

CARACTERIZAÇÃO E ANÁLISE DOS PESQUISADORES BRASILEIROS COM ATUAÇÃO NA ÁREA DE CIÊNCIA DA INFORMAÇÃO

Email:

thiagomagela@gmail.com

lanilucas@gmail.com

adilson.pinto@ufsc.br

patriciamdias@gmail.com

Thiago Magela Rodrigues Dias¹, Elaine Rosângela Oliveira Lucas², Adilson Luiz Pinto³, Patrícia Mascarenhas Dias

RESUMO

Os estudos bibliométricos e cientométricos tem contribuído de forma significativa para a compreensão sobre a evolução das pesquisas científicas. Tais estudos que visam quantificar a produção científica, contribuem de forma significativa para análises que possibilitam a comparação entre conjuntos de pesquisadores, de instituições e ainda, sobre possíveis direcionamentos para o fomento de novas pesquisas. Neste contexto, compreender a produção científica, de um determinado conjunto de indivíduos, de uma instituição ou de uma determinada área de atuação, possibilita além de identificar o panorama dos trabalhos realizados, traçar estratégias para alavancar a produção científica do conjunto em análise. A grande dificuldade ou limitador deste tipo de análise, está nos repositórios de dados que em sua maioria é limitado a algumas áreas do conhecimento ou possuem apenas informações de repositórios de dados internacionais que podem negligenciar a produção científica de alguns conjuntos de pesquisadores, comprometendo de forma impactante nas análises realizadas. Logo, este trabalho tem como objetivo analisar o conjunto de pesquisadores que têm atuado na área de Ciência da Informação no Brasil, tendo como fonte de dados seus currículos cadastrados na Plataforma Lattes. Como resultado, é possível verificar como o conjunto tem publicado os resultados de suas pesquisas, como estão distribuídos pelas diversas instituições de ensino e pesquisa do país, além de analisar como eles têm colaborado. Os resultados aqui apresentados são relevantes pois apresentam de forma inédita um retrato geral de todos os indivíduos que têm atuado na área.

Palavras-Chave: *Ciência da Informação; Produção Científica; Plataforma Lattes.*

ABSTRACT

Bibliometric and scientometric studies have contributed significantly to the understanding of the evolution of scientific research. Such studies that aim to quantify the scientific production contribute significantly to analyzes that make it possible to compare groups of researchers, institutions and also possible directions for the promotion of new research. In this context, understanding the scientific production of a particular set of individuals, an institution or a specific area of activity, allows not only to identify the panorama of the works carried out, but also to outline strategies to leverage the scientific production of the set under analysis. The major difficulty or limitation of this type of analysis lies in data repositories that are mostly limited to some areas of knowledge or only have information from international data repositories that may neglect the scientific production of some sets of researchers, impact analysis. Therefore, this paper aims to analyze the set of researchers who have worked in the area of Information Science in Brazil, having as data source their curricula registered in the Lattes Platform. As a result, it is possible to see how the group has published the results of their research, as they are

¹ Programa de Pós-Graduação em Modelagem Matemática e Computacional.

Pesquisa em andamento apresentada no XIX ENANCIB – GT 7

Acesso ao Trabalho: <http://enancib.marilia.unesp.br/index.php/XIXENANCIB/xixenancib/paper/view/1582>

distributed by the various teaching and research institutions in the country, and analyze how they have collaborated. The results presented here are relevant because they present in an unprecedented way a general picture of all the individuals who have acted in the area.

Keywords: *Information Science; Scientific production; Lattes Platform.*

1 INTRODUÇÃO

A Ciência da Informação (CI) é o campo científico que busca resolver problemas de acesso, uso, controle, armazenamento e disponibilidade da produção informacional.

Para alguns teóricos sua origem se dá com influência da Documentação, com as bibliotecas públicas e o controle por sistemas de representação descritiva e temática, além de reunir o repertório mundial bibliográfico, conforme desejo de Otlet.

