

Um processo para a identificação e análises de artigos científicos em anais de congressos

A process for the identification and analysis of scientific articles in conference proceedings

Fernanda Silva Coimbra

CEFET-MG, Brasil.

E-mail: coimbra.sfernanda@gmail.com

ORCID: <https://orcid.org/0000-0002-9231-7739>

Thiago Magela Rodrigues Dias

CEFET-MG, Brasil.

E-mail: thiagomagela@cefetmg.br

ORCID: <https://orcid.org/0000-0001-5057-9936>

RESUMO

Os eventos científicos cumprem várias funções no contexto das ciências, é um excelente meio de comunicação entre os cientistas, dão visibilidade interna e externa aos trabalhos apresentados. A produção científica vem crescendo de forma significativa, e, conseqüentemente compreender o comportamento desta tendência se torna importante. Os trabalhos apresentados em eventos, geralmente publicados em anais, visam apresentar o que foi veiculado e debatido durante a realização do evento. Diante do exposto o presente estudo objetiva-se em apresentar características da produção científica brasileira publicadas em anais de eventos, utilizando como fonte de dados os currículos cadastrados na Plataforma Lattes. Com esta análise foi possível identificar e discutir algumas características como as grandes áreas que possuem maior representatividade no conjunto analisado; uma análise temporal permitindo verificar os anos que houve maior número de publicações, e ainda, foi possível verificar se é utilizado algum identificador persistente nos registros destas publicações.

Palavras-chave: produção científica; eventos; anais; plataforma lattes.

Como citar: Coimbra, F. S.; & Dias, T. M. R. (2022). Um processo para a identificação e análises de artigos científicos em anais de congressos. En T. M. R. Dias (Ed.), *Informação, Dados e Tecnologia. Advanced Notes in Information Science, volume 2* (pp. 74-81). Tallinn, Estonia: ColNes Publishing. DOI: 10.47909/anis.978-9916-9760-3-6.93.

Copyright: © 2022, The author(s). This is an open access work distributed under the terms of the CC BY-NC 4.0 license which permits copying and redistributing the material in any medium or format, adapting, transforming and building upon the material as long as the license terms are followed.

ABSTRACT

Scientific events fulfil several functions in the context of science. It is an excellent means of communication between scientists, giving internal and external visibility to the works presented. Scientific production has been growing significantly, and, consequently, understanding the behavior of this trend has become essential. The results presented at events, usually published in proceedings, aim to show what was broadcast and debated during the event. Given the above, this study seeks to give characteristics of Brazilian scientific production published in event proceedings, using the curricula registered in the Lattes Platform as a data source. From our analysis, we identified and discussed the large areas that have more excellent representation in the analyzed set; a temporal analysis is allowed us to verify the years with the most significant number of publications. It was possible to confirm if any persistent identifier was used in the records of these publications.

Keywords: *scientific production; events; proceedings; lattes platform*

1. INTRODUÇÃO

A INTERNET se tornou um grande repositório de conhecimento, propiciando acesso aos usuários de forma simples e intuitiva, permitindo disponibilizar e acessar os trabalhos científicos, que engloba suas análises e seus resultados, assim como, possibilitando registrar suas informações pessoais, profissionais e acadêmicas. Atualmente, existem vários sites para registros dessas informações e vários repositórios institucionais.

O volume de dados disponíveis atualmente possui características próprias, padrões únicos, quantidade e diversidade de dados, ocasionando assim uma tarefa complexa a realização de estudos que visam explorar esses dados. Neste contexto, a bibliometria surge com o intuito de quantificar os processos de comunicação escrita, através da utilização de métodos que geram análises estatísticas sobre a produção e disseminação do conhecimento aplicadas a fontes de dados científicos (Araújo, 2006).

