

# COMUNICAÇÃO CIENTÍFICA DOS DADOS DE PESQUISA SOBRE BIODIVERSIDADE GLOBAL: EVOLUÇÃO DOS REGISTROS DE METADADOS<sup>1</sup>

E-mail:  
isabelapc@gmail.com  
fpassini@gmail.com

Isabela Cardoso<sup>2</sup>; Farnanda Passini Moreno<sup>3</sup>

## RESUMO

O acesso aos dados de pesquisa permite o desenvolvimento de pesquisas adicionais, fornece informações para subsidiar decisões e contribui para a transparência na ciência. Este estudo analisa a comunicação científica dos conjuntos de dados relacionados à temática biodiversidade, sob a perspectiva do uso dos metadados ao longo do fluxo de informação científica. Para tanto, será realizada a análise comparativa entre os padrões de metadados internacionalmente recomendados e as evidências de uso dos metadados nos registros de conjuntos de dados de pesquisa. Trata-se de uma pesquisa qualitativa e descritiva, com estratégia de coleta de dados bibliográfica e documental, considerando como fontes os artigos da literatura científica e os registros de metadados dos conjuntos de dados relativos à biodiversidade extraídos em repositórios de acesso livre. As buscas por evidências de uso dos metadados serão guiadas pelo método chamado recuperação de informação realizada pouco-a-pouco (*bit-at-a-time* ou *berry picking*). Para coleta e análise adota-se a revisão sistematizada e a meta-síntese. Como resultado, espera-se delinear o panorama de uso dos metadados no registro dos dados de pesquisa sobre biodiversidade. Além disso, há a expectativa de que seja possível também identificar práticas internalizadas na gestão, comunicação e governança desses dados.

**Palavra-chave:** Padrões de Metadados. Gestão de Dados. Dados de Pesquisa. Biodiversidade. Comunicação Científica.

## ABSTRACT

Making data accessible allows further research, provides information for decision-making and contributes to transparency in science. This paper presents an ongoing study about the scientific communication of the global biodiversity data sets, analyzed under the perspective of metadata use throughout the scientific information flow. The proposed methodology aims to establish a baseline by identifying the internationally recommended metadata standards and compare it with existent metadata evidences. This study is grounded in a qualitative and descriptive research of metadata evidences that will be guided by the berry picking method. Furthermore, the systematic review and meta-synthetic analysis were defined as methods to collect and analyze metadata evidences extracted from scientific articles and metadata records related to biodiversity research data sets of public repositories. The expected outcome is to outline metadata use in biodiversity data sets in a comparative analysis view to recommended standards. It is also expected to identify practices that have been adopted in the management, communication and governance of biodiversity research data.

**Keyword:** Metadata standards. Data Management. Research data. Biodiversity. Scientific Communication.

<sup>1</sup> Relatório intermediário de mestrado aprovado em exame de qualificação e desenvolvido no Programa de Pós-graduação em Ciência da Informação da Universidade de Brasília (PPGCINF/UNB).

<sup>2</sup> Universidade Federal de Brasília, Brasília-DF. Brasil.

<sup>3</sup> Universidade Federal de Brasília, Brasília-DF. Brasil. <https://orcid.org/0000-0002-6652-0565>

## 1 INTRODUÇÃO

O acesso aos dados de pesquisa é uma demanda crescente, não apenas para a comunidade acadêmica, mas também em diferentes esferas do governo e nos mais diversos segmentos da sociedade em geral. “Os pesquisadores, as instituições acadêmicas e as agências de fomento à pesquisa começam a compreender que esses dados, se devidamente tratados, preservados e gerenciados, podem constituir uma fonte inestimável de recursos informacionais” (SAYÃO e SALES, 2015, p. 5).

A capacidade de compartilhamento e de uso dos dados de pesquisa se tornaram fatores determinantes para o desenvolvimento científico e econômico. Segundo a Fundação Americana de Ciência (NATIONAL SCIENCE FOUNDATION - NFS, 2007), a liderança mundial em ciência dependerá cada vez mais de capacidade de alavancar os reservatórios de dados científicos capturados em formato digital e transformar os dados em informação e conhecimento, auxiliados por ferramentas sofisticadas de mineração, integração, análise e visualização de dados.