Seguindo a evolução terminológica e conceitual para designar o campo, temos como principais associações que tiveram nomes alterados, o International Institute of Bibliography alterado para International Institute for Documentation, em 1931, na Bélgica e depois em 1988 para International Federation for Information and Documentation (BUCKLAND, 2002); o American Documentation Institute passa a responder como American Society for Information Science (BORKO, 1968), nos Estados Unidos; e o Instituto Brasileiro de Bibliografia e Documentação que passa a chamar-se Instituto Brasileiro de Informação em Ciência e Tecnologia (CUNHA, 2005) no Brasil em 1976.

Um importante marco científico para a Ciência da Informação veio com o estudo de Borko, em 1968, “Ciência da Informação: o que é isto?” desenvolvido para atender um pedido da então Associação Americana para a Ciência da Informação. O autor afirma que CI “é aquela disciplina que investiga as propriedades e o comportamento da informação, as forças que governam o fluxo da informação e os meios de processamento para acesso e uso otimizados”. (BORKO, 1968, p. 3, Tradução nossa).

A Ciência da informação tem seu escopo focado no aporte interdisciplinar e dependente da tecnologia da informação, e que tem sua ativação na sociedade da informação, visando solucionar os problemas da grande massa documentação/informacional, ocasionada pela própria tecnologia de informação.

No cenário nacional brasileiro, a Ciência da Informação foi introduzida em 1970, por meio do curso de mestrado do chamado IBBD - que a partir de 1976 passa a se chamar IBICT.

Para Pinheiro e Loureiro (1995) é possível identificar o amadurecimento da Pós-Graduação da área no Brasil por meio da identificação de três fases do mestrado implantado: 1ª, de implantação (1970 – 1982); 2ª, transitória (1983 a 1986); e 3ª, de consolidação (a partir de 1987).

Ainda sobre os cursos, em especial no Brasil, o seu desenvolvimento teve estreita relação com os veículos de publicação da área, como a origem da revista Ciência da Informação e da Revista da Escola de Biblioteconomia da UFMG, ambas em 1972; e o Encontro Nacional de Pesquisa em Ciência da Informação (Enancib) com sua origem em 1994, em Belo Horizonte.

Na atualidade a Ciência da Informação brasileira está representada por 15 Mestrados Acadêmicos; oito Mestrados Profissionais e 11 Doutorados. Este conjunto é ofertado por 16 Universidades Federais e quatro Universidades Estaduais. O fortalecimento da pós-graduação em CI brasileira tem contribuído significativamente para a capacitação de profissionais que têm atuado na área nos diversos segmentos de atuação.

Neste contexto, diante da expansão no número de profissionais que têm atuado na área de CI no Brasil, esta pesquisa tem como objetivo realizar uma caracterização deste conjunto de profissionais apresentando informações sobre o nível de formação, características da produção científica e ainda, dados sobre a colaboração científica do conjunto analisado. Para tanto, foram utilizados como fonte de dados os currículos cadastrados na Plataforma Lattes.

2 TRABALHOS CORRELATOS

Mugnaini, Leite e Leta (2011) destacam que muito embora não se apresente como uma base de indexação e catalogação de periódicos, a Plataforma Lattes se constitui em uma fonte inesgotável de informações sobre a ciência brasileira, sob diversos aspectos e abordagens. Os autores ressaltam que, apesar de todo o volume de informações, o que se observa é ainda uma baixa frequência de estudos cientométricos, de especialistas brasileiros, que utilizam a Plataforma Lattes. E que isso tem reflexo nas limitações dos mecanismos de recuperação e extração das informações, constituindo-se em obstáculos para os estudos. Porém, a correlação de um grande volume de variáveis pessoais, acadêmicas e de produtividade, somadas a um acesso livre e gratuito, é um estímulo significativo para uso deste repositório pelos pesquisadores brasileiros da área. Destaca-se, ainda, o fato do repositório de dados reunir toda a produção científica brasileira, viabilizando análises que só seriam possíveis se conduzidas em diversas bases internacionais, o que representaria um custo considerável para tais estudos.

