

A Plataforma Lattes como ferramenta para a padronização do nome de autores em estudos bibliométricos

Eixo temático: Técnicas e Ferramentas para Estudos Bibliométricos e Cientométricos

Modalidade: Pôster

1 INTRODUÇÃO

A Plataforma Lattes representa a experiência do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) na integração de bases de dados de Currículos, de Grupos de Pesquisa e de Instituições das áreas de Ciência e Tecnologia em um único Sistema de Informações. Criada em 1999, a base reúne atualmente cerca de 1,6 milhão de currículos, sendo 8% de doutores e cerca de 13% de mestres, além de contar com 4 mil instituições cadastradas (CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO, 2009; LANE, 2010).

Por ser uma base de dados de grande credibilidade, “mesmo que não se apresente como uma base de indexação e catalogação de periódicos, a Base Lattes se constitui em uma fonte inesgotável de informações sobre a ciência brasileira” (LETA; LEITE; MUGNAINI, 2008). É ferramenta essencial para a execução de estudos que necessitam de dados precisos, tais como a identificação dos autores de pesquisas científicas, seus objetos de estudo, filiação e suas relações de co-autoria. Segundo Café e Brascher (2008), a eficácia da descrição dos dados é o elemento fundamental da organização da informação, possibilitando o acesso a itens específicos que se encontram armazenados. Desta forma a padronização de dados é essencial para a realização dos estudos bibliométricos.

O presente estudo tem por objetivo demonstrar a importância da utilização da Plataforma Lattes como ferramenta na padronização dos dados que constituem as pesquisas bibliométricas. Quando não há uma padronização, a apresentação dos dados indexados nas bases pode incluir inúmeras imperfeições - como erros de grafia, ausência de componentes importantes do sobrenome dos autores, dentre outros - e levar o pesquisador a conclusões precipitadas, exigindo correção manual e acarretando no envolvimento de tempo em excesso na função de limpeza de dados.

2 METODOLOGIA

A partir de uma coleta de dados na *Web of Science* (WoS) para o projeto de pesquisa intitulado “A Ciência no Rio Grande do Sul: mapeamento da produção e colaboração nos anos de 2000 a 2009”, realizada em 19 de janeiro de 2012, foram obtidos 29.486 documentos em 144 áreas de conhecimento diferentes e 49.019 nomes de autores.

Gerou-se uma lista contendo as diferentes formas de como os nomes em citações de cada autor estão apresentados na WoS e, a partir da relação dos nomes em planilhas no *software Excel*, foram feitas buscas na Plataforma Lattes para checagem e correção dos nomes. No processo, percebeu-se que o Lattes constitui-se numa importante ferramenta para confrontar nomes e proceder às correções necessárias, conforme se apresenta a seguir.

3 RESULTADOS

No processo de pesquisa observou-se que o nome de alguns autores se apresenta de diferentes formas, de acordo como os registros são indexados nas bases de dados. Foram encontrados, por exemplo, erros de grafia ou dados incompletos para a identificação do autor. Através da busca na plataforma foi possível reconhecer a forma correta de apresentação da identificação dos autores, como demonstrado no quadro a seguir:

Quadro 1- Diferentes formas de identificação para uma mesma autoria:

Nº artigos	Identificação	Filiação
X	Rocha J	Univ Fed Santa Maria, Ctr Ciencias Naturais & Exatas, Dept Quim, BR-97
Y	Rocha JB	Univ Fed Santa Maria, Ctr Ciencias Naturais & Exatas, Dept Quim, BR-97
Z	Rocha JT	Univ Fed Santa Maria, Ctr Ciencias Naturais & Exatas, Dept Quim, BR-97
295	Rocha JBT	Univ Fed Santa Maria, Ctr Ciencias Naturais & Exatas, Dept Quim, BR-97

Fonte: Plataforma Lattes. Dados dos autores.

Nota: o somatório de X, Y e Z resulta em 295 artigos para João Batista Teixeira da Rocha, cujo Currículo Lattes está disponível em: <<http://buscatextual.cnpq.br/buscatextual/visualizacv.do?id=K4782281H2>>.

A plataforma contribui para resolver a dúvida sobre quem é o autor de um determinado trabalho quando há um conflito entre dois nomes com mesmo sobrenome, mesmas iniciais, e mesma instituição de filiação, como o exemplo abaixo. Nesse caso, a checagem é feita através do título do trabalho.

