

XX ENANCIB

21 a 25 Outubro/2019 – Florianópolis

A Ciência da Informação e a era da Ciência de Dados

ISSN 2177-3688

GT-7 – Produção e Comunicação da Informação em Ciência, Tecnologia & Inovação

INDICADORES DA PRODUÇÃO CIENTÍFICA A PARTIR DE CONTROLE DE AUTORIDADE: O CASO UFSCAR

SCIENTIFIC PRODUCTION INDICATORS FROM AUTHORITY CONTROL: THE UFSCAR CASE

José Eduardo dos Reis - Universidade Federal de São Carlos - UFSCar
Rafaela Cristina de Camargo - Universidade Federal de São Carlos - UFSCar
Denilson de Oliveira Sarvo - Universidade Federal de São Carlos - UFSCar
Roniberto Morato do Amaral - Universidade Federal de São Carlos - UFSCar

Modalidade: Trabalho Completo

Resumo: A utilização das bases de dados científicas e ferramentas computacionais na mineração de dados potencializam os estudos métricos. Porém, a falta de um controle de autoridade, para a correta identificação da Instituição de Ciência e Tecnologia (ICT) e dos autores, no ato de indexação dos artigos científicos nas bases de dados, dificulta o desenvolvimento e compromete a qualidade dos resultados desses estudos. O objetivo do trabalho foi elaborar um tesouro de forma automatizada para o controle de autoridade aplicado às ICTs. O estudo de caso exploratório foi utilizado como método de pesquisa e a bibliometria automatizada como técnica de análise da informação. A unidade-caso foi a Universidade Federal de São Carlos e a amostra analisada abrangeu a produção científica dos docentes da unidade-caso registrada na Plataforma Lattes e na base de dados multidisciplinar *Web of Science*. Como resultado foi possível: 1) a partir do *Digital Object Identifier* (DOI) e da mesclagem dos campos "ISSN, ano, volume e página inicial", identificados na produção científica registrada na Plataforma Lattes e cotejados na *Web of Science*, localizar de forma automatizada as variações de redação do nome da unidade-caso, e gerar um tesouro para um efetivo controle de autoridade; 2) gerar um conjunto de indicadores de produção científica a partir da aplicação do tesouro da unidade-caso e analisar a sua eficácia. Considera-se que os resultados alcançados poderão contribuir para a legitimidade dos estudos métricos, e ainda, para a visibilidade da ICT, ao potencializar a recuperação da sua produção científica nas bases de dados.

Palavras-Chave: Indicadores de ciência e tecnologia. Controle de autoridade. Bibliometria. Estudos métricos. Produção científica.

Abstract: *The use of scientific databases and computational tools in data mining enhances metric studies. However, the lack of authority control for the correct identification of the Institute of Science and Technology (IST) and the authors, in the act of indexing the scientific articles in the databases, hinders the development and compromises the quality of the results of these studies. The aim of this work was to elaborate an automated thesaurus for the authority control applied to ISTs. The exploratory case study was used as a research method and automated bibliometrics as the information analysis technique. The case unit was the Federal University of São Carlos and the sample analyzed covered the scientific production of its professors, which was registered in the Lattes Platform and the multidisciplinary Web of Science database. As a result it was possible to: 1) automate the wording variations of the name of the case unit, and generate a thesaurus for effective authority control from the Digital Object Identifier (DOI) and the merge of the "ISSN, year, volume and*

home page" fields, identified in the scientific production registered in the Lattes Platform and collated in the Web of Science; 2) generate a set of indicators of scientific production from the application of the thesaurus of the case unit and analyze their effectiveness. It is considered that the results achieved may contribute to the legitimacy of metric studies, and also to the visibility of the IST, by enhancing the retrieval of its scientific production.

Keywords: Science and Technology Indicators; Authority control; Bibliometrics; Metric studies; Scientific production.

1 INTRODUÇÃO

Em 1977, Lancaster previu que não usaríamos mais papel. As bases de dados são possivelmente um dos “braços” dessa mudança, contribuindo para o aumento e visibilidade da produção científica, já que a disponibilidade online facilita o acesso e a recuperação da informação em uma diversidade de formatos (LANCASTER, 1978).

Atualmente a comunidade científica é atendida por iniciativas públicas e/ou privadas, de acesso aberto e/ou controlado, multidisciplinar e/ou especializada, texto completo e/ou referencial entre outras especificidades, que deverão ser avaliadas de acordo com as demandas informacionais da comunidade, no momento da sua aquisição. De acordo com Macias-Chapula (1998), além da disponibilização e divulgação da comunicação científica, as bases de dados são utilizadas, por exemplo, como fontes de informações para avaliar a atuação de pesquisadores, instituições e países, através dos estudos métricos da ciência, em especial a bibliometria.

Para Vanz (2009), os indicadores bibliométricos, quando utilizados na avaliação da ciência constituem-se em uma das mais importantes ferramentas para auxiliar a análise da produção científica, seja de uma determinada instituição, cidade ou até mesmo de um país. Esses indicadores auxiliam os órgãos do governo e as agências de fomento a tomarem melhores decisões em relação ao planejamento e inferem diretamente na produção intelectual e nos avanços científicos proporcionados à sociedade, através dos resultados das pesquisas (VANZ, 2009).

A base de dados *Web of Science (WOS)* (BRASIL, 2019a) está entre as plataformas mais influentes no contexto da ciência, se não, a mais influente por indexar parte significativa da produção científica internacional, sob um rigoroso processo de avaliação. Essa base disponibiliza dados de aproximadamente 9.000 periódicos distribuídos em 5 coleções que estão disponíveis para consulta por meio do pagamento de uma assinatura. Além do acesso às referências bibliográficas e textos completos, a *WOS* disponibiliza ferramentas para análise bibliométrica do seu conteúdo, como por exemplo, análise de citações e índice h. Devido a essas características é uma das bases de dados mais utilizada como fonte de informações para os estudos métricos da ciência. Dentre os estudos, pode-se citar a elaboração de *rankings* universitários, que têm o objetivo de auxiliar futuros

estudantes na escolha das melhores universidades, como por exemplo, no Brasil o *Ranking* Universitário da Folha RUF¹, que tem em sua metodologia a WOS como fonte de informação.