Para Ferraz, Quoniam e Maccari (2014), até o presente momento, não existe no mundo um repositório curricular nacional único semelhante à Plataforma Lattes, sendo que somente repositórios de dados referenciais, de onde se podem extrair referências bibliográficas e fontes de informação secundárias estão disponíveis para livre acesso. Dessa forma, a Plataforma Lattes pode servir como um instrumento de maior importância para o estudo da produção científica brasileira.

Cañibano e Bozeman (2009) destacam currículos acadêmicos como uma potencial fonte de informação extremamente abrangente e foco de trabalhos recentes para análises de grupos de pesquisadores. Estudos que utilizam currículos para a análise de redes sociais são ainda menos frequentes, porém deve-se considerar a gama de estudos sobre análise de coautoria e os efeitos das colaborações científicas na carreira do pesquisador (DIGIAMPIETRI; SANTIAGO; ALVES, 2013; LIMA et al., 2013; MENA-CHALCO; CESAR-JUNIOR, 2013).

É importante destacar que, apesar da riqueza de dados registrados na Plataforma Lattes, são disponibilizadas pela plataforma, apenas, simples consultas que possibilitam recuperar conjuntos de currículos, como, por exemplo, por formação acadêmica ou proficiência em algum idioma. Isto dificulta o processo de obtenção e conseqüentemente, análise dos dados, sejam de grupos específicos ou de todos os currículos cadastrados. Logo, algumas ferramentas têm sido propostas para auxiliar no processo de extração e análise dos currículos Lattes.

Sidone, Haddad e Mena-Chalco (2016) apresentam o papel da geografia na evolução da produção e colaboração científica no Brasil entre 1992 e 2009. Para os estudos, foram utilizados dados dos currículos da Plataforma Lattes de um milhão de pesquisadores. Os autores destacam o processo de desaceleração da produção científica brasileira a partir dos últimos triênios analisados. No trabalho, foi possível identificar que o padrão de crescimento das grandes áreas é semelhante nos triênios iniciais da análise, quando praticamente todas apresentaram produção em taxas duplicadas entre 1992-1994 e 1995-1997. No entanto, percebe-se que o crescimento desacelerado

ao longo do período é comum a todas as grandes áreas, porém em intensidades distintas. Como exemplo, verifica-se que a desaceleração no ritmo da produção científica ocorreu de maneira mais forte nos últimos triênios, primordialmente nas grandes áreas de Engenharias, Ciências Exatas e da Terra, Ciências da Saúde e Ciências Biológicas.

Diversos outros trabalhos têm explorado a Plataforma Lattes como principal fonte de dados para análise bibliográfica e de redes de colaboração, dentre os quais os de Pacheco et al. (2007), Cardoso e Machado (2008), Arruda et al. (2009), Alves, Yanasse e Soma (2011), Dias et al. (2013), Digiampietri, Mugnaini e Alves (2013), Roos et al. (2014), Araújo et al. (2014), Furtado et al. (2015), e Brito, Quoniam e Mena-Chalco (2016).

3 DESENVOLVIMENTO

Estudo de natureza quantitativa e aplicado, com foco em dois aspectos básicos: sendo o primeiro na reunião dos Currículos Lattes, recuperados da Plataforma Lattes; e o segundo em uma sistematização dos dados em um sistema de recuperação, armazenamento e tratamento dos dados.

A escolha do repositório da Plataforma Lattes está relacionada ao fato de tratar a integração de dados acadêmicos e profissionais de mais de 5 milhões de indivíduos distribuídos nos mais diversos níveis de formação e atuação (DIAS, 2016). A Plataforma Lattes caracteriza-se como uma fonte extremamente rica, pois integra dados científicos de currículos e de instituições de pesquisa, registrando os dados acadêmicos e as produções científicas dos pesquisadores e instituições, possibilitando que a atualização dos dados seja realizada pelos próprios pesquisadores.