A Plataforma Lattes permite incluir e recuperar informações inseridas nos currículos pelos próprios indivíduos, é uma plataforma de acesso aberto e gratuito. Gerando assim, um rico conjunto de dados que viabiliza realizar inúmeras análises, possibilitando visualizar e compreender a evolução da ciência, compreender a trajetória acadêmica e profissional dos pesquisadores que possuem currículos cadastrados na Plataforma Lattes.

No decorrer dos últimos anos é possível verificar a evolução da divulgação científica nas diversas áreas do conhecimento e nesse âmbito tem sido possível verificar, como este aumento tem sido impulsionado pelo aumento de eventos científicos no mundo. Sendo classificado como um dos principais meios de comunicação científica os eventos, apontados

como meio efetivo de comunicação oral do conhecimento, vem ganhando notoriedade (Mello, 1996). Os eventos de cunho científico desempenham funções como: aperfeiçoamento de trabalhos, reflexo do estado da arte e comunicação, conforme apresentado por Campello *et al.* (2007). Nos eventos científicos são gerados documentos com os trabalhos apresentados, popularmente conhecidos como anais. As publicações geradas nos eventos são consideradas por alguns estudos, como as produções acadêmicas mais atuais (Carmona e Pereira, 2018).

Araújo e Alvarenga (2011) ressaltam que entender os comportamentos de publicações científicas de determinado campo de pesquisa, possibilita novos olhares para a compreensão do mesmo possibilitando uma nova vertente na avaliação da ciência. Diante disso, analisar como as publicações em anais de eventos vem sendo realizadas, se apresenta como um importante mecanismo para a compreensão da evolução dos eventos científicos em um contexto geral ou em determinadas áreas do conhecimento. Entretanto, de forma geral as informações relacionadas à produção científica em anais de eventos estão presentes em inúmeros repositórios de dados, dificultando assim a recuperação e análise dos dados, principalmente em uma grande escala. Neste contexto, no Brasil, a Plataforma Lattes se tornou um padrão para registro de dados curriculares da comunidade científica. De acordo com Lane (2010), a Plataforma Lattes é um poderoso repositório de dados científicos brasileiro, que possui dados de alta qualidade e permite acesso aos dados dos indivíduos cadastrados. É de responsabilidade do próprio indivíduo inserir e atualizar os dados.

Diante disso, este trabalho tem como objetivo compreender traços da produção científica brasileira em anais de congresso, utilizando como fonte de dados os artigos cadastrados nos currículos da Plataforma Lattes.

2. PROCEDIMENTOS METODOLÓGICOS

A metodologia utilizada teve como base à análise bibliométrica concomitante com metodologia quantitativa. Foi utilizado o *LattesDataXplorer* desenvolvido por Dias (2016), esse arcabouço permite realizar o processo de extração e seleção dos dados curriculares da Plataforma Lattes, que envolve um conjunto de técnicas e métodos, que possibilita a coleta, seleção, tratamento e análise dos dados, neste trabalho foi utilizado apenas os módulos de coleta e seleção do *LattesDataXplorer*, com o intuito de extrair e selecionar os dados curriculares da Plataforma Lattes.

O arcabouço foi utilizado para realizar a coleta de todos os currículos, cerca de 7 milhões de registros em janeiro de 2021. Para obter uma análise acurada dos artigos publicados em anais de eventos, preferiu-se estabelecer um conjunto de indivíduos através do nível de formação

acadêmica/titulação para indivíduos que possuem o nível de formação doutorado concluído.

Com o intuito de auxiliar as análises dos dados, após a utilização do arcabouço, foi necessário criar métodos para a seleção, tratamento e visualização conforme Figura 1.