Outra fonte de informação utilizada de forma significativa no contexto brasileiro é a Plataforma Lattes, que compreende bases de dados sobre grupos de pesquisa e currículos dos pesquisadores que atuam no desenvolvimento científico e tecnológico no Brasil (BRASIL, 2019b). Atualmente essa Plataforma é utilizada como fonte de informação por organizações de governo, agências de fomento, grupos de pesquisadores entre outros, na avaliação do desempenho dos pesquisadores e instituições, visando racionalizar o processo de tomada de decisão a respeito da distribuição e alocação de recursos, e para a elaboração e manutenção de políticas públicas para a ciência brasileira (BRASIL, 2019b).

Porém, a extração de informações da Plataforma Lattes ficava comprometida pelos mecanismos de proteção, que restringem o acesso aos currículos. Matias (2015) desenvolveu uma solução computacional, denominada SynLattes, que avançou na resolução do problema de extração de informações institucionais de forma robusta, ágil e confiável, viabilizando a utilização dessa base como fonte de informações para os estudos métricos.

Durante a elaboração dos estudos métricos, o tratamento da informação pode ser realizado de forma manual ou automática, mas o que realmente importa para a qualidade e legitimidade dos resultados desses estudos é a capacidade de recuperar as informações pertinentes, uma amostra que consiga ser representativa de uma determinada realidade.

Entre os vários desafios que devem ser superados, está a dificuldade de identificação e recuperação da produção científica institucional, como por exemplo, a falta de padronização na identificação da instituição por parte dos seus pesquisadores. As diversas formas de redação do nome de uma Instituição de Ciência e Tecnologia (ICT) comprometem a recuperação da sua produção científica e conseqüentemente a qualidade dos resultados dos estudos métricos institucionais. A base de dados WOS, visando superar esse desafio, disponibiliza na sua interface um mecanismo que possibilita identificar um conjunto de variações de nomes de uma determinada ICT, porém essa solução, apesar de ser utilizada na construção do *Ranking* Universitário da Folha, ainda é insuficiente para resolver toda a problemática (BASSOLI, 2017).

¹ RUF 2018 - <https://ruf.folha.uol.com.br/2018/ranking-de-universidades>

Na ciência da informação a solução para a falta de padronização na redação de nomes de autores e instituições é alcançada por meio do controle de autoridade. Esse controle é uma facilidade importante no processo de catalogação, por isso é muito utilizado nos sistemas informatizados de gerenciamento de acervos bibliográficos, no entanto é pouco implementada em outros sistemas. Cenário que tem mudado com o surgimento de novas ferramentas informatizadas que disponibilizam documentos digitais em sua forma completa, o que requer descrição acurada do documento, tal como ocorre nos repositórios digitais (SEGUNDO et al., 2013).

No contexto dos autores individuais, atualmente vêm sendo desenvolvidas soluções para os problemas com a falta de padronização de nomes, como por exemplo, o *Digital Object Identifier* (DOI), que possibilita a correta identificação dos autores de um determinado documento digital (INTERNATIONAL DOI FOUNDATION, 2018). No Brasil, há um contrato de certificação digital para a Plataforma Lattes que, segundo a Assessoria de Comunicação do CNPq, garante maior confiabilidade à informação cadastrada e acesso integral ao artigo publicado pelos pesquisadores, que é identificado através do DOI. Porém, não são capazes de diferenciar as diversas formas de redação do nome de uma ICT, mas é possível intuir que podem ser utilizados para a sua identificação, por intermédio dos identificadores únicos dos autores registrados na Plataforma Lattes.

Com o intuito de contribuir para a eficácia e legitimidade dos resultados dos estudos métricos da ciência ao nível institucional, o objetivo geral desse trabalho foi elaborar um tesouro² de forma automatizada, para o controle de autoridade aplicado às ICT, o qual compreendeu os seguintes objetivos específicos:

- a) Identificar e recuperar a produção científica da unidade-caso na Plataforma Lattes;
- b) Identificar e recuperar a produção científica da unidade-caso na base de dados *Web of Science*;
- c) Elaborar e aplicar um tesouro, de forma automatizada, na análise da produção científica da unidade-caso;

² O termo tesouro é adotado pelo *software* Vantage Point para denominar o recurso que permite a aplicação de dicionários de sinônimos em uma lista (SEARCH TECHNOLOGY, 2009) e difere-se do termo utilizado para denominar uma lista estruturada de termos relacionados semântica e genericamente, utilizados por áreas como a Ciência da Informação (CUNHA; CAVALCANTI, 2008).

- d) Analisar os indicadores bibliométricos institucionais, com base na eficácia do tesouro proposto.

2 DESENVOLVIMENTO

2.1 Coleta e análise de dados

Para alcançar o objetivo geral da pesquisa foi realizado um estudo de caso exploratório (MIGUEL, 2007), com abordagem quantitativa, sendo utilizada a bibliometria como técnica de análise de dados (ALAVI; CARLSON, 1992), aplicada com o auxílio do *software* Vantagepoint³.

A Universidade Federal de São Carlos (UFSCar) foi selecionada como unidade-caso, devido à facilidade de acesso às informações a respeito da sua produção científica institucional registrada na Plataforma Lattes. A produção científica da unidade-caso foi obtida por meio de duas fontes de informação, sendo 37.667 registros bibliográficos de artigos de periódicos registrados na Plataforma Lattes, e 17.240 registrados na base de dados WOS. O desenvolvimento desta pesquisa envolveu quatro etapas:

- a) Etapa 1: coleta da produção científica da unidade-caso na Plataforma Lattes;
- b) Etapa 2: coleta da produção científica da unidade-caso na WOS;
- c) Etapa 3: identificação de registros em comum entre Plataforma Lattes e WOS por meio da identificação do DOI e de um novo campo-chave, denominado ISSN_ANO_VOLUME_PÁGINA_INICIAL; e
- d) Etapa 4: Elaboração e aplicação de um tesouro da unidade-caso, utilizado para a confecção de indicadores bibliométricos institucionais, a partir do controle de autoridade institucional.