Para a extração de todo o conjunto de currículos da Plataforma Lattes e seleção dos indivíduos com área de atuação vinculada a Ciência da Informação, foi utilizado o LattesDataXplorer, proposto e desenvolvido por Dias (2016), que foi responsável pela coleta, tratamento e padronização das análises.

No contexto deste trabalho, foram selecionados todos os indivíduos que possuíam como principal área de atuação, Ciência da Informação. Como resultado, este conjunto selecionado foi caracterizado e analisado.

4 RESULTADOS

Os dados de todos os currículos cadastrados na Plataforma Lattes foram coletados em janeiro de 2018, totalizando 5.502.175 currículos de todos os indivíduos, em que 14.030 destes currículos têm a Ciência da Informação como principal área de atuação. Logo, tendo como base o endereço profissional dos indivíduos, foi possível verificar as suas respectivas distribuições geográficas (Figura 1).

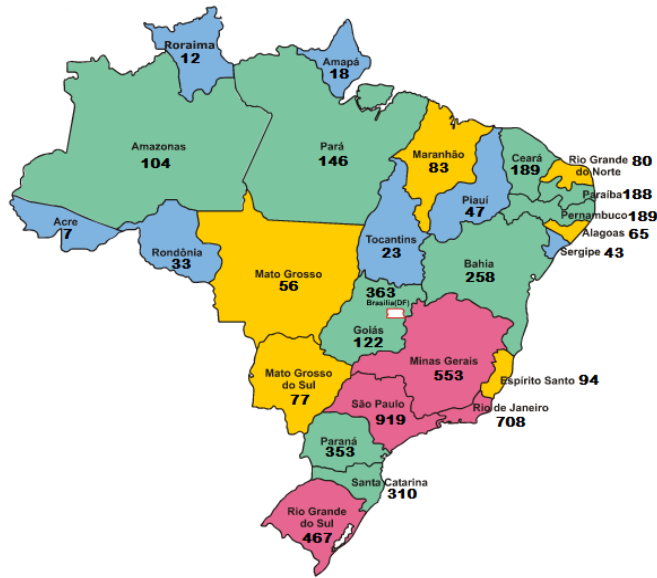


Figura 1: Distribuição geográfica dos indivíduos

Do conjunto selecionado, aproximadamente 40% dos currículos possuem endereço profissional registrado. É possível observar que o estado de São Paulo concentra a maior quantidade de profissionais, seguido dos estados do Rio de Janeiro e Minas Gerais, destacando-se ainda os estados do Rio Grande do Sul, Paraná e Santa Catarina. Tal distribuição é influenciada significativamente pela quantidade de cursos de pós-graduação de relevância na área de CI, em instituições de ensino consolidadas que se concentram nestes estados.

Já com relação ao perfil de publicação do conjunto, percebe-se que existe uma tendência de publicações de artigos em anais de congressos, seguido dos artigos em periódicos, bem como, textos em jornais e revistas, capítulos de livro e livros. Ao se analisar de forma temporal a produção científica do conjunto selecionado, percebe-se que, ao longo dos últimos anos houve uma queda na quantidade de publicações científicas (Figura 2).