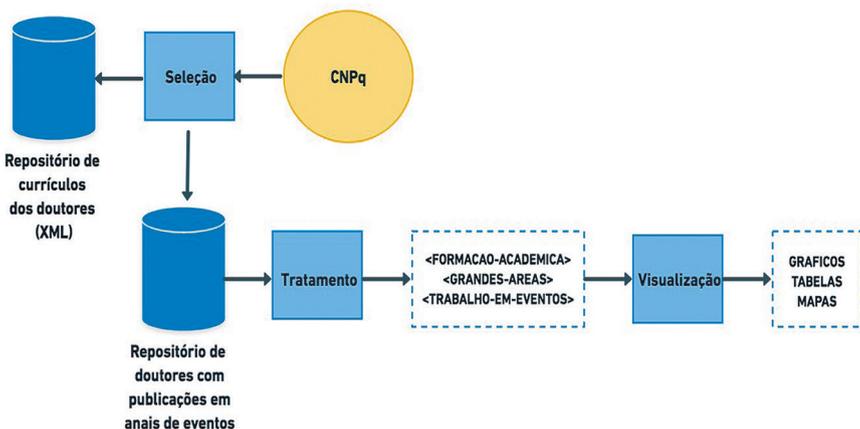


Figura 1. Processo de seleção, tratamento e visualização dos dados

Fonte: Elaborado pelo autor, 2021.

A etapa de tratamento baseou-se em identificar a estrutura do arquivo XML, para cada currículo foi extraído um arquivo XML. Cada currículo é único e possui informações próprias; esses dados estão reunidos por grandes áreas, formação acadêmica, orientações, produções, entre outras.

Após realizar o tratamento do XML, foi realizada a etapa de visualização em que é realizada a caracterização dos dados, permitindo a análise dos dados inseridos na Seção de Trabalhos em Eventos, seção que contém trabalhos publicados em anais de eventos, caso o indivíduo tenha informado em seu currículo.

3. RESULTADOS

A caracterização inicial resultou em um conjunto de 360.888 currículos, este quantitativo representa aproximadamente 5% de currículos cadastrados na Plataforma Lattes (dados de janeiro de 2021). Dentre esses currículos foi verificado que 57.403 currículos não possuíam nenhum artigo informado na seção de trabalhos publicados em anais de eventos, correspondendo a 16% do conjunto de doutores.

Após as análises dos dados foi possível compreender como acontece a produção científica brasileira em anais de congresso, podendo ser

analisados por grandes áreas de atuação e realizar uma análise temporal das publicações.

4. ARTIGOS EM ANAIS DE EVENTOS POR GRANDE ÁREA

A Plataforma Lattes permite inserir as informações de grandes áreas de atuação dos indivíduos seguindo a classificação de áreas do CNPq. Assim, o indivíduo insere a sua classificação em nove grandes áreas, sendo estas: Ciências Agrárias, Ciências Biológicas, Ciências da Saúde, Ciências Exatas e da Terra, Ciências Humanas, Ciências Sociais Aplicadas, Engenharias, Linguística, Letras e Artes e Outra. Caso, o indivíduo não informe este campo, o mesmo estará em branco e no contexto deste trabalho foi caracterizado como “Não informado”.

Com intuito de compreender em quais grandes áreas os indivíduos considerados neste estudo estão distribuídos, a Figura 2 apresenta tal distribuição. Ressalta-se que como é possível inserir mais de uma grande área de atuação em um currículo, quando este fenômeno aconteceu, foi considerado o primeiro registro como sendo a principal grande área.

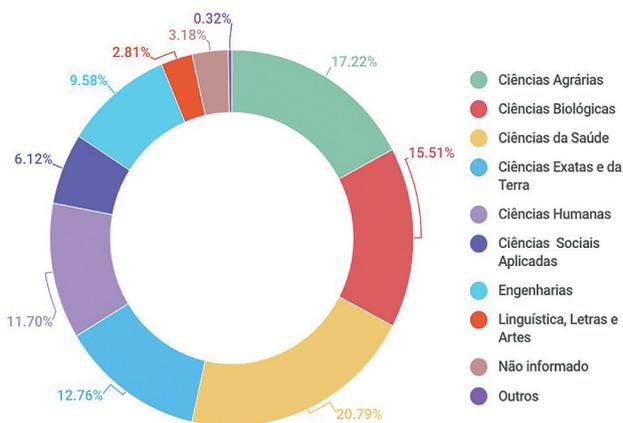


Figura 2. Quantitativo geral dos indivíduos com publicação por grandes áreas.