A Etapa 1 envolveu a coleta da produção científica da UFSCar (unidade-caso) por meio da aplicação da ferramenta Synclattes (MATIAS, 2015) junto à Plataforma Lattes. Como resultado foi possível coletar a produção científica, de 1.287 pesquisadores docentes que atuaram na UFSCar e compreendeu 37.667 registros bibliográficos. Os dados extraídos foram importados para o *software* Vantagepoint.

Os registros obtidos a partir da Plataforma Lattes foram utilizados como referência da produção científica da UFSCar e, uma vez que a Plataforma Lattes não apresenta metadados

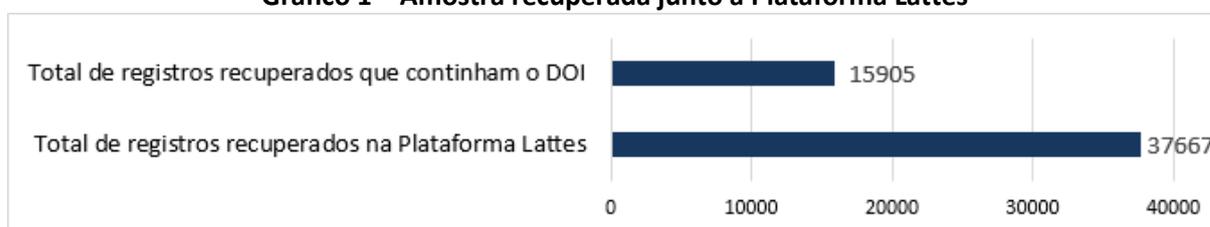
³<https://www.thevantagepoint.com>

sobre o vínculo institucional, optou-se por identificar os mesmos registros em bases de dados de informação científica e tecnológica que estavam indexados. Deste modo, a base de dados *WOS* foi selecionada para a comparação de registros e identificação de variações do nome institucional.

Para a identificação de registros entre a Plataforma Lattes e a base de dados foi necessária a coleta da produção científica da UFSCar na *WOS* (Etapa 2) e posterior comparação e fusão desses registros, permitindo a correlação e análise de metadados provenientes de ambas as fontes de informação. A identificação e coleta da produção científica da UFSCar indexada na *WOS* foi realizada a partir da expressão de busca utilizada por Bassoli (2017). Foram recuperados 17.240 registros da produção científica da UFSCar indexados na base de dados.

Para a identificação de registros em comum entre as fontes de informação, Etapa 3⁴, foi utilizado inicialmente o *Digital Object Identifier* (DOI). Na identificação do DOI da produção científica coletada junto a Plataforma Lattes, do total de registros recuperados (37.667), 42,2% continham DOI (15.905) (Gráfico 1).

Gráfico 1 – Amostra recuperada junto a Plataforma Lattes

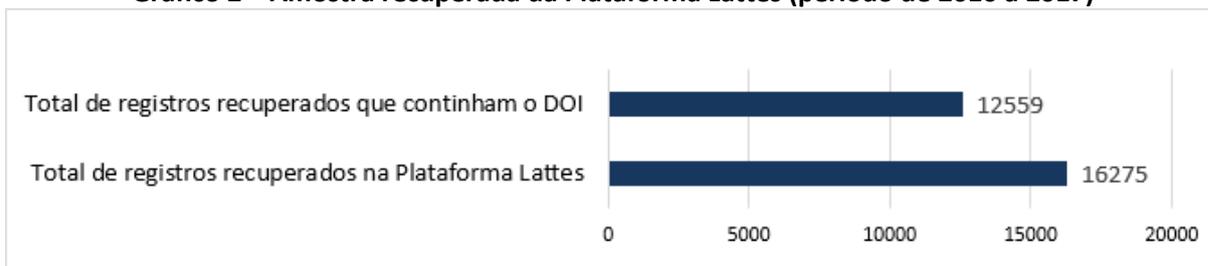


Fonte: elaborado pelos autores.

Levando em conta o período de 2010 a 2017, evidenciou-se uma maior quantidade de documentos que continham o DOI, deste modo optou-se em realizar a análise da produção científica da UFSCar neste período, contabilizando uma amostragem de 16.275 registros, dos quais 77,2% continham DOI (12.559) (Gráfico 2).

⁴ O passo-a-passo para unir metadados de bases de dados diferentes no VantagePoint está disponível em: <http://www.nit.ufscar.br/ferramentas/uniao-de-bases-e-tesauro-ufscar>

Gráfico 2 – Amostra recuperada da Plataforma Lattes (período de 2010 a 2017)



Fonte: elaborado pelos autores.

Nessa etapa, foi necessário o tratamento dos metadados coletados, especialmente o texto do campo DOI, pois notou-se diferenças na formatação entre os metadados do Lattes e da WOS. Após o tratamento, as amostras puderam ser cotejadas com o auxílio do *software* Vantagepoint, visando a identificação do conjunto de registros em intersecção nas duas amostras. O resultado desta etapa foi a identificação de 5.956 registros com informação do DOI indexados na Plataforma Lattes e na WOS, no período de 2010 a 2017.

Ainda, optou-se por elaborar um novo campo-chave, além do DOI, visando uma maior abrangência no que diz respeito à comparação de registros entre a Plataforma Lattes e WOS. Nesse sentido, criou-se um campo-chave único, formado por outros quatro campos da publicação: ISSN, ano, volume e página inicial. O processo de mesclagem dos campos, que objetivou formar uma nova *string* de caracteres, se deu através do *software* Vantagepoint e foi realizado nos registros de ambas as bases.

De forma semelhante à comparação de registros utilizando o DOI, de posse do novo campo-chave, as amostras puderam ser cotejadas com o auxílio do *software* Vantagepoint, visando a identificação do conjunto de registros em intersecção nas duas amostras. O resultado desta etapa foi a identificação de 3.987 registros com informação do ISSN_ANO_VOL_PAG_INICIAL, indexados na Plataforma Lattes e na WOS. Em seguida, verificou-se os registros que ficaram de fora da intersecção das amostras, obtendo-se desta forma a identificação de mais 894 registros exclusivos, conforme ilustra a Figura 1.

