

# Utilização de dados abertos para análise da publicação de artigos em eventos científicos

## Use of open data to analyze the publication of articles in scientific events

Fernanda Silva Coimbra<sup>1</sup>, Thiago Magela Rodrigues Dias<sup>2</sup>

**Como citar (APA):** Coimbra, F.S., & Dias, T.M.R. (2021). Utilização de dados abertos para análise da publicação de artigos em eventos científicos. *Iberoamerican Journal of Science Measurement and Communication*; 1(3), 1-13. <https://doi.org/10.47909/ijsmc.123>

**Recebido:** 16-10-2021

**Aceito:** 18-12-2021

**Editor:** Carlos Luis González-Valiente

**Copyright:** © 2021 Coimbra & Dias. This is an open access article 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.

### RESUMO

**Objetivo.** Neste artigo é realizada a análise das publicações de artigos em eventos científicos utilizando dados abertos da Plataforma Lattes.

**Design/Metodologia/Abordagem.** A ferramenta LattesDataXplorer foi utilizado para extrair dados curriculares da Plataforma Lattes, os dados utilizados foram extraídos em janeiro de 2021, para que dessa forma fosse realizado toda a seleção e tratamento dos dados de interesse. A etapa de seleção consistiu-se em verificar quais currículos possuíam trabalhos publicados em eventos científicos; já na etapa de tratamento foi realizada uma caracterização dos artigos, permitindo assim a análise dos artigos publicados em eventos científicos.

**Resultados/Discussão.** Foi possível realizar algumas caracterizações como: análise temporal, análise por grande área de atuação e quais artigos utilizam identificadores persistentes. A análise temporal, possibilitou verificar quantos artigos foram publicados por ano. Através da análise por grande área de atuação, identificou-se os indivíduos que apresenta maior e menor taxa de publicação em anais de eventos. Aproximadamente 3% dos artigos informam o identificador persistente (DOI) em seus artigos publicados em anais de eventos.

**Conclusões.** O ápice das publicações é em 2011, após 2011 nota-se uma queda significativa. Os indivíduos da grande área de Ciências da Saúde apresentam maior taxa de publicação em anais de eventos, já os indivíduos da grande área Ciências Exatas e da Terra apresenta menor volume de publicações. Apenas 30.936 dos artigos em anais de eventos cadastrados nos currículos do conjunto selecionado possuem

1 Centro Federal de Educacao Tecnologica de Minas Gerais, Av. Amazonas, 5253, Belo Horizonte 30421-169, Brazil. Email: [coimbra.sfernanda@gmail.com](mailto:coimbra.sfernanda@gmail.com).

2 Centro Federal de Educacao Tecnologica de Minas Gerais, Divinópolis, Brazil. Email: [thiagomagela@cefetmg.br](mailto:thiagomagela@cefetmg.br), ORCID: <https://orcid.org/0000-0001-5057-9936>.

identificador persistente, porém os indivíduos da grande área de Ciências Exata e da Terra (30,68%) são os que mais utilizam identificador persistente.

**Originalidade/Valor.** A partir de dados extraídos da Plataforma Lattes, são apresentados os resultados de uma pesquisa original que busca caracterizar a produção científica brasileira por meio de eventos.

**Palavras-chave:** Produção Científica; Eventos; Anais; Plataforma Lattes

## ABSTRACT

**Objective.** This article has analyzed the publications of articles in scientific events using open data from the Lattes Platform.

**Design/Methodology/Approach.** The LattesDataXplorer tool was used to extract curricular data from the Lattes Platform. The selection stage consisted of verifying which curricula had works published in scientific events. In the treatment stage, a characterization of the articles was carried out, thus allowing the analysis of articles published in scientific events.

**Results/Discussion.** It was possible to carry out some characterizations such as temporal analysis, analysis by a large area of expertise, and which articles use persistent identifiers. The temporal analysis made it possible to verify how many articles were published per year. Through the analysis by a large area of action, the individuals with the highest and lowest rate of publication in annals of events were identified. Approximately 3% of articles report the persistent identifier (DOI).