Fonte: Elaborado pelo autor, (2021).

A Figura 2 demonstrou que a grande área de Ciências da Saúde (20.79%) apresenta maior taxa de publicação em anais de eventos. A grande área de Ciências Agrárias (17.22%) é apresentada com a segunda maior, e com um valor próximo, a grande área de Ciências Biológicas (15.51%).

Conjuntamente, foi possível verificar quais grandes áreas têm menor representatividade em participação em eventos (Figura 3).

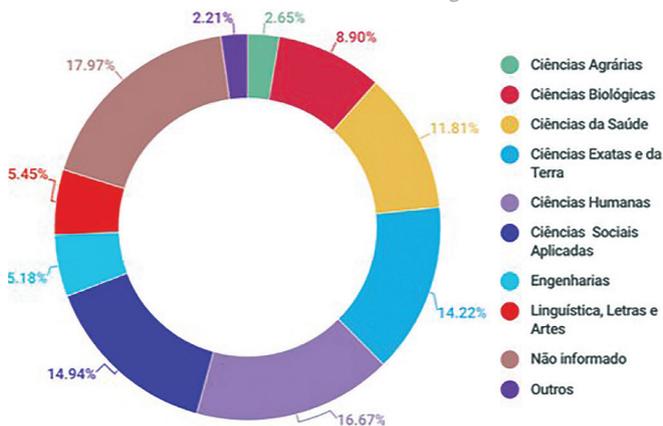


Figura 3. Quantitativo geral dos indivíduos sem publicação por grandes áreas.
Fonte: Elaborado pelo autor, 2021.

Pode-se observar que a grande área que possui maior valor referenciado sem publicações corresponde a Ciências Humanas (16.67%), na segunda e na terceira posição Ciências Sociais Aplicadas (14.94%) e Ciências Exatas e da Terra (14.22), respectivamente, com valores próximos.

5. ANÁLISE TEMPORAL

A análise temporal dos artigos publicados em anais de eventos foi realizada com o intuito de verificar a quantidade de artigos publicados por ano, foi traçada uma linha de corte de 30 anos para melhor representação gráfica (Figura 4).

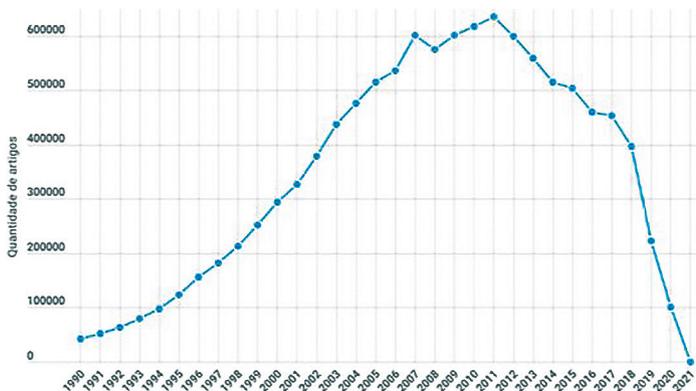


Figura 4. Análise temporal dos artigos em anais de eventos por ano.
Fonte: Elaborado pelo autor, 2021.

É possível verificar assimetrias ao longo do período de análise, pode-se observar uma queda de publicações em eventos a partir de 2011 e acentuada em 2018, em que é apresentado uma queda significativa. Uma hipótese para a queda significativa no quantitativo de publicações nos últimos anos pode estar relacionada a falta de atualização de alguns currículos, que mesmo o autor tendo publicado um trabalho, pode ainda não ter registrado o artigo em seu currículo.

6. IDENTIFICADOR PERSISTENTE

Todos os dados dos trabalhos publicados em eventos são inseridos pelos próprios indivíduos de forma manual, contudo existe uma forma alternativa de inserir tais dados de forma automática, utilizando o identificador persistente DOI, ao digitar o identificador é realizada a consulta e indexado os dados automaticamente no currículo na Plataforma Lattes.

Logo em 2007, a Plataforma Lattes passou a aceitar o DOI, possibilitando a inserção do mesmo, através de digitação manual, permitindo que a Plataforma acesse a base de dados da DOI, preenchendo os dados da publicação de forma automática.

Ao realizar a análise dos dados das publicações cadastradas nos currículos dos doutores que possuem o identificador persistente (DOI), é possível verificar que apenas 30.936 artigos em anais de eventos cadastrados nos currículos do conjunto selecionado possuem identificador persistente, representado assim aproximadamente 3% dos artigos. Demonstrando assim que ainda é um identificador pouco utilizado, tendo em vista que a sua utilização poderia agregar uma maior quantidade de dados referente aos artigos que estão sendo registrados.

7. CONSIDERAÇÕES FINAIS

Com a realização deste trabalho observou-se que os dados extraídos da Plataforma Lattes são uma excelente fonte para compreender como acontece a produção científica brasileira em anais de congresso.

Através das análises realizadas pode-se observar algumas características como: participação em eventos por grande área, sazonalidade das publicações e a escassa utilização de identificador persistente.

CONFLITOS DE INTERESSE

Os autores declaram não haver conflito de interesses.

DECLARAÇÃO DE CONTRIBUIÇÃO

Conceptualização, curadoria de dados, análise formal, investigação, metodologia, visualização, redação - rascunho original: Coimbra, F. S.

Administração do Projeto, Supervisão, Validação, Redação - revisão e edição: Dias, T. M. R.

DECLARAÇÃO DE CONSENTIMENTO DE DADOS

Os repositório de doutores/xml e análises/csv gerados durante o desenvolvimento deste estudo não podem ser disponibilizados gratuitamente devido às restrições impostas por (Fernanda Silva Coimbra) 

REFERÊNCIAS

- ARAÚJO, R. F.; ALVARENGA, L. (2011). A bibliometria na pesquisa científica da pós-graduação brasileira de 1987 a 2007. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação*, 16(13), 51–70. <https://doi.org/10.5007/1518-2924.2011v16n31p51>
- ARAÚJO, C. A. (2006). Bibliometria: evolução histórica e questões atuais. Em *Questão*, 12(1), 11-32. Obtido em <https://seer.ufrgs.br/EmQuestao/article/view/16>
- CAMPELLO, B. S., CENDÓN, B. V., KREMER, J. M. (2007). *Fontes de informação para pesquisadores e profissionais*. Belo Horizonte: Editora UFMG.
- CARMONA, I. V., PEREIRA, M. V. (2018). Ciência, tecnologia e sociedade e educação ambiental: uma revisão bibliográfica em anais de eventos científicos da área de ensino de ciências. *Revista Ciências & Ideias*, 8(3), 94-14. <https://doi.org/10.22407/2176-1477/2017.v8i3.752>
- DIAS, T. M. R. (2016). *Um Estudo da Produção Científica Brasileira a partir de Dados da Plataforma Lattes* (Unpublished doctoral dissertation). Centro Federal de Educação Tecnológica de Minas Gerais, Belo Horizonte.
- LANE, J. (2010). Let's make science metrics more scientific. *Nature*, 464(7288), 488-489. <https://doi.org/10.1038/464488a>
- MELLO, L. L. C. C. (1996). Os anais de encontros científicos como fonte de informação: relato de pesquisa. *Revista de biblioteconomia de Brasília*, 20(1), 53-68. Obtido em <http://www.brapci.inf.br/index.php/article/download/8832>