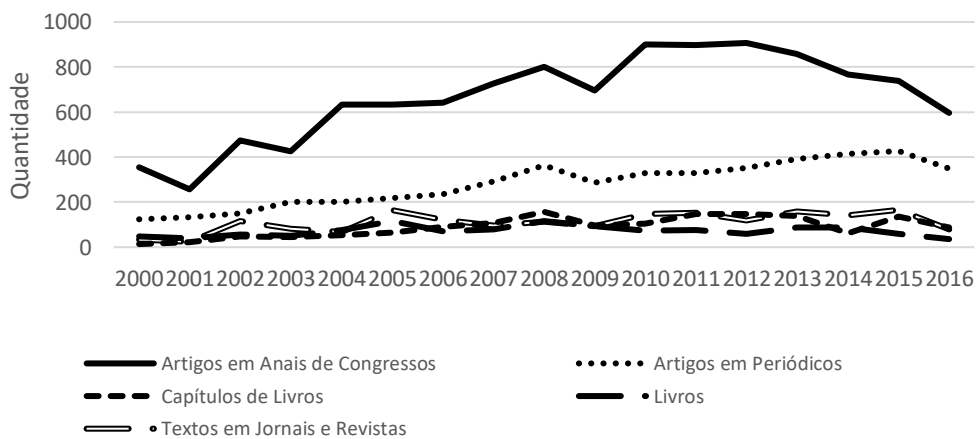


Figura 2: Evolução da produção científica dos profissionais da área de CI

Como pode ser observado, a maioria das publicações são de artigos em anais de congressos e artigos em periódicos. Especificamente com relação a publicação de artigos em anais de congressos, percebe-se que o auge de publicações aconteceu no ano de 2012 e posteriormente, inicia uma queda na quantidade de publicações, sendo bem acentuada entre os anos de 2015 e 2016, especificamente pela falta de recurso federal e estaduais neste período.

Já com relação aos artigos em periódicos, a quantidade de publicações tem crescimento linear a partir de 2009, tendo seu ápice em 2015, apresentando queda no ano de 2016. Tal situação pode estar relacionada a falta de atualização de alguns currículos, que na época da extração, poderiam não ter ainda, os registros das publicações mais recentes dos profissionais analisados.

Os outros tipos de publicações possuem crescimento semelhante, também apresentando queda no ano de 2016. Com ressalva para o ano de 2014 em que a publicação de Livros, diferentemente das outras produções, apresentou uma redução de aproximadamente 56% se comparado ao ano anterior.

Ao analisar a colaboração do conjunto, percebe-se que a rede de colaboração é pouco conectada, principalmente por ser influenciada pela distribuição geográfica do conjunto analisado e a distinção nos níveis de formação, em que uma grande quantidade de indivíduos possui grau 0, ou seja, nunca tiveram trabalhos em coautoria com outros indivíduos do conjunto (Figura 3).

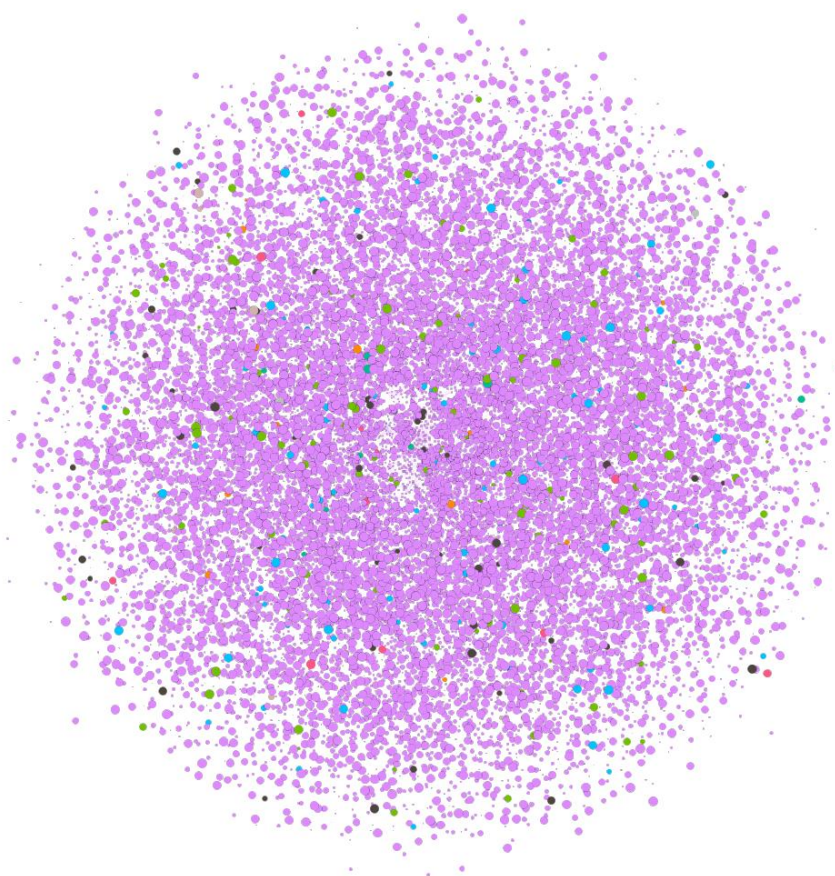


Figura 3: Rede de colaboração do conjunto analisado

Como pode ser observado, uma pequena quantidade de indivíduos, estão na componente conexas, em que, na grande maioria são de pesquisadores que estão atuando em instituições de ensino, trabalhando em colaboração com outros pesquisadores.

O grau médio da rede é baixo (0,482) influenciado de forma significativa pelos indivíduos com grau 0 (87,9%). Logo, uma análise considerando a componente gigante foi realizado.

A componente gigante é composta por apenas 3,71% dos indivíduos e eles em praticamente sua totalidade são de pesquisadores que já orientaram alunos em programas de pós-graduação. Foi possível ainda observar que grande parte dos grupos mais conectados são de pesquisadores de uma mesma instituição ou que tiveram algum relacionamento de orientação.

Logo, apesar da análise considerar um conjunto considerável de pesquisadores que estão atuando na área de Ciência da Informação, percebe-se que no conjunto analisado a colaboração científica não é tão representativa, reflexo do nível distinto de formação acadêmica e localização geográfica de atuação.

5 CONSIDERAÇÕES FINAIS

As análises aqui apresentadas são importantes pois possibilitaram obter uma visão sobre o conjunto de profissionais que têm atuado na área de CI, bem como a interferência que os Programas de Pós-Graduação têm no processo de produtividade e distribuição geográfico da área.

O comportamento científico da área está focado em publicar em publicações de maior impacto, como os eventos científicos reconhecidos e revistas científicas. Este fato se dá por dois motivos. O primeiro porque são os meios de publicações mais rápidos, e o segundo por que muitos destes meios de publicação não têm custo de publicação. Os eventos exercem os custos das matrículas e das viagens, mas não da submissão. Também existe uma lógica do decréscimo a partir de 2015, fato salientado pelo corte drástico de recursos financeiros do governo federal e estaduais para financiar a pesquisa científica nacional.

Finalmente, a distribuição geográfica de atuação se dá em especial pela implementação do maior número possível de instituições que contemplam cursos de Biblioteconomia, Arquivologia e Museologia, bem como os Programas de Pós-Graduação, que categoricamente estão nas regiões Sudeste e Sul do país, por este motivo o destaque para instituições destes centros.

REFERÊNCIAS

ALVES, A. D.; YANASSE, H. H.; SOMA, N. Y. LattesMiner: a multilingual DSL for information extraction from Lattes Platform. In: PROCEEDINGS OF THE COMPILATION OF THE CO-LOCATED WORKSHOPS ON DSM'11, TMC'11, AGERE!'11, AOPES'11, NEAT'11; VMIL'11. Portland, Oregon, USA: ACM, p. 85-92. 2011.

ARAÚJO, E. B. et al. Collaboration networks from a large cv database: dynamics, topology and bonus impact. PloS one, v. 9, n. 3, p. e90537, 2014.

ARRUDA, D. et al. Brazilian computer science research: Gender and regional distributions. *Scientometrics*, v. 79, n. 3, p. 651-665, 2009.

BORKO, H. Information science: what is it? *American Documentation*, v. 19, n. 1, p. 3-5, 1968. Disponível em: <https://goo.gl/RXwggR>. Acesso em: 16 jan. 2018.

BRITO, A. G. C.; QUONIAM, L.; MENA-CHALCO, J. P. Exploração da Plataforma Lattes por assunto: proposta de metodologia. *TransInformação*, v. 28, n. 1, p. 77-86, 2016.

BUCKLAND, M. 1895-2002 FID Achievements. California: Berkeley, 2002. Disponível em: <http://people.ischool.berkeley.edu/~buckland/fidhist.html>. Acesso em: 16 jan. 2018.