**Conclusions.** The peak of publications is in 2011, afterwards, there was a significant drop. Individuals from Health Sciences have a higher rate of publication in event proceedings, while individuals from the Exact and Earth Sciences have a lower volume of publications. Only 30,936 of the articles have a persistent identifier, but individuals from the Exact and Earth Sciences (30.68%) are the ones who use persistent identifiers the most.

**Originality/Value.** From data extracted from the Lattes Platform, the results of original research that seeks to characterize Brazilian scientific production utilizing events are presented.

**Keywords:** Scientific Production; Events; Proceedings; Lattes platform

## INTRODUÇÃO

A internet se tornou um grande repositório de conhecimento científico, propiciando acesso a este valioso bem aos usuários de forma simples e intuitiva, permitindo disponibilizar e acessar os trabalhos científicos, que engloba suas análises e seus resultados, assim como, também possibilitando por parte destes usuários divulgar suas informações pessoais, profissionais e acadêmicas. Atualmente, existem vários sítios para registros dessas informações e vários repositórios institucionais que também viabilizam a divulgação científica por grupos de indivíduos.

São inúmeros os estudos disponibilizados na internet que utilizam os dados das produções científicas e os trabalhos que realizam análises de publicações científicas têm ganhado destaque. As produções científicas são partes integrantes do processo de produção de conhecimento do indivíduo, no qual pode disponibilizar esse conhecimento adquirido através de artigos em anais de congressos e em periódicos, livros, capítulos de livros, resumos, teses, dissertações e monografias, entre outros meios de divulgação e comunicação da ciência (Domingues, 2014).

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).

Atualmente existe várias técnicas quantitativas para avaliar a produtividade científica como a bibliometria, cientometria, informetria e webometria. São técnicas distintas, mas que tem um fator principal em comum, mensurar a propagação do conhecimento científico e o fluxo da informação (Vanti, 2002). Todas as técnicas possibilitam diversas análises. Dias (2016) destaca que, a análise bibliométrica pode ser utilizada para mensurar a produção científica de indivíduos em particular, grupos de pesquisa, instituições, regiões geográficas, organizações ou eventos.

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, Cendón e Kremer (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). Cardoso (2018) expõe que os anais apresentam o que foi veiculado durante realização do evento, podendo ser divulgado em termos de resumo, resumo expandido ou trabalho completo.

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.

No intuito de se obter um grande conjunto de dados sobre publicações científicas em anais de eventos, a Plataforma Lattes do CNPq, surge como uma excelente alternativa para coleta de dados. A Plataforma Lattes possui como um de seus elementos os currículos, em que é possível incluir e recuperar informações inseridas nos currículos pelos próprios indivíduos, se caracteriza como uma plataforma de acesso aberto.

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. Por se tratar de um repositório valioso, é possível encontrar na literatura diversos autores que utilizam a Plataforma Lattes como fonte de estudos bibliométricos.

O trabalho de Dias (2016) é inédito na literatura, visto que além de um estudo amplo sobre a produção científica brasileira utilizando como fonte de dados de todos os currículos cadastrados na Plataforma Lattes, o autor desenvolve um arcabouço, implementando técnicas bibliométricas e métricas baseadas em análises de redes sociais, sendo o mesmo responsável pela extração de todo o conjunto de dados dos currículos coletados em abril de 2015. Após a coleta, o tratamento e a análise dos dados, o autor realiza diversas análises, gerando várias visões e estatísticas. Se trata um trabalho valioso e relevante de caráter inédito sobre produção científica brasileira, tendo em vista a abrangência das análises realizadas e pela grande quantidade de indivíduos analisados.

Inúmeros são os trabalhos que utilizam a Plataforma Lattes como recurso de fonte de dados, por se tratar de um rico repositório inclusive com a temática de artigos publicados em eventos. Porém, observa-se uma quantidade limitada de trabalhos que utilizam como fonte principal de dados as informações curriculares da Plataforma Lattes no contexto dos trabalhos publicados em anais de eventos. Recuperar trabalhos publicados em eventos é uma tarefa complexa, devido a gama de eventos, formatos entre outros pontos. À vista disso, o presente estudo apresenta-se como uma análise no intuito de compreender o comportamento dos trabalhos publicados em anais de eventos registrados nos currículos cadastrados na Plataforma Lattes de forma global, considerando todos os currículos cadastrados, através de análises bibliométricas, caracterização geral dos eventos, análises temporais e análises quantitativas, possibilitando verificar representatividade por áreas do conhecimento.