Figura 1 - Recuperação de registros das bases por intersecção



Fonte: elaborado pelos autores.

Portanto, a amostra final resultou de 5.956 identificados a partir do DOI, somados a 894 registros identificados com o uso no novo campo-chave, totalizando 6.850 registros do montante das duas fontes de informação.

2.2 Construção de Vocabulário Controlado

Por fim, na Etapa 4, ocorreu a elaboração e aplicação de um tesauro da unidade-caso, a partir das nomenclaturas identificadas nos 5.956 registros bibliográficos referentes a produção científica institucional indexada na Plataforma Lattes e na WOS. Foi possível identificar 151 variações utilizadas pelos pesquisadores da UFSCar na amostra analisada.

As produções científicas recuperadas na Plataforma Lattes e na WOS foram importadas para o *software* Vantagepoint e, por meio dos campos-chave DOI e ISSN_ANO_VOL_PAG_INICIAL, foram identificados os registros em comum entre as fontes de informação, removidas as duplicidades e excluídos os registros bibliográficos que não continham campos-chave. A partir da identificação da diversidade de nomenclaturas utilizadas pelos docentes da UFSCar foi elaborado um controle de autoridade, através da aplicação do tesauro (Apêndices A e B), que possibilitou a elaboração de indicadores bibliométricos institucionais da unidade caso.

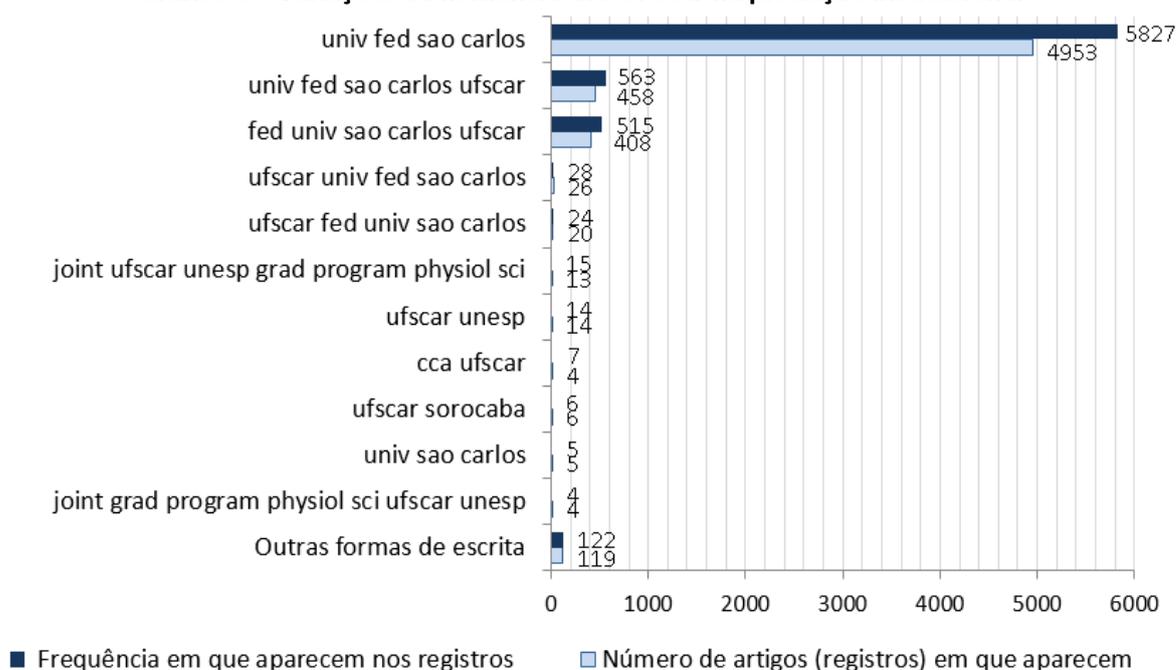
2.3 Resultados

2.3.1 Indicadores sobre as variações de nomenclatura da UFSCar

Foi possível a partir do *Digital Object Identifier System* (DOI) e do novo campo-chave ISSN_ANO_VOL_PAG_INICIAL, identificados na produção científica registrada na Plataforma Lattes e cotejados na Web of Science, identificar de forma automatizada as variações de redação do nome da unidade caso, e ainda, gerar um tesouro para um efetivo controle de autoridade (Apêndice A). Também, foi possível gerar um conjunto de indicadores institucionais sobre a produção científica, a partir da aplicação do tesouro da unidade caso e analisar a sua eficácia.

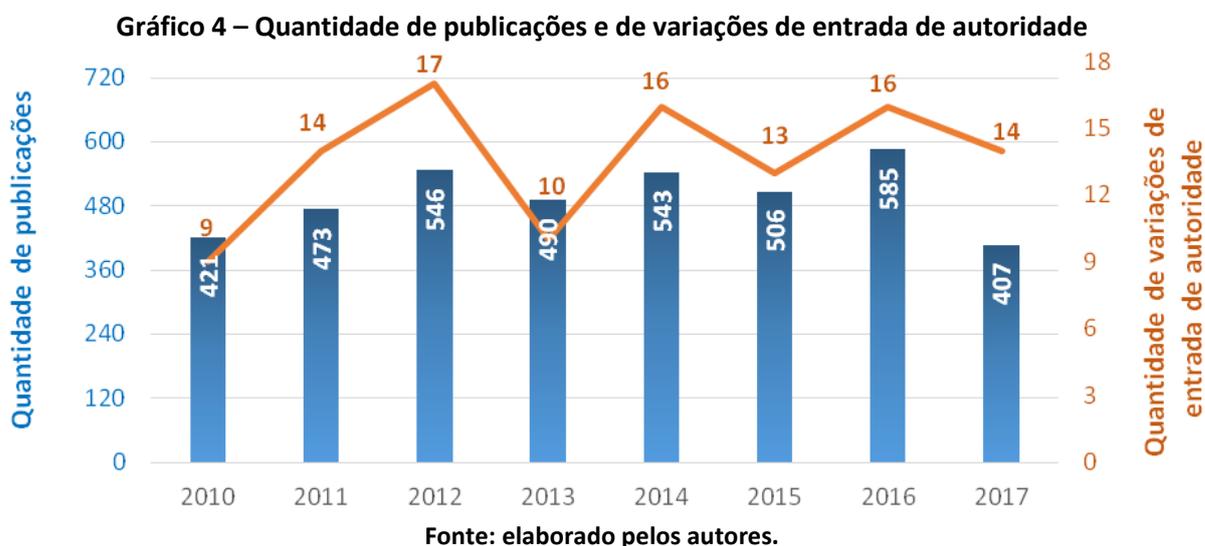
Após a análise e criação do tesouro da unidade caso, chegou-se às variações de nomenclatura utilizadas pelos 1.287 pesquisadores da UFSCar em suas publicações. A partir da aplicação deste tesouro foi elaborado um *ranking* (Figura 2) com base em duas dimensões, sendo uma com o número de artigos (registros) em que aparecem e a frequência em que aparecem nos registros, respectivamente. A amostra analisada compreendeu 5.956 registros, sendo possível identificar as variações de nomes em 5.947 desses. O Gráfico 3 ilustra o *ranking* elaborado com as variações da nomenclatura da UFSCar.

Gráfico 3 – Variações da nomenclatura UFSCar na produção institucional



Fonte: elaborado pelos autores.

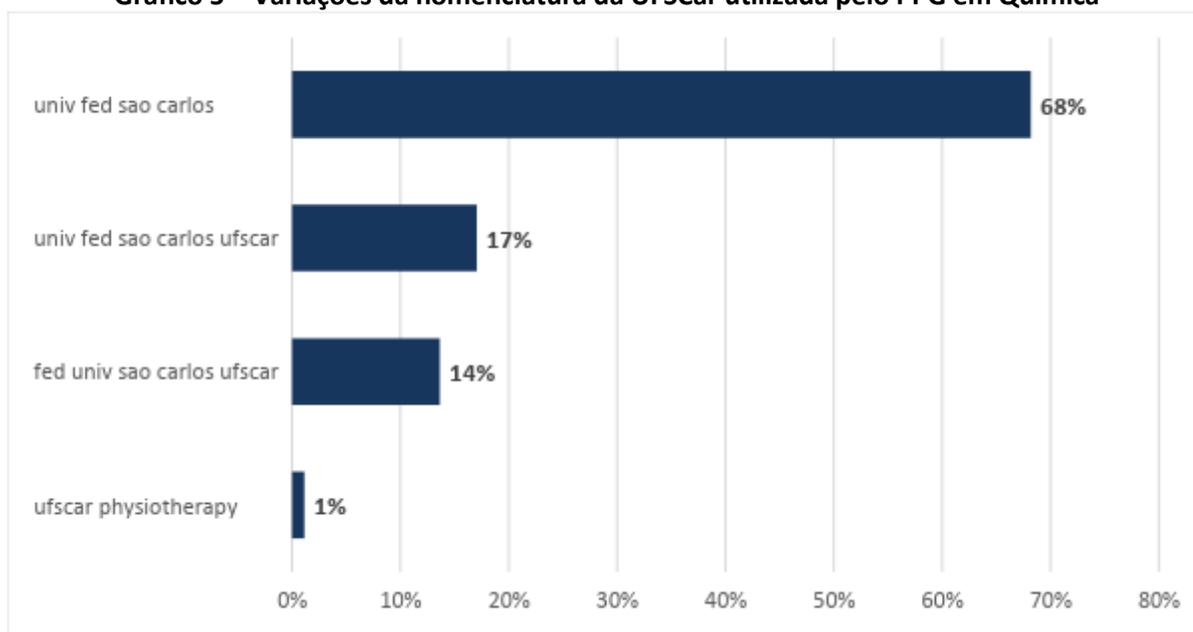
Verificou-se também a quantidade de publicações e de variações de entrada de autoridade (Gráfico 4), visto que este resultado mostrou uma grande variação de entradas ao longo dos anos. Essa variação pode estar relacionada com a forma como os periódicos científicos e bases de dados indexam/cadastram as instituições em suas bases.



2.3.2 Identificação das variações de nomenclatura da UFSCar utilizadas por unidade organizacional: PPG Química

A estratégia utilizada possibilitou a identificação das unidades organizacionais da unidade caso e as variações na nomenclatura por elas utilizadas na amostra de 5.956 registros. Por exemplo, foi possível identificar as variações utilizadas pelo PPG em Química, conforme ilustrado no Gráfico 5 e no Apêndice B. A prática utilizada pelos pesquisadores dessa unidade na identificação da UFSCar compreende o uso de nomenclaturas referentes, por exemplo, ao departamento, ao grupo de pesquisa, ou ainda, a uma temática específica. Esta prática pode estar relacionada ao esforço dos pesquisadores em dar visibilidade às iniciativas da sua comunidade. Porém, prejudicam a visibilidade institucional, ao contribuir com a falta de padronização e consequentemente para maximizar a dificuldade de se identificar e recuperar a produção científica institucional. Este indicador, aplicado a essa e a outras unidades organizacionais da unidade caso, poderá nortear ações de melhorias específicas, possibilitando aos gestores identificarem e estabelecerem prioridades nas ações de melhoria, visando o controle de autoridade institucional e a conscientização dos pesquisadores.

Gráfico 5 – Variações da nomenclatura da UFSCar utilizada pelo PPG em Química

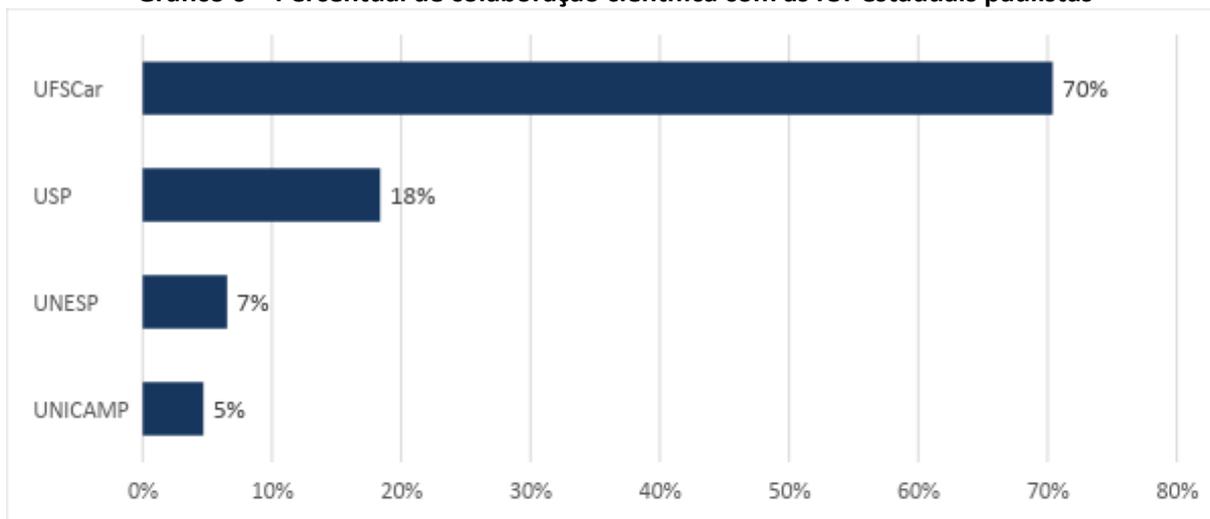


Fonte: elaborado pelos autores.

2.3.3 Identificação das variações das instituições parceiras

Uma vez que a Universidade Federal de São Carlos apresentou na amostra analisada de 5.956 registros bibliográficos, três Instituições de Ciência e Tecnologia (ICT) com significativa participação em sua produção institucional, identificadas através da análise de coautorias de artigos em colaboração, sendo elas: Universidade de São Paulo (USP), Universidade Estadual Paulista (UNESP) e a Universidade Estadual de Campinas (UNICAMP), considerou-se relevante o levantamento das variações de suas nomenclaturas. O indicador apresentado no Gráfico 6 ilustra o percentual de colaboração com essas ICT. Este indicador poderá no futuro auxiliar os pesquisadores na busca por parcerias estratégicas para o desenvolvimento de suas iniciativas em pesquisa, pois além de possibilitar a identificação das ICT, também é possível identificar os assuntos em que as colaborações foram realizadas, uma vez que essas informações estão registradas nos metadados da produção científica registrada na WOS.

Gráfico 6 – Percentual de colaboração científica com as ICT estaduais paulistas



Fonte: elaborado pelos autores.

3 CONSIDERAÇÕES FINAIS

Os resultados alcançados poderão contribuir para aumentar a visibilidade da UFSCar nos *rankings*, ao facilitar a sua identificação e padronização nas bases de dados analisadas, através da construção e uso de um tesouro automatizado. Essa estratégia maximiza a eficácia da recuperação da produção científica institucional nas mais diversas bases de dados, em especial na *Web of Science*. E ainda, viabiliza a implementação de ações de melhoria no processo de comunicação e divulgação científica, contribuindo para um melhor posicionamento nos *rankings* e reconhecimento pela comunidade científica.

Como observado através da identificação das principais ICT parceiras da unidade caso: USP, UNESP e UNICAMP, a estratégia aqui proposta poderá ser replicada para outras instituições, contribuindo para a legitimidade dos resultados dos estudos métricos.

Como limitação, verificou-se o uso recente do DOI pela Plataforma Lattes, o que poderá prejudicar estudos que visam analisar períodos anteriores a 2010. A ampliação da estratégia para a identificação da produção indexada em outras bases poderia melhorar os resultados. Por outro lado, a identificação das variações de nomenclatura através de outras estratégias além do uso do DOI, como por exemplo, a utilização do novo campo-chave ISSN_ANO_VOL_PAG_INICIAL em conjunto com o DOI, melhorou o retorno de registros. Contudo, ainda há a necessidade de divulgação e uso da padronização pela instituição, por periódicos científicos e por bases de dados.

Conclui-se que os resultados alcançados poderão contribuir para resultados mais legítimos dos estudos métricos, e ainda, para a visibilidade da ICT, ao potencializar a recuperação da produção científica institucional.

REFERÊNCIAS

- ALAVI, M.; CARLSON, P. A review of MIS research and disciplinary development. **Journal of Management Information Systems**, v. 8, n. 4, p. 45-62, 1992. Disponível em: <<https://doi.org/10.1080/07421222.1992.11517938>>. Acesso em: 12 jul. 2019.
- BASSOLI, M. **Avaliação do Currículo Lattes como fonte de informação para construção de indicadores**: o caso da UFSCar. 2017. 128 f. Dissertação (Mestrado em Ciência, Tecnologia e Sociedade) - Centro de Ciências Humanas, Universidade Federal de São Carlos, São Carlos, 2017. Disponível em: <<https://repositorio.ufscar.br/handle/ufscar/8908>>. Acesso em: 12 jul. 2019.
- BRASIL. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. **Portal de Periódicos CAPES**. Brasília, DF: CAPES, 2019a. Disponível em: <http://buscador-periodicos-capes-gov-br.ez31.periodicos.capes.gov.br/V/6HJ7U95RVC6I6IPTBPHB72BU86AY342K3D5YA8U7LH9TD RP1SJ-02795?func=find-db-info&doc_num=000002653>. Acesso em: 05 mar. 2017.
- BRASIL. Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). **Plataforma Lattes**. Brasília, DF: CNPq, 2019b. Disponível em: <<http://lattes.cnpq.br>>. Acesso em: 17 mar. 2019.
- CUNHA, M. B.; CAVALCANTI, C. R. O. **Dicionário de Biblioteconomia e Arquivologia**. Brasília: Briquet de Lemos, 2008.
- INTERNATIONAL DOI FOUNDATION. **Digital Object Identifier (DOI)**. 2018. Disponível em: <<http://www.doi.org>>. Acesso em: 03 mar. 2019.
- LANCASTER, F. W. **Toward paperless information systems**. New York, NY: Academic Press, 1978.
- MACIAS-CHAPULA, C. A. O papel da informetria e da cienciometria e sua perspectiva nacional e internacional. **Ciência da Informação**, v. 27, n. 2, p. 134-40, mai./ago. 1998. Disponível em: <<http://dx.doi.org/10.1590/S0100-19651998000200005>>. Acesso em: 12 jul. 2019.
- MATIAS, M. S. de O. **Base referencial para o povoamento de repositórios institucionais**: coleta automatizada de metadados da plataforma lattes. 2015. 94 f. Dissertação (Mestrado Profissional em Gestão de Organizações e Sistemas Públicos) - Universidade Federal de São Carlos, São Carlos: UFSCar, 2015. Disponível em: <<https://repositorio.ufscar.br/handle/ufscar/6932>>. Acesso em: 12 jul. 2019.

XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC

MIGUEL, P. A. C. Estudo de caso na engenharia de produção: estruturação e recomendações para sua condução. **Produção**, v. 17, n. 1, p. 216-229, jan./abr. 2007. Disponível em: <<http://dx.doi.org/10.1590/S0103-65132007000100015>>. Acesso em: 12 jul. 2019.

SEARCH TECHNOLOGY. **Vantage Point**: user guide. Norcross, GA, 2015. Disponível em: <https://www.thevantagepoint.com/_Analyst_Guide_Online_/Users_Guide/VP%20Users%20Guide.pdf>. Acesso em: 24 set. 2019.

SEGUNDO, W. R. et al. A implementação de controle de autoridade em sistemas baseados em DSpace. In: CINFORM - ENCONTRO NACIONAL DE ENSINO E PESQUISA EM INFORMAÇÃO, 11., 2013, Salvador. **Anais...** Salvador: Escola Politécnica/UFBa, 2013. Disponível em: <<http://repositorio.ufes.br/bitstream/10/2025/1/artigo%20cinform%20pelo%20modelo.pdf>>. Acesso em: 12 jul. 2019.

VANZ, S. A. S. **As redes de colaboração científica no Brasil (2004-2006)**. 2009. 204 f. Tese (Doutorado em Comunicação e Informação) - Faculdade de Biblioteconomia e Comunicação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2009. Disponível em: <<http://hdl.handle.net/10183/17169>>. Acesso em: 12 jul. 2019.

APÊNDICE A – Tesouro para o controle de autoridade da UFSCar⁵

**#ufscar - universidade federal de saocarlos	100 1 ^joint ufscarunesp\$
100 1 ^agroecol desenvolvimento rural	100 1 ^lab biogeoquimambientaldqufscar\$
ufscarpesquis\$	100 1 ^lab rmnufscar\$
100 1 ^cacsorfedunivsaocarlos\$	100 1 ^liecdquniv fed saocarlos\$
100 1 ^cacsorfedunivsaocarlos\$	100 1 ^liecufscar\$
100 1 ^ccaufscar\$	100 1 ^liecuniv fed saocarlos\$
100 1 ^ccbsufscar\$	100 1 ^matemat&educufscar\$
100 1 ^ccdmufscar\$	100 1 ^nanofaelliecdqufscar\$
100 1 ^ccetfedunivsaocarlos\$	100 1 ^physinstsaocarlos\$
100 1 ^cctspgdbucufscar\$	100 1 ^pipgcfufscarunesp\$
100 1 ^cdmfufscarunivfedsaocarlos\$	100 1 ^ppgcemufscar\$
100 1 ^cechufscar\$	100 1 ^ppgernufscar\$
100 1 ^cmdmc\$	100 1 ^ppgcemufscar\$
100 1 ^cmdmclieufscar\$	100 1 ^ppgcemunivfedsaocarlos\$
100 1 ^demaufscar\$	100 1 ^programa posgradciencia&engnmatppg
100 1 ^deptbotufscar\$	cem ufscas\$
100 1 ^dept enfermagem ufscar\$	100 1 ^psychobiolgrpufscar\$
100 1 ^deptengn mat\$	100 1 ^san carlos fed univ\$
100 1 ^deptengn mat ufscar\$	100 1 ^sancarlosuniv\$
100 1 ^deptfisufscar\$	100 1 ^saocarlosfeduniv\$
100 1 ^depthidrobiolufscar\$	100 1 ^saocarlosfedunivufscar\$
100 1 ^dept mat engn\$	100 1 ^saocarlosuniv\$
100 1 ^dept mat engnufscar rod\$	100 1 ^joint ufscarunesp grad program
100 1 ^deptmatematufscar\$	physiolsci\$
100 1 ^deptpsicolufscar\$	100 1 ^ufscar\$
100 1 ^dept quim ufscar\$	100 1 ^ufscar\$
100 1 ^dequfscar\$	100 1 ^ufscarcca\$
100 1 ^des ufscar\$	100 1 ^ufscarcmdmcliec\$
100 1 ^desufscar\$	100 1 ^ufscarctrcienciasagr\$
100 1 ^dfufscar\$	100 1 ^ufscardema\$
100 1 ^dfufscaruniv fed saocarlos\$	100 1 ^ufscardes\$
100 1 ^dipartimentofisdqufscar\$	100 1 ^ufscarfedunivsaocarlos\$
100 1 ^dipartimentofis&dqufscar\$	100 1 ^ufscarfedunivsaocarlos\$
100 1 ^doutor ciencias sociais ufscar\$	100 1 ^ufscarphysiotherapy\$
100 1 ^doutora engnproducaoufscar\$	100 1 ^ufscarppg diversidade biolconservacao\$
100 1 ^dqufscar\$	100 1 ^ufscar programa posgradgenetevolucao\$
100 1 ^dqufscarunivfedsaocarlos\$	100 1 ^ufscarrod\$
100 1 ^drnpaufscar\$	100 1 ^ufscarrodwashingtonluis\$
100 1 ^fdunivsaocarlos\$	100 1 ^ufscar rodovia washingtonluis\$
100 1 ^fdunivfedsaocarlos\$	100 1 ^ufscarsorocaba\$
100 1 ^feduniv s ao carlosufscar\$	100 1 ^ufscarsp\$
100 1 ^feduniv s carlos\$	100 1 ^ufscarsp brazil\$
100 1 ^fedunivsaocarlosufscar\$	100 1 ^ufscarunesp\$
100 1 ^fedunivsaocarlos\$	100 1 ^ufscarunespararaquara\$
100 1 ^fedunivsaocarlosdcufscar\$	100 1 ^ufscarunesprod\$
100 1 ^fedunivsaocarlosdepufscar\$	100 1 ^ufscarunespsaocarlos\$
100 1 ^fedunivsaocarlosdepsufscar\$	100 1 ^ufscarunivfedsaocarlos\$