Quadro 2 - Identificação de autoria com conflito no nome em citações e mesma filiação:

Nº artigos	Identificação	Nome	Filiação
2	Krewer C	Carina da Costa Krewer	Univ Fed Santa Maria, Bacteriol Lab, BR-97105900 Santa Maria, RS, Brazil
5	Krewer CC	Cristina da Costa Krewer	Univ Fed Santa Maria, Bacteriol Lab, BR-97119900 Santa Maria, RS, Brazil

Fonte: Plataforma Lattes. Dados dos autores.

Nota: dados disponíveis em: <<http://buscatextual.cnpq.br/buscatextual/visualizacv.do?id=K4559462T1>> e em <<http://buscatextual.cnpq.br/buscatextual/visualizacv.do?id=K4772903U7>>, respectivamente.

A base também auxilia para a identificação dos autores através do seu conteúdo e de suas redes de co-autoria. De acordo com Geraldo Sorte, através da funcionalidade das redes de colaboração é possível visualizar graficamente a rede de co-autoria de um pesquisador que também possua o Currículo Lattes (CNPq, 2009). O quadro a seguir demonstra a relação de co-autoria vinculada ao título do trabalho, identificando o nome completo destas duas autoras:

Quadro 3: Identificação do nome completo através da rede de co-autoria:

Currículo do Sistema de Currículos Lattes (Carina da Costa Krewer) - Windows Internet Explorer	
http://buscatextual.cnpq.br/buscatextual/visualizacv.do?id=K4559462T1	
Dados Gerais	Projetos Áreas Produção em C,T & A Eventos
Resumos publicados em anais de congressos	
1.	VARGAS, A. C. ; KOLLING, L. ; COSTA, M. M. ; KREWER, Carina da Costa ; KREWER, Cristina da Costa ; MABONI, F. . Phylogenetic and pathotype analysis of <i>Escherichia coli</i> swine isolates from Southern Brazil. In: <i>VIA International Conference</i> , 2009, Londres. <i>VIA International Conference</i> , 2009.
2.	FARIAS, L. D. ; KOLLING, L. ; KREWER, Carina da Costa ; COSTA, M. M. ; VARGAS, A. C. . Susceptibilidade antimicrobiana de amostras de <i>Rhodococcus equi</i> do Brasil. In: <i>25º Congresso Brasileiro de Microbiologia</i> , 2009, Porto de Galinhas. <i>Anais do 25º Congresso Brasileiro de Microbiologia</i> , 2009.

Fonte: Plataforma Lattes.

A falta de padronização dos dados influencia no resultado de estudos bibliométricos, implicando na necessidade da limpeza manual dos dados coletados. Autores com diferentes formas na apresentação de seu nome elevam o número total de registros de autoria indexados, porém o número de documentos produzidos pelos mesmos não se modifica em função das discrepâncias nominais.

Constatou-se até o momento uma redução que se aproxima a 10% no número total de autores da pesquisa partindo da comparação do número total com o dado numérico referente às diferentes formas de apresentação do nome de um mesmo autor.

4 CONSIDERAÇÕES FINAIS

Além de afetar a qualidade da recuperação da informação, dados incorretos também influenciam na contagem do número total de autores durante um estudo bibliométrico e, desta forma, podem levar o pesquisador à conclusões equivocadas.

A padronização da apresentação dos nomes dos autores nas bases de dados resulta em melhorias na qualidade de dados disponibilizados para estudos bibliométricos e na economia de tempo de pesquisa empreendido na tarefa de limpeza de dados.

REFERÊNCIAS

CAFE, L; BRASHER, M. Organização da Informação e Bibliometria. **Encontros Bibli**, Florianópolis, 2008, p. 54-75. Número Especial.

CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO. **CNPq lança nova versão do Currículo Lattes**. Brasília: 2009. Disponível em: <<http://memoria.cnpq.br/saladeimprensa/noticias/print/09/1119c.htm>>. Acesso em: 24 maio 2012.

CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO. **Plataforma Lattes**. Disponível em: <<http://lattes.cnpq.br/>>. Acesso em: 24 maio 2012.

LANE, J. Let's make science metrics more scientific. **Nature**, Virginia, v. 464, n. 25, p. 488-489, Mar. 2010. Disponível em: <<http://www.nature.com/nature/journal/v464/n7288/full/464488a.html>>. Acesso em: 24 maio 2012.

LETA, J; LEITE, P; MUGNAINI, R. Fontes de informação para análise de internacionalização da produção científica brasileira. **PontodeAcesso**, Salvador, v. 5, n. 3, p. 87-102, ago. 2011. Disponível em: <<http://www.brapci.ufpr.br/documento.php?dd0=0000011715&dd1=8bd98>>. Acesso em: 24 maio 2012.