São diversos os pontos que motivam compreender a produção científica por meio de anais de eventos, um deles é a possibilidade de uma visão geral sobre como as diversas áreas do conhecimento têm explorado este meio de divulgação para apresentar os resultados de suas pesquisas; em quais períodos temporais houve maior número de publicações e quais são as áreas que possuem maior volume de publicações. Diante disso, este trabalho tem como objetivo geral compreender traços da produção científica brasileira em anais de eventos, utilizando como fonte de dados os artigos cadastrados nos currículos da Plataforma Lattes.

### **Trabalhos relacionados**

É possível encontrar diversos trabalhos que analisam eventos específicos como o artigo de Moreira e Moraes (2019) que tem como fonte de dados os anais do Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB), realizados no período de 2003-2014. Os autores analisam os anais do Enancib, que abordam o assunto “classificação” na literatura brasileira de ciência da informação. A análise fundamentou-se na identificação de indicadores métricos, os quais são baseados na utilização de recursos quantitativos. Foi realizada a análise de um corpus documental composto por 350 trabalhos em onze edições do evento. Os resultados apresentaram caracterizações e comparações da produção em diversos aspectos. Através das análises conseguiram identificar um pequeno núcleo de pesquisadores mais produtivos pertencentes nas edições do Enancib realizadas nos últimos doze anos, cuja produção no âmbito desse evento é bastante significativa em termos quantitativos, revelando forte presença e atuação.

Fathala *et al.* (2020) realizam um estudo a fim de obter uma melhor compreensão das

características de eventos acadêmicos em quatro campos da ciência, analisam os metadados de eventos acadêmicos de quatro grandes campos da ciência; eventos acadêmicos de renome pertencentes a cinco subcampos. Através das análises os autores encontram resultados expressivos que possibilitam observar a evolução geral e os fatores de sucesso dos eventos acadêmicos, permitindo assim que os organizadores dos eventos julguem o progresso de seu evento ao longo do tempo e o comparem com outros eventos no mesmo campo; também é apresentada análises que possibilitam tomadas de decisões para os pesquisadores escolherem os locais adequados para a apresentação de seus trabalhos.

Em uma outra vertente, Gomes (2018) apresenta uma análise temporal das principais palavras-chave existentes nos artigos científicos. Foram utilizadas as palavras-chave de aproximadamente 14 milhões de artigos científicos do grupo de doutores referentes a todas as grandes áreas de atuação, as análises foram apresentadas como caracterização dos dados e a análise temporal dos principais tópicos de pesquisa da ciência brasileira. Os resultados iniciais são apresentados a partir da aplicação de análises bibliométricas e técnicas baseadas em análises de redes sociais sobre as palavras-chave de artigos publicados em anais de congressos e em periódicos do conjunto selecionado. Os resultados gerados podem servir de base para outros estudos que visam entender o desenvolvimento da ciência brasileira nas diversas áreas do conhecimento.

Damaceno *et al.* (2019), examinaram os resultados da genealogia acadêmica brasileira dos estudantes de mestrado e doutorado. Apresentam como as árvores genealógicas funcionam e a evolução dos pesquisadores e de seus respectivos grupos, auxiliando agências de financiamento e instituições acadêmicas para avaliar a formação dos recursos humanos, além de complementar a avaliação dos resultados da pesquisa.

## METODOLOGIA

A Plataforma Lattes possui uma integração das bases de dados em único sistema de Currículos, Grupos de pesquisa e de Instituições. Porém é de responsabilidade do indivíduo a inserção de todas suas informações curriculares na Plataforma Lattes, sendo que, após a inclusão, todos estes dados estão disponíveis em acesso aberto na internet. É um rico repositório, contemplando o registro da trajetória profissional, acadêmica e produção científica, no qual possibilita diversas e diferentes análises, justificando assim, a escolha desse repositório como fonte de dados para este trabalho. Porém, não é possível recuperar todos os currículos de uma única vez.