⁵Os resultados desta pesquisa foram utilizados para atualizar o tesouro utilizado pelo NIT/Materiais (<http://www.nit.ufscar.br/ferramentas/uniao-de-bases-e-tesouro-ufscar>).

XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC

100 1 ^fedunivsaocarlosppequfscar\$	100 1 ^univfed s carlos\$
100 1 ^fedunivsaocarlosppgequfscar\$	100 1 ^univfedsacarlos\$
100 1 ^fedunivsaocarlosppgequfscar\$	100 1 ^univfedsancarlos\$
100 1 ^fedunivsaocarlossufscar\$	100 1 ^univfedsaocarlos\$
100 1 ^fedunivsaocarlosufscar\$	100 1 ^univfedsaocarlosccaufscar\$
100 1 ^fedunivsaocarlosufscar\$	100 1 ^univfedsaocarlosccbsufscar\$
100 1 ^fedunivsaocarlosufscar rodovia\$	100 1 ^univfedsaocarloscrrufscarsorocaba\$
100 1 ^fedunivsaocarols\$	100 1 ^univfedsaocarlosdbeeufscar\$
100 1 ^fedunivsaocarlosufscar\$	100 1 ^univfedsaocarlosdcufscar\$
100 1 ^fedunivsaocarlosufscar\$	100 1 ^univfedsaocarlosdemaufscar\$
100 1 ^fedunivselo carlosufscar\$	100 1 ^univfedsaocarlosdema\$
100 1 ^feduniv silo carlosufscar\$	100 1 ^univfedsaocarlosdgtufscar\$
100 1 ^feduniv soo carlosufscar\$	100 1 ^univfedsaocarlos dm\$
100 1 ^fedunivufscar\$	100 1 ^univfedsaocarlosppgern\$
100 1 ^federal univsaocarlos\$	100 1 ^univfedsaocarlos rodovia
100 1 ^fisiolunivfedsaocarlos\$	washingtonluis\$
100 1 ^fundacaounivfedsaocarlos\$	100 1 ^univfedsaocarlossp\$
100 1 ^grad program mat sci&engnppgem\$	100 1 ^univfedsaocarlosufscar\$
100 1 ^grad program psycholufscar\$	100 1 ^univfedsaocarlosufscar\$
100 1 ^grad schbiotechnol&environm	100 1 ^univfedsaocarlosufscarsp\$
monitoring ufscar\$	100 1 ^univfedsarcarlosufscar\$
100 1 ^grad schevolut genet &molbiolufscar\$	100 1 ^univfedsclocarlosufscar\$
100 1 ^inctmndqunivfedsaocarlos\$	100 1 ^univfed silo carlos\$
100 1 ^inctmndfedunivsaocarlos\$	100 1 ^univfedsocarlos\$
100 1 ^instfissaocarlos\$	100 1 ^univfedsocarlosufscar\$
100 1 ^instquimsaocarlos\$	100 1 ^univfed soa carlos\$
100 1 ^joint fedunivsaocarlosufscar\$	100 1 ^univsancarlos\$
100 1 ^joint grad program	100 1 ^univsaocarlos\$
physiolscipipgcfufscarunes\$	100 1 ^univsaocarlosufscar\$
100 1 ^joint grad program	100 1 ^univsaocarlus\$
physiolsciufscarunesp\$	100 1 ^universidade federal de são carlos \$
100 1 ^joint grad program psycholufscar\$	100 1 ^universidade federal de são carlos\$
	100 1 ^vitreousmatlabufscar\$

Fonte: elaborado pelos autores.

APÊNDICE B – Nomenclaturas usadas pelo Departamento de Química da UFSCar.

Número de registros > Nomenclatura	1 dqufscar
1381 univfedsaocarlos	1 fedunivsaocarlosufscar rodovia
92 univfedsaocarlosufscar	1 feduniv selo carlosufscar
69 fedunivsaocarlosufscar	1 feduniv soo carlosufscar
11 ufscarfedunivsaocarlos	1 inctmndqunivfedsaocarlos
9 ufscarunivfedsaocarlos	1 inctmnedunivsaocarlos
3 dqufscarunivfedsaocarlos	1 instfissaocarlos
3 liecunivfedsaocarlos	1 instquimsaocarlos
2 cdmfufscarunivfedsaocarlos	1 liecdqunivfedsaocarlos
2 fedunivsaocarlosspufscar	1 liecufscar
2 ppgcemufscar	1 ppg cem ufscar
2 sancarlosfeduniv	1 ppgcemunivfedsaocarlos
2 saocarlosfedunivufscar	1 saocarlosfedunivufscarsp
2 ufscarunesp	1 ufscardema
2 univfedsaocarlosufscarsp	1 univfed s carlos
1 ccdmufscar	1 univfedsaocarlosccbsufscar
1 deptfisufscar	1 univfedsaocarlosdema
1 deptquimufscar	1 univfedsarcarlosufscar
1 dfufscarunivfedsaocarlos	1 univfedsclocarlosufscar
1 dipartimentofis&dqufscar	1 univfed silo carlos

Fonte: elaborado pelos autores.