“O reconhecimento do potencial informacional dos dados de pesquisa para a ciência contemporânea transforma a visão que os caracterizava como simples subprodutos dos processos de pesquisa” (SAYÃO e SALES, 2015, p. 5). Diante disso, há um movimento crescente de valorização dos dados de pesquisa como insumos para impulsionar novos estudos e gerar conhecimento, permitindo desenvolver pesquisas adicionais, subsidiar análises e decisões, bem como contribuir para a transparência na ciência.

No entanto, apesar da rapidez em que ocorrem os avanços tecnológicos, os processos tradicionais que organizam o fluxo da informação e conhecimento científicos ainda carecem de soluções focadas em ampliar a comunicação científica dos conjuntos de dados de pesquisa. As tecnologias que conduziram à explosão digital da informação têm facilitado à expansão acelerada do volume de dados de pesquisas produzidos, aumentando também a necessidade de promover a integração, uso e circulação adequada desses recursos. Isotani e Bittencourt (2015, p. 16) apontam que:

Contudo, grande parte destes dados não está disponível ao público. Estes dados tampouco estão estruturados para facilitar sua compreensão mesmo por aqueles que podem acessá-los e manipulá-los. Como resultado, a extração de informações e a produção de conhecimentos que poderiam ser úteis para a sociedade não acontecem com a agilidade e a eficácia necessárias para lidar com as questões sociais e econômicas do século 21.

Essa obstrução do fluxo de informação científica possui diversas causas, que reunidas representam prejuízos ao processo de criação, compartilhamento e uso do conhecimento científico, limitando, por esta razão, a velocidade das descobertas, o impacto dos resultados das pesquisas e o progresso da ciência (LEITE, 2011, p. 3). Dentre essas causas verifica-se que a ausência, insuficiência ou inadequado registro dos metadados representam um dos obstáculos para aperfeiçoar a gestão e comunicação dos dados de pesquisa. Para Silva e Tomaél (2009), nas “três últimas décadas, a exemplo dos demais campos da ciência, o campo da política da informação atravessou cenários repletos de transformações provocadas pela emergência e profusão das tecnologias que abarcam todas as esferas da vida social”.

Considerando esse cenário, a presente pesquisa visa analisar a comunicação científica dos dados de pesquisa sob a perspectiva do uso dos metadados ao longo do fluxo de informação científica. Devido à dimensão expressiva do universo dos dados científicos e a diversidade de práticas de comunicação científica adotadas pelas diferentes áreas do conhecimento, o escopo desse estudo restringe-se a analisar o uso dos metadados nas pesquisas sobre biodiversidade. Essa temática foi escolhida por tratar-se de área com alta relevância científica e econômica globalmente, que envolve diferentes comunidades científicas (e.g. ciências biológicas, ecologia, biomedicina, farmacologia e outras). Além disso, “o Brasil está entre os países que detém alta diversidade biológica — chamados de países megadiversos —, abrigando em seu território cerca de 20% das espécies vivas conhecidas em todo o mundo” (BRASIL, 2016, p. 29) e, portanto, tem papel decisivo e de vanguarda na geração de conhecimento sobre biodiversidade.

Assim, o objetivo principal desse estudo é avaliar a adoção dos padrões de metadados no fluxo de informação científica relacionada à biodiversidade. Para tanto, foram delineados os seguintes objetivos específicos: 1) identificar os padrões de metadados definidos para a temática biodiversidade e reconhecidos pela comunidade internacional; 2) identificar em artigos da literatura científica o uso de metadados e demais práticas adotadas no registro de dados de pesquisa sobre biodiversidade; 3) identificar o uso de metadados adotados no registro de dados das pesquisas sobre biodiversidade em repositórios com acesso livre; e 4) comparar os padrões identificados com as evidências de uso dos metadados efetivamente adotados nas pesquisas sobre biodiversidade.

## *2 REVISÃO DE LITERATURA*

O estudo sobre os padrões de metadados e sua aplicação nos conjuntos de dados de pesquisa, bem como o papel dos metadados no aperfeiçoamento do fluxo da informação científica, demanda a compreensão de conceitos inseridos nos seguintes campos da CI: Gestão, Comunicação e Organização da Informação.