A metodologia utilizada neste estudo 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. Para este 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. Deste modo, o módulo Coleta, foi realizado em etapas, sendo elas:

1. Coleta de URLs: nesta etapa é realizada a obtenção dos códigos únicos de todos os currículos cadastrados, viabilizando o acesso individual ao currículo. Todos os currículos da Plataforma Lattes dispõem de um código único constituído de oito dígitos; através de uma requisição é gerada uma lista que contém todos esses códigos únicos, de todos os currículos cadastrados.

2. Coleta de identificadores e data: é feito o acesso ao cabeçalho do currículo para obter o identificador único de dezesseis dígitos e a última data de atualização do currículo. Tais dados são extraídos e armazenados localmente em um arquivo.
3. Extração dos currículos: é realizada a extração e armazenamento dos currículos. É realizado uma validação de datas; se a data de atualização na Plataforma Lattes for divergente da data de atualização do currículo armazenado localmente, o currículo local é substituído pelo mais recente.

O extrator LattesDataXplorer foi utilizado em janeiro de 2021, para realizar a coleta de todos os currículos, cerca de 7 milhões de registros. O formato dos currículos é XML (*eXtensible Markup Language*), tal formato permite delimitações e é apropriado para o processamento automático, possibilitando uma melhor manipulação dos dados. O resultado da coleta pelo extrator são os XML dos currículos em pastas, identificadas de 00 a 99. Incorporado a cada pasta os arquivos, os nomes dos arquivos são definidos de acordo com seu identificador único de 16 dígitos. Uma vez que, os dois primeiros números do identificador correspondem ao nome do repositório e os 14 dígitos restantes correspondem ao nome do arquivo XML salvo.

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 que possuem o nível de formação doutorado concluído. A escolha é feita mediante ao que Dias (2016) menciona em seu trabalho: os doutores são responsáveis por 74,51% dos artigos publicados em periódico e 64,67% dos artigos publicados em anais de congresso, além de possuir em geral data de atualização de seus currículos recente e notadamente são responsáveis pelo mais alto nível de formação.

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.

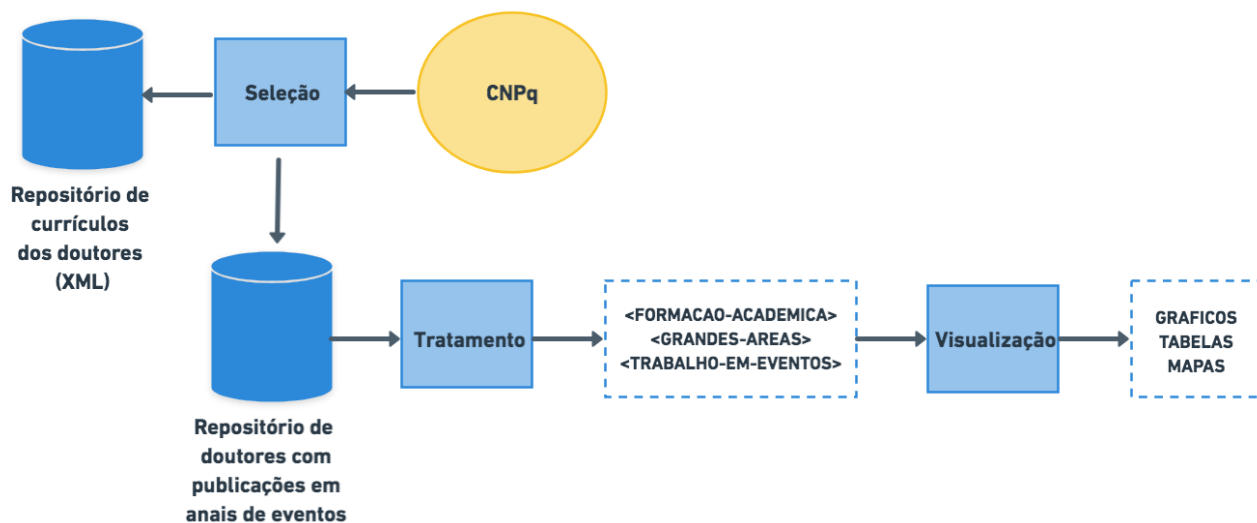


Figura 1. Processo de seleção, tratamento e visualização dos dados (Fonte: Elaborado pelo autor, 2021).

