentado pelo CNPq.

Dos pesquisadores do nível 1A, grande parte possui de 20 a 40 anos de doutorado, sendo que 36,31% possuem entre 30 e 40 anos. Já os níveis 1B e 1C possuem um número maior de pesquisadores entre 15 e 20 anos de doutorado, entretanto se concentram na faixa de 20 a 25 anos. O nível 1B possui maior homogeneidade entre as faixas. O nível 1D possui pesquisadores entre 10 e 25 anos de doutorado. Por fim, a categoria 2 engloba pesquisadores entre 5 de 25 anos de doutorado, sendo que 33,62% possuem de 10 a 15 anos.

É importante ressaltar os resultados obtidos no tempo de titulação da categoria 2 e do nível 1D. Para que um pesquisador possa ser elegível à categoria 2 são necessários 3 anos

de titulação de doutorado e para a categoria 1, 8 anos. Observou-se que no nível 1D apenas 2% dos pesquisadores possuem menos de 10 anos de doutorado. Na categoria 2, apenas 0,39% possuem menos de 5 anos de doutorado e 16,02% possuem entre 5 e 10 anos.

A quantidade mestres orientados é bem próxima entre todas as modalidades da bolsa. Por outro lado, os valores dos doutores orientados foi consideravelmente mais alto no nível 1A.

Em relação à distribuição da produção científica entre as modalidades, foi observado uma distribuição esperada na produção de artigos em periódicos, em que o valor da mediana diminui na direção dos níveis mais baixos, sendo que o nível 1A se destaca. Entretanto, ao considerar a produção de artigos em anais de congressos, os valores das medianas ficaram próximos nos níveis 1A-1D.

A produção de capítulos de livros mostrou valores de medianas próximas entre as modalidades 1B-1D, sendo que o nível 1A se destacou e a categoria 2 mostrou um valor de mediana consideravelmente mais baixo. Já a produção de livros foi observado que os valores das medianas próximos, mas diminuindo em direção dos níveis mais baixos. É notável que os valores ficaram próximos de 0 e que a maior parte dos pesquisadores publicou poucos livros durante a carreira.

Como melhoramento metodológico, em estudos futuros será considerado a análise de um perfil mais amplo dos pesquisadores utilizando variáveis qualitativas como volume de citações e índice-h.

Bibliography

ALVES, A. D.; YANASSE, H. H.; SOMA, N. Y. Perfil dos bolsistas pq das áreas de engenharia de producao e de transportes do cnpq: enfoque na subárea de pesquisa operacional. **XLIII Simpósio Brasileiro de Pesquisa Operacional, Ubatuba, SP**, v. 8, 2011.

ALVES, A. D.; YANASSE, H. H.; SOMA, N. Y. Profile of cnpq productivity research fellows in the chemistry area based on the lattes platform. **Química Nova**, SciELO Brasil, v. 37, n. 2, p. 377–383, 2014.

AMORIN, C. V. Curriculum vitae organization: the lattes software platform. **Pesquisa Odontológica Brasileira**, SciELO Brasil, v. 17, p. 18–22, 2003.

CNPQ. **CNPq: Sobre a Plataforma Lattes**. 2014. Disponível em: <http://lattes.cnpq.br/>.

DIAS, T. M. R.; MOITA, G. F. A method for the identification of collaboration in large scientific databases. **Em Questão**, UFRGS, 2015.

DIAS, T. M. R.; MOITA, G. F.; DIAS, P. M.; MOREIRA, T. H.; SANTOS, L. R. F. Modelagem e caracterização de redes científicas : Um estudo sobre a plataforma lattes. In: **Brazilian Workshop on Social Network Analysis and Mining (BraSNAM)**. [S.l.: s.n.], 2013.

FERREIRA, A. A.; GONÇALVES, M. A.; LAENDER, A. H. A brief survey of automatic methods for author name disambiguation. **Acm Sigmod Record**, ACM, v. 41, n. 2, p. 15–26, 2012.

HIRSCH, J. E. An index to quantify an individual's scientific research output. **Proceedings of the National academy of Sciences of the United States of America**, National Acad Sciences, v. 102, n. 46, p. 16569–16572, 2005.

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

LEE, D.; KANG, J.; MITRA, P.; GILES, C. L.; ON, B.-W. Are your citations clean? **Communications of the ACM**, ACM, v. 50, n. 12, p. 33–38, 2007.

LIMA, H.; SILVA, T. H.; MORO, M. M.; SANTOS, R. L.; JR, W. M.; LAENDER, A. H. Assessing the profile of top brazilian computer science researchers. **Scientometrics**, Springer, v. 103, n. 3, p. 879–896, 2015.

MARTELLI-JUNIOR, H.; MARTELLI, D. R. B.; QUIRINO, I. G.; OLIVEIRA, M. C. L. A.; LIMA, L. S.; OLIVEIRA, E. A. d. Cnpq researchers in medicine: a comparative study of research areas. **Revista Da Associacao Medica Brasileira**, SciELO Brasil, v. 56, n. 4, p. 478–483, 2010.

MENA-CHALCO, J. P.; JUNIOR, R. M. C. Prospecção de dados acadêmicos de currículos lattes através de scriptlattes. **Capítulo do livro Bibliometria e Cientometria: reflexões teóricas e interfaces (in press). São Carlos: Pedro e João**, p. 1–20, 2011.

PICININ, C. T.; PILATTI, L. A.; KOVALESKI, J. L.; PEDROSO, B. Critérios gerais para a concessão de bolsa produtividade no brasil: um estudo na engenharia de produção. **Interciencia**, Asociación Interciencia, v. 38, n. 11, p. 785–792, 2013.

SANTOS, N. C. F.; CANDIDO, L. F. d. O.; KUPPENS, C. L. Produtividade em pesquisa do cnpq: análise do perfil dos pesquisadores da química. **Química Nova**, SciELO Brasil, v. 33, n. 2, p. 489–95, 2010.