“A ciência e a produção do conhecimento científico têm sido objeto de estudo de várias áreas do conhecimento, entre elas a Filosofia da Ciência, a Sociologia da Ciência e a Ciência da Informação – CI” (MUELLER e PASSOS, 2000, p. 13). A CI tem estudado, desde a sua gênese, questões relacionadas à comunicação científica como fenômeno central de interesse, uma vez que a disciplina surgiu em razão da preocupação de cientistas, tecnólogos e documentalistas de meados do século passado com os fluxos da informação em ciência e tecnologia (COSTA, LEITE e TAVARES, 2018, p. 17)

A comunicação da informação é responsável pelos fluxos de informação, e fenômenos associados, que viabilizam a atividade humana em diferentes contextos; já a gestão da informação e a gestão do conhecimento são as responsáveis por sua sistematização e dão conta de sua complexidade na contemporaneidade (COSTA, LEITE e TAVARES, 2018, p. 15). Tais definições induzem à consideração de que há relação de complementaridade e interdependência entre as duas práticas. Isso porque, segundo os autores, por um lado, a gestão do conhecimento disciplina, sistematiza e torna mais efetivos os processos de comunicação. Por outro, a comunicação permite que a gestão do conhecimento seja viabilizada, na medida em que possibilita, entre outros processos, a interação entre indivíduos e, por consequência, a criação, o compartilhamento e o uso do conhecimento.

O processo de comunicação científica é um ciclo contínuo e regenerativo (LANCASTER e SMITH, 1978, p. 369), que inclui atividades associadas com a produção, disseminação e uso da informação, desde a hora em que o cientista teve a ideia da pesquisa até o momento em que os resultados de seu trabalho são aceitos formalmente como parte integrante do conhecimento científico (GARVEY, 1979). “Enquanto de alguma forma os canais informais de comunicação são tão antigos quanto à própria ciência, outros canais têm origem muito mais recente” (LANCASTER e SMITH, 1978, p. 370), devido aos avanços tecnológicos e o crescimento da comunicação informal.

De forma geral, a comunicação científica constitui tópico bastante explorado e discutido na CI, gerando vários modelos teóricos, abordagens e contextos encontrados na literatura, que evoluíram para absorver as mudanças tecnológicas (COSTA, 2005, p. 165). Já a gestão da informação, considerando diversos posicionamentos encontrados na literatura, Costa e Leite (2018, p. 27) relatam que o entendimento predominante é de que esta área da CI deve se ocupar do conhecimento registrado. Os autores também afirmam que “tanto para a investigação quanto para a aplicação de conhecimentos na solução de problemas reais relacionados com a melhoria da comunicação da informação entre seres humanos, a Ciência da Informação se vale da perspectiva de processos de gestão da informação” (COSTA e LEITE, 2018, p. 32)

Ao encontro dessa perspectiva, Campos et al (2005, p. 58-59) entendem que a gestão da informação se insere como um conceito em construção, contemplando o gerenciamento das informações por meio do tratamento de conteúdos semânticos. As autoras explicam que a qualidade desse tratamento tem como eixos fundamentais as investigações relativas a padrões de intercâmbio, controle de linguagem e modelos de representação, utilizando metadados, vocabulários, ontologias, tesouros, taxonomias, entre outras. Por fim, as autoras concluem que a gestão da informação envolve uma série de procedimentos de uniformização e padronização de informações, baseados em uma visão integradora, fundamentada em dois conceitos essenciais: o de organização e o de comunicação.

Dessa forma, compreende-se que o uso dos metadados no fluxo da informação científica é assunto circunscrito entre os procedimentos de uniformização e padronização de informações. Por isso, também é relevante compreender essa temática pela ótica do campo da Organização da Informação.