A etapa de tratamento baseou-se em analisar o conjunto de arquivos XML. Para cada currículo foi extraído um arquivo em formato XML. Os arquivos são organizados em uma estrutura de elementos e atributos, com uma hierarquia de registros do tipo pais e filhos. O currículo XML extraído da Plataforma Lattes apresenta sua estrutura o elemento raiz,

denominado como "Currículo" e possui cinco elementos filhos que possuem seus próprios elementos e atributos. Cada currículo é único e possui informações próprias; esses dados agregam informações sobre grandes áreas, formação acadêmica, orientações, produções, entre outras. Nesta etapa da análise é realizada a divisão das informações do currículo, posteriormente acessar as informações de interesse e desprezar algumas informações que para este trabalho são irrelevantes.

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.

Através da caracterização dos dados foi possível obter indicadores gerais como: total de artigos publicados em anais de eventos, ano de publicação dos artigos, totais de indivíduos que possuem artigos publicados em anais de eventos, publicações por grande área, indivíduos sem publicações, publicações com identificadores persistentes; tais indicadores são apresentados na seção de Resultados.

## RESULTADOS E DISCUSSÕES

A caracterização inicial resultou em um conjunto de 360.888 currículos de indivíduos com doutorado concluído, 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 tendo os currículos cadastrados na Plataforma Lattes como fonte de dados, podendo ser analisados por grandes áreas de atuação e realizar uma análise temporal das publicações. Como o foco deste trabalho é caracterizar os artigos publicados em anais de eventos, para as próximas análises será utilizado apenas o conjunto de dados que correspondem aos 11.416.655 artigos publicados em anais de eventos identificados neste estudo.

### **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 2).



Figura 2. Análise temporal dos artigos em anais de eventos por ano (Fonte: Elaborado pelo autor, 2021).

É possível observar o aumento constante na quantidade de publicações, especialmente a partir do final da década de 1990, sendo o ápice das publicações em 2011. Após, percebe-se uma queda significativa de publicações em eventos a partir de 2011 e em 2018 é apresentada uma queda de forma abrupta. No entanto, 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 da Plataforma Lattes.

### Artigos em anais de eventos por grande área

Os currículos da Plataforma Lattes permitem 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 3 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.



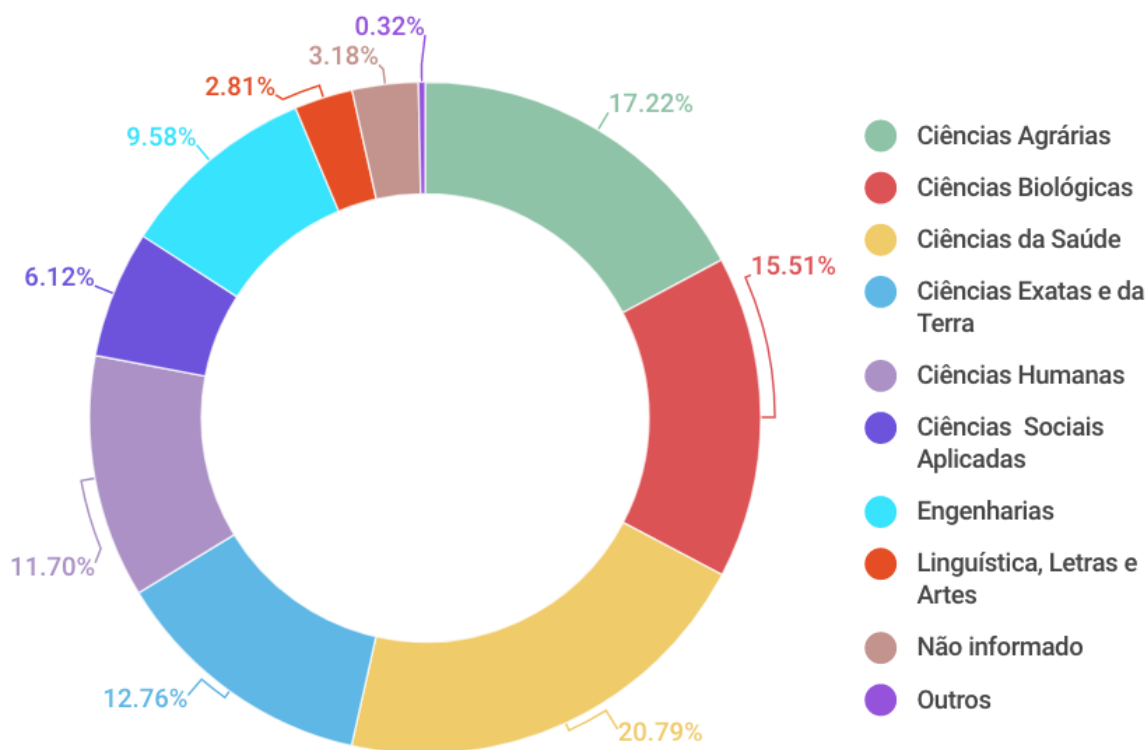


Figura 3. Quantitativo geral dos indivíduos com publicação por grandes áreas (Fonte: Elaborado pelo autor, 2021).

Percebe-se que a grande área de Ciências da Saúde apresenta maior taxa de publicação em anais de eventos (20.79%). Já a grande área de Ciências Agrárias (17.22%) é apresentada com o segundo maior percentual, e com um quantitativo próximo, a grande área de Ciências Biológicas (15.51%).

Foi possível ainda observar que a grande área que possui maior valor referenciado sem publicações em anais de eventos corresponde a Ciências Humanas (16.67%), seguido da grande área de Ciências Sociais Aplicadas (14.94%) e Ciências Exatas e da Terra (14.22), respectivamente, com quantitativos próximos.

### Identificadores persistentes

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 parcialmente automática, utilizando o identificador persistente DOI. Ao digitar o identificador é realizada a consulta e indexados os dados automaticamente em campos específicos nos currículos.

Logo em 2007, os currículos da Plataforma Lattes passaram a aceitar o DOI, possibilitando a inserção do mesmo, através de digitação manual, permitindo que a Plataforma faça uma consulta e preencha 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 dos artigos em anais de eventos cadastrados nos currículos do conjunto selecionado possuem

identificador persistente, representado assim aproximadamente 3% dos artigos. Demonstrando 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. Assim, ao recuperar as informações que possuem DOI cadastrado nas produções de seus currículos, foi possível distribuir estes por grandes áreas, conforme apresentado na Figura 4.

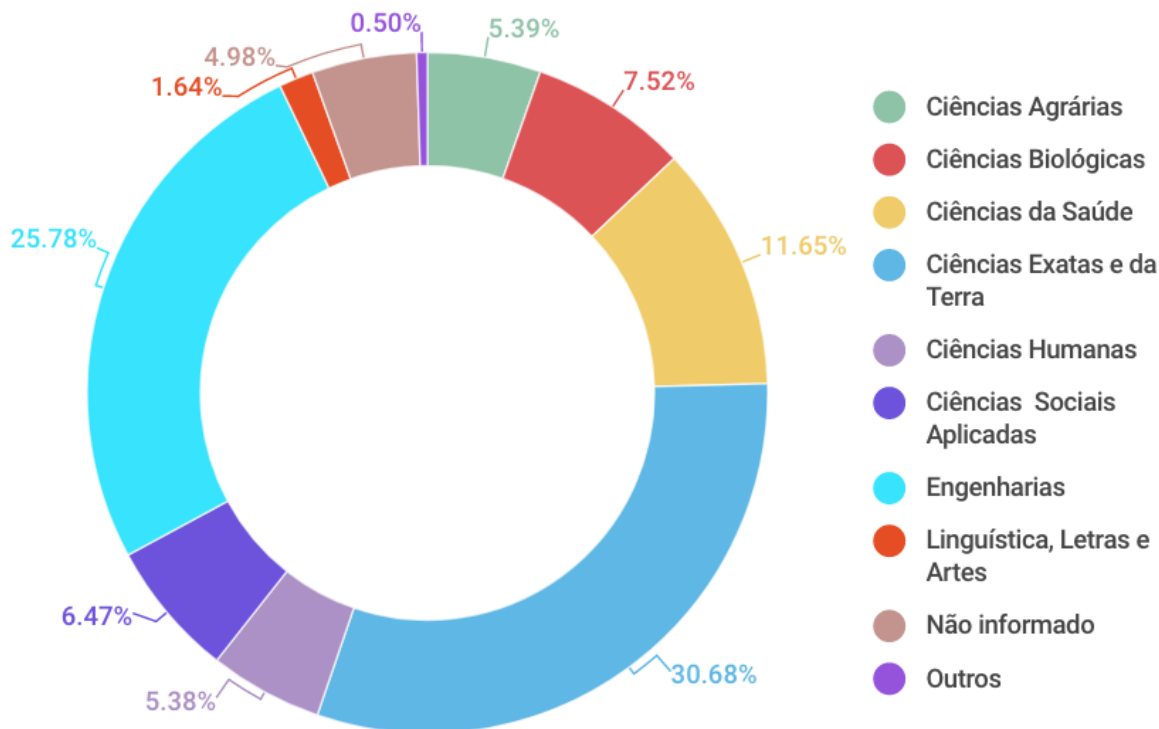


Figura 4. Quantitativo geral dos artigos com identificador persistente por grandes áreas (Fonte: Elaborado pelo autor, 2021).

Verificando os artigos publicados em anais de eventos por grande área, que possuem este identificador persistente vinculado a produção bibliográfica, pode-se verificar que ocorre uma diferença notória referente a Figura 3 sendo que, os indivíduos das grandes áreas de Ciências Exata e da Terra (30,68%) e Engenharias (25.78%) são os que mais utilizam identificador persistentes, e assim, propiciam maior facilidade de acesso aos seus trabalhos, o que pode resultar em maior visibilidade de seus estudos.

Realizando a análise de forma temporal dos artigos publicados em anais de eventos que possuem o DOI, é possível verificar alguns artigos publicados antes de 1994 (proposição do DOI) que informaram o identificador, conforme apresentado na Figura 5.

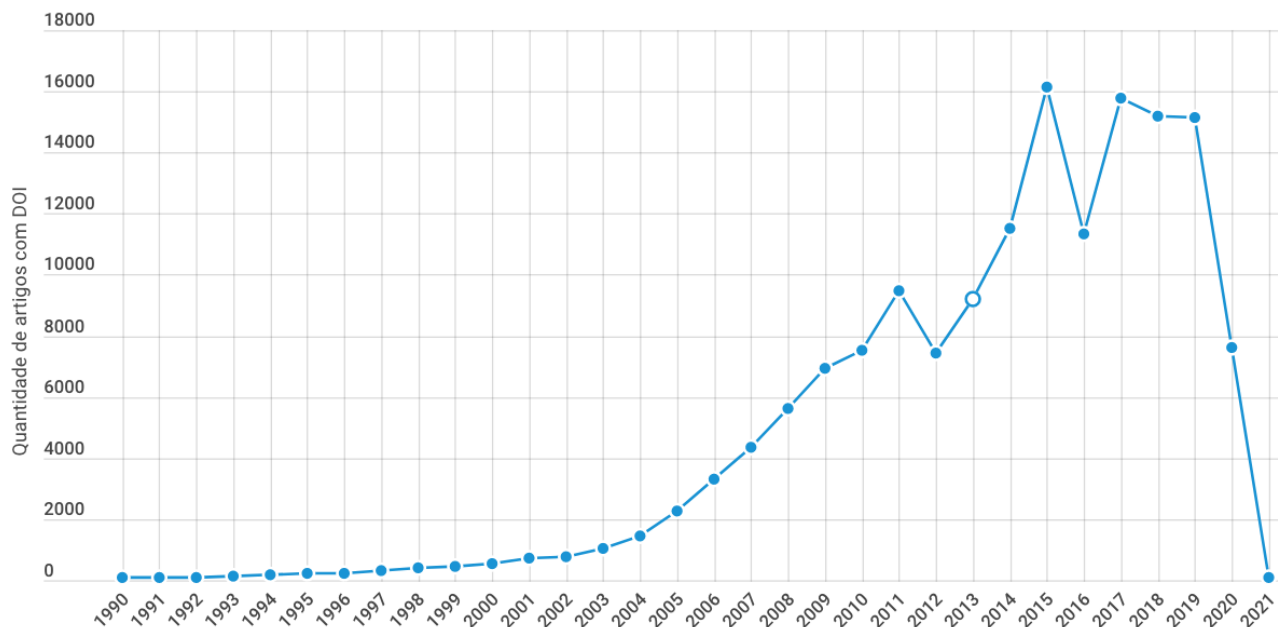


Figura 5. Análise temporal dos artigos com identificador persistente por grandes (Fonte: Elaborado pelo autor, 2021).

Observa-se que ocorre uma variação temporal significativa na utilização do identificador. Essa variação tem como uma das hipóteses o volume de eventos naquele ano, em que a quantidade de publicações possa ter influenciado nos quantitativos. Ou ainda, da forma em que estes anais foram publicados, tendo em vista que alguns eventos, sob responsabilidade de alguns organizadores, prezam pela incorporação do DOI nos anais de suas publicações.

## 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. Estudos que focam em analisar os artigos publicados em anais de eventos de forma geral são relevantes e são capazes de prover diversas análises, no âmbito de compreender o comportamento dos indivíduos participantes dos eventos, bem como, das publicações que estão sendo realizadas.

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

### 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.

### Conflitos de interesse

Os autores declaram não haver conflito de interesses.

## Declaração de consentimento de dados

Os dados gerados durante o desenvolvimento deste estudo foram incluídos no manuscrito.

## 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), 51–70. <https://doi.org/10.5007/1518-2924.2011v16n31p51>
- Araújo, C. A. (2006). Bibliometria: evolução histórica e questões atuais. *Em Questão*, (12), 11-32. Obtido em <https://seer.ufrgs.br/EmQuestao/article/view/16>
- Cardoso, T. M. M. (2018) Congresso brasileiro de biblioteconomia e documentação: análises a partir de sua produção científica (2011-2017). *Revista Brasileira de Biblioteconomia e Documentação*, (14), 69–89. Obtido em <https://rbbd.febab.org.br/rbbd/article/view/814>
- Santos, C. B., Valadares, C. B., & Marguerite, K. J. (2000). *Fontes de informação para pesquisadores e profissionais*. Belo Horizonte: 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), 94-14. <https://doi.org/10.22407/2176-1477/2017.v8i3.752>
- Damaceno, R. J. et al. (2019). The Brazilian academic genealogy: evidence of advisor–advisee relationships through quantitative analysis. *Scientometrics*, (119), 303–333. <https://doi.org/10.1007/s11192-019-03023-0>
- 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.
- Domingues, I. (2014). O sistema de comunicação da ciência e o taylorismo acadêmico: questionamentos e alternativas. *Estudos avançados*, (28), 225-250. Obtido em <https://www.revistas.usp.br/eav/article/view/88929>
- Fathalla, S., Vahdati, S., Lange, C., & Auer, S. (2020). Scholarly event characteristics in four fields of science: a metrics-based analysis. *Scientometrics*, 123(2), 677-705. <https://doi.org/10.1007/s11192-020-03391-y>
- Gomes, J. O. (2018). *Uma análise temporal dos principais tópicos de pesquisa da ciência brasileira a partir das palavras-chave de publicações científicas* (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), 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), 53-68. Obtido em <http://www.brapci.inf.br/index.php/article/download/8832>
- Moreira, W., Moraes, I. S. (2019). O assunto “classificação” na literatura brasileira de ciência

da informação: uma análise nos anais do Enancib (2003-2014). *Informação & Informação*, 24(1), 226-246. Obtido em <http://dx.doi.org/10.5433/1981-8920.2019v24n1p226>

Vanti, N. A. P. (2002). Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. *Ciência da informação*, (31), 152-162. Obtido em <http://revista.ibict.br/ciinf/article/view/970>