CAÑIBANO, C.; BOZEMAN, B. Curriculum vitae method in science policy and research evaluation: the state-of-the-art. *Research Evaluation*, v. 18, n. 2, p. 86-94, 2009.

CARDOSO, O. N. P.; MACHADO, R. T. M. Gestão do conhecimento usando data mining: estudo de caso na Universidade Federal de Lavras. *Revista de Administração Pública*, v. 42, n. 3, p. 495-528, 2008.

CUNHA, M. B. IBICT: 51 anos. *Ciência da Informação*, v. 34, n. 1, p. 7-8, 2005.

DIAS, T. M. R. et al. Modelagem e Caracterização de Redes Científicas: Um Estudo Sobre a Plataforma Lattes. In: BRAZILIAN WORKSHOP ON SOCIAL NETWORK ANALYSIS AND MINING, 2, Anais ... Maceió, 2013.

DIAS, T. M. R. Um estudo da produção científica brasileira a partir de dados da Plataforma Lattes. (Tese de Doutorado) - Pós-Graduação em Modelagem Matemática e Computacional, Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG). 2016.

DIGIAMPIETRI, L. A.; SANTIAGO, C. R. N.; ALVES, C. M. Predição de coautorias em redes sociais acadêmicas: um estudo exploratório em Ciência da Computação. In: BRAZILIAN WORKSHOP ON SOCIAL NETWORK ANALYSIS AND MINING, 2, 2013, Anais... Maceió, 2013.

DIGIAMPIETRI, L. A.; MUGNAINI, R.; ALVES, C. Analysis of Participation in Supervised Production of Advisors: A Case Study in Computer Science. In: BRAZILIAN WORKSHOP ON SOCIAL NETWORK ANALYSIS AND MINING, 2, 2013, Anais ... Maceió, 2013.

FERRAZ, R. R. N.; QUONIAM, L.; MACCARI, E. A. The use of ScriptLattes tool for extraction and on line availability of academic production from a departament of stricto sensu in management. In: INTERNATIONAL CONFERENCE ON INFORMATION SYSTEMS AND TECHNOLOGY MANAGEMENT, 11, São Paulo, 2014, Anais... São Paulo, p. 663-679, 2014.

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. e0141528, 2015.

LIMA, H. et al. Aggregating productivity indices for ranking researchers across multiple areas. In: PROCEEDINGS OF THE 13TH ACM/IEEE-CS JOINT CONFERENCE ON DIGITAL LIBRARIES, ACM, p. 97-106, 2013.

MENA-CHALCO, J. P.; CESAR-JUNIOR, R. M. Prospecção de dados acadêmicos de currículos Lattes através de scriptLattes. In: HAYASHI, M. C. P. I.; LETA, H. E. J. (Orgs.). Bibliometria e Cientometria: reflexões teóricas e interfaces. São Carlos: Pedro & João, p. 109-128, 2013.

MUGNAINI, R.; LEITE, P.; LETA, J. Fontes de informação para análise de internacionalização da produção científica brasileira. PontodeAcesso, v. 5, n. 3, p. 87-102, 2011.

PACHECO, R. C. et al. Uma análise da pesquisa em engenharia e ciências mecânicas no Brasil a partir dos dados da Plataforma Lattes. Rio de Janeiro, ABCM Engenharia, v. 12, n. 1, p. 16-22, 2007.

PINHEIRO, L. N. V. R.; LOUREIRO, J. M. M. Traçados e limites da ciência da informação. Ciência da Informação, v. 24, n. 1, p. 42-53, 1995.

ROOS, D. H. et al. Brazilian scientific production in areas of biological sciences: a comparative study on the modalities of full doctorate in Brazil or abroad. Scientometrics. v. 98, n. 1, p. 415-427, 2014.

SIDONE, O. J. G.; HADDAD, E. A.; MENA-CHALCO, J. P. A Ciência nas Regiões Brasileiras: Evolução da Produção e das Redes de Colaboração Científica. Transinformação. v. 28, n. 1, p. 15-31, 2016.