Organizar significa categorizar, incluir, em função de algum critério, determinados elementos em uma classe e, nesse mesmo movimento, excluir outros elementos, pressupõe, portanto, a identificação de semelhanças e de diferenças (SMIT e KOBASHI, 2002, p. 11-12). Segundo Lima e Cunha (2006, p. 1), “os esforços para a organização da informação geram modelos artificiais que tentam representar o mundo real, criando sistemas e esquemas de Para Café e Sales (2010) a organização da informação é um processo de arranjo de acervos tradicionais ou eletrônicos realizado por meio da descrição física e de conteúdo (assunto) de seus objetos informacionais. Um mesmo objeto pode apresentar diversas descrições sucessivas e interdependentes, que se constituem em um enunciado de suas propriedades ou das relações desse objeto com outros que o identificam. Assim, “os fundamentos da catalogação de assuntos, da classificação, da indexação e da análise documental são abordagens do tratamento temático da informação responsáveis pela formação da organização da informação como a conhecemos hoje” (CAFÉ e SALES, 2010, p. 116).

Entre os sistemas de catalogação, segundo Alves (2010, p. 38) o período de desenvolvimento dos padrões de metadados foi marcado pelo aumento de estudos sobre a

recuperação da informação, que se tornou foco de investigação nas áreas de Ciência da Informação e Ciência da Computação. Essas áreas convergiram na busca de soluções, criando uma infraestrutura tecnológica e representacional para diversos ambientes científicos.

A descrição do assunto na literatura apresenta-se com diversas nomenclaturas, como Análise documentária, Análise Temática, Análise de Assunto, Descrição de Conteúdo e Tratamento Temático da Informação (CAFÉ e SALES, 2010, p. 118). Essa descrição produz a chamada informação documentária por meio dos processos de classificação, indexação e resumo ou condensação documental (KOBASHI, 2018; CAFÉ e SALES, 2010).

Para compreender o papel dos metadados, Boughida (2005, p. 49) entende que é importante especificar e diferenciar os padrões de linguagem documentária envolvidos na catalogação, classificação e indexação. Para tanto o autor, propõe o agrupamento em quatro categorias: 1) conteúdo, as regras que orientam o preenchimento de um determinado campo de metadados; 2) valor, os termos e outros valores com os quais se devem preencher um campo de metadados; 3) estrutura, as categorias, 4) organização e relação entre os elementos de metadados; e formato, a expressão técnica ou codificação de dados em um arquivo ou uma tabela de banco de dados.

Harpring (2016, p. 203) explica que a anotação dos valores nos campos ou elementos de metadados dedicados especificamente a certos elementos de conteúdo garante que os dados sejam consistentemente registrados e indexados no mesmo lugar, utilizando as mesmas convenções. A autora ressalta que sem adequados metadados descritivos, a informação permanece inacessível e, por isso, uma solução completa demanda o uso de vocabulários controlados (HARPRING, 2016, p. 22).

“Nos últimos cem anos, pelo menos, a criação e o gerenciamento de metadados foram principalmente de responsabilidade de profissionais da informação envolvidos em catalogação, classificação e indexação” (GILLILAND, 2016). Segundo Sayão (2010, p. 2), os metadados foram usados inicialmente no contexto dos bancos de dados para descrever e controlar a gestão e o uso dos dados, mas é comum associar o termo ao seu uso em bibliotecas, considerando seu papel em um esquema formal para descrição de todo tipo de objetos informacionais, digitais e não digitais. Entretanto, foi sua aplicação nos meios digitais e na Web aumentaram suas interpretações e uso para muito além dos profissionais da informação.

Embora o termo metadados não seja muito familiar entre usuários comuns, criadores e consumidores de conteúdo digital na Web, todos esses indivíduos são cada vez mais estimulados a criar, explorar e avaliar metadados nos ambientes digitais, desde títulos de páginas da Web, folksonomias, marcadores sociais, até outros meios voltados à descrição ou avaliação da qualidade dos recursos informacionais. Entretanto, “esse processo social não é necessariamente colaborativo, porque o esforço geralmente não é organizado; na realidade, os indivíduos não trabalham em conjunto ou de forma coordenada, e não empregam a padronização e um vocabulário comum” (HARPRING, 2016, p. 52).

Gilliland (2016) apresenta metadados como “a soma total do que se pode dizer em um dado momento sobre qualquer objeto de informação em qualquer nível de agregação”. Para a autora, “um objeto de informação é qualquer coisa que possa ser endereçada e manipulada como uma entidade discreta por um ser humano ou um sistema de informação”. Por fim, a autora destaca que, “em geral, todos os objetos de informação, independentemente da forma física ou intelectual, possuem três recursos - conteúdo, contexto e estrutura -, os quais podem e devem ser refletidos por meio de metadados”. Segundo a Organização Americana de Padronização da Informação, “metadado é a informação estruturada que descreve, explica, localiza um recurso informacional,

ou possibilita sua recuperação, uso e gerenciamento. O termo metadados frequentemente designa dados sobre dados” (NISO - NATIONAL INFORMATION STANDARD ORGANIZATION, 2017, p. 1).

Diante da expansão da abordagem e uso, tem sido cada vez mais importante que, além dos profissionais de informação, pesquisadores, criadores e utilizadores de conteúdo digital, entendam os principais papéis dos diferentes tipos de metadados, de forma a assegurar acessibilidade, autoria, interoperação, granulação e preservação dos dados científicos e seus sistemas de registro.

As atuais práticas de financiamento e publicação exigem cada vez mais o compartilhamento dos conjuntos de dados resultantes das pesquisas e a maioria dos pesquisadores entende que a publicação dos dados é, portanto, não apenas desejável, mas necessária. De fato, roteiros para eventualmente institucionalizar a publicação de conjuntos de dados foram propostos (MORITZ, KRISHNAN, *et al.*, 2011) e muitos desafios para tal mudança cultural já foram identificados (CHAVAN, PENEV e HOBERN, 2013). No entanto, tornar os dados acessíveis, compreensíveis e verdadeiramente reutilizáveis continua a ser um desafio (COSTELLO e WIECZOREK, 2014), mas esses dados não podem ser compartilhados e usados sem critérios e precisam ser mobilizados de acordo com padrões estabelecidos, amplamente conhecidos e aceitos (PETERS, 2010).

Assim, verifica-se a necessidade de desenvolver estudos voltados a aperfeiçoar o fluxo da informação científica para adequá-lo a esse novo contexto. Desenvolver o conhecimento visando avanço científico e tecnológico demanda o aperfeiçoamento do fluxo de informação científica, com destaque às necessidades de redução de custos, compartilhamento da informação e aumento da produção científica. Segundo Leite (2011, p. 2), quanto mais rápida e completamente pesquisadores receberem a informação científica necessária às suas atividades, mais produtos científicos eles gerarão a custos menores. Frente a essa crescente necessidade agilizar a descoberta de novos conhecimentos e agregar valor a qualquer informação, Isotani e Bittencourt (2015, p. 17) afirmam que diversas empresas, governos e institutos de pesquisa têm realizado esforços para disponibilizar dados e produzir tecnologias que permitam criar um ecossistema de produção e consumo de dados. Desse modo, apesar da existência de diversos modelos e abordagens voltados à comunicação científica e à gestão da informação desenvolvidos na CI, verifica-se que ainda é necessário ampliar o conhecimento em busca de soluções voltadas a aperfeiçoar o uso, circulação e compartilhamento de conjuntos de dados.

### 3 PROCEDIMENTOS METODOLÓGICOS

Esta pesquisa apresenta propósito descritivo (KOTHARI, 2004, p. 37), finalidade básica (APPOLINÁRIO, 2006) e sua abordagem é qualitativa e indutiva (KAUARK, MANHÃES e MEDEIROS, 2010, p. 26; CRESWELL, 2010, p. 26 e 92). A estratégia de coleta de dados se caracteriza como bibliográfica e documental, que tem como fontes artigos disponíveis na literatura científica e evidências documentais extraídas dos repositórios de dados analisados. A figura 1 a seguir sintetiza as estratégias, métodos e instrumentos utilizados para: 1) busca e coleta de material bibliográficos e documental; 2) coleta, síntese e estruturação dos dados; e 3) análise dos dados coletados.



**Figura 1.** Relação entre as estratégias adotadas na pesquisa para busca, coleta e análise.

As buscas para identificar os padrões de metadados utilizam a ferramenta *Disciplinary Metadata* disponibilizada pelo Digital Curation Center – DCC (DCC, 2020). Essa ferramenta permite a busca desses padrões por áreas temáticas e campos do conhecimento relacionados. As buscas por artigos científicos e registros dos metadados são realizadas na ferramenta de “Busca por Assunto” do Portal de Periódicos da Capes (BRASIL. MINISTÉRIO DA EDUCAÇÃO. CAPES, 2019), que permite realizar a investigação multidisciplinar em diversas áreas do conhecimento.

As escolhas dos termos de busca são guiadas pelo método chamado recuperação de informação realizada pouco a pouco ou a cada instante (*bit-at-a-time* ou *berrypicking*) proposto por Bates (1989), que inicia a partir de um tópico amplo ou com uma referência relevante e, em seguida, movimentar-se por diversas fontes. Nesse método, cada nova informação encontrada oferece novas ideias e orientações e, conseqüentemente, nova concepção da consulta que pode levar a modificação dos termos de busca. Dessa forma, as consultas não são estáticas, mas estão em constante evolução, sendo realizada em partes, até que seja atendido o objetivo da pesquisa.

Após encontrar e selecionar o material bibliográfico e documental, os procedimentos de extração, síntese e estruturação dos dados são iniciados. Os métodos utilizados são a revisão sistematizada e análise meta-sintética (MENDES, SILVEIRA e GALVAO, 2008) de artigos científicos, aplicados para obtenção de informações sobre os padrões de metadados identificados, bem como extração das evidências de uso dos metadados nas pesquisas. Os instrumentos definidos para auxiliar a coleta e estruturação dos dados definiu-se um roteiro de estudo para os artigos científicos e uma lista de verificação para as evidências documentais dos repositórios de conjuntos de dados de pesquisa.

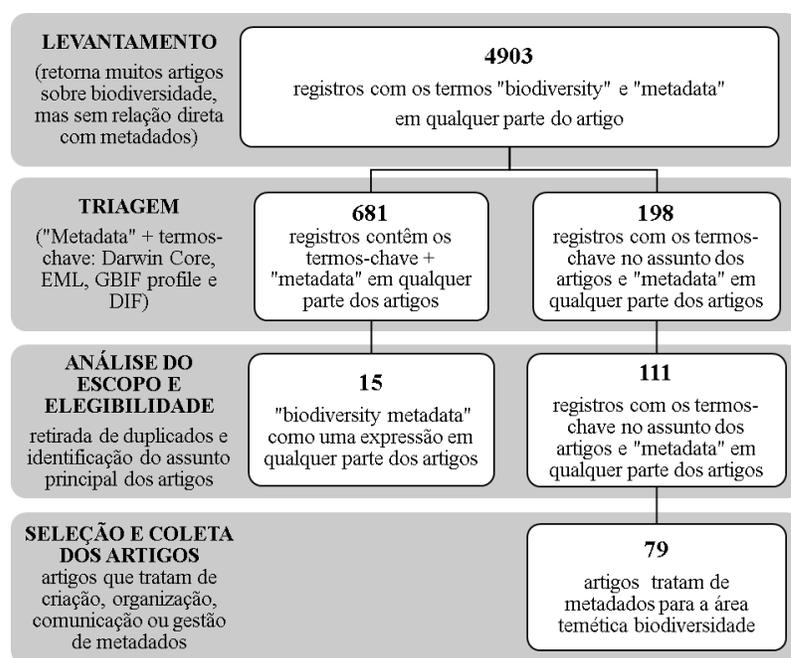
Por fim, será realizada a análise integrada das evidências de uso dos metadados e a análise comparativa dessas evidências com os padrões de metadados recomendados para biodiversidade.

#### 4 CONSIDERAÇÕES PRELIMINARES

O uso dos metadados em conjuntos de dados de pesquisas e os padrões de metadados recomendados são assuntos circunscritos às áreas de Organização, Gestão e de Comunicação da Informação, estabelecendo relações integradoras entre esses campos da CI. Sob essa perspectiva, a fundamentação teórica apresentada tem foco em identificar os conceitos fundamentais à

compreensão da aplicação dos padrões de metadados na pesquisa científica e na disseminação de seus resultados, com ênfase no aperfeiçoamento do uso e circulação dos dados no fluxo da informação científica. Além disso, a literatura revisada permite identificar as lacunas e desafios relacionados ao uso, circulação e compartilhamento de conjunto de dados.

Foram realizados pré-teste de busca e coleta de artigos que tratam dos padrões e uso de metadados na temática biodiversidade. A metodologia proposta permitiu identificar 79 artigos relevantes, publicados entre os anos de 1999 e 2019 – ver figura 2. Todos os artigos selecionados são revisados por pares e a maior parte deles trata três tópicos dentro do tema estudado: 1) metadados no processo de digitalização e publicação de conjuntos de dados da biodiversidade; 2) relevância do uso de metadados na pesquisa de biodiversidade; 3) o papel central dos padrões de metadados no desenvolvimento de portais, sistemas ou soluções de integração de conjuntos de dados da biodiversidade.



**Figura 2** – Resultado do pré-teste realizado para a busca e seleção de artigos científicos.

Também foi possível identificar três padrões metadados recomendados e reconhecidos internacionalmente para a temática biodiversidade: 1) *Access to Biological Collections Data* (ABCD) é um padrão baseado em XML estabelecido para o acesso e intercâmbio de dados primários sobre a biodiversidade, incluindo espécimes e dados primários (TDWG, 2019); 2) *Darwin Core* (DwC) se tornou o padrão de acesso aberto mais utilizado para dados de biodiversidade, pois foi desenvolvido para fornecer uma maneira simples de documentar e compartilhar informações sobre ocorrências de espécies, coletadas em campo ou a partir de uma coleção de museu. Sua última versão foi publicada em outubro de 2015 (TDWG, 2015); e 3) *Ecological Metadata Language* (EML) é um padrão de metadados que registra informações sobre conjuntos de dados ecológicos em documentos XML modulares e extensíveis. Trata-se de uma especificação de metadados inicialmente desenvolvida para a disciplina de ecologia, embasada em trabalhos realizados pela *Ecological Society of America and Associated Efforts*

(ECOINFORMATICS, 2019). Além disso, identificou-se também a existência do GBIF *Metadata Profile* (GMP), que é o perfil de metadados desenvolvido pela rede *Global Biodiversity Information Facility* (GBIF) criado para padronizar a descrição dos conjuntos de dados. Destaca-se que existem várias maneiras de escrever um documento de metadados em conformidade com esse perfil e que seu desenvolvimento é embasado nos três padrões de metadados identificados (GBIF, 2017).

A identificação dos padrões e perfil de metadados permitiu melhor delineamento dos parâmetros utilizados para as buscas por artigos, aperfeiçoando a abordagem para seleção do material mais adequado ao foco desta pesquisa. Destaca-se que partir da análise dos artigos serão selecionados os repositórios para a coleta de evidências sobre a prática de uso de metadados.

Por fim, verifica-se que os resultados do pré-teste apontam para a conformidade da metodologia para subsidiar as análises necessárias ao alcance do objetivo desse estudo. A partir da configuração metodológica proposta, espera-se delinear o panorama de adoção uso dos metadados no registro dos conjuntos de dados sobre biodiversidade e compreender se os padrões recomendados têm sido utilizados. Há também a expectativa de identificar as práticas internacionalmente internalizadas na gestão, comunicação e governança dos dados científicos sobre biodiversidade.

## REFERÊNCIAS

ALVES, R. C. V. **Metadados como elementos do processo de catalogação - Tese (Doutorado em Ciência da Informação)**. Universidade Estadual Paulista. Marília, p. 132. 2010.

APPOLINÁRIO, F. **Metodologia da ciência: filosofia e prática da pesquisa**. São Paulo: Cen-gage Learning, v. xii, 2006. 209 p.

BATES, M. J. The design of browsing and berrypicking techniques for the online search interface". **Online Review**, v. 13, n. 5, p. 407-424, 1989.

BOUGHIDA, K. B. CDWA Lite for Cataloging Cultural Objects (CCO): A New XML Schema for the Cultural Heritage Community, Amsterdam, p. 49-54, 2005. ISSN 90-6984-456-7.

BRASIL. **Conhecendo a Biodiversidade**. Ministério da Ciência, Tecnologia, Inovações e Comunicações - MCTIC; Programa de Pesquisa em Biodiversidade - PPBio; Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq. Brasília, p. 196. 2016. (ISBN 978-85-63100-08-5).

BRASIL. MINISTÉRIO DA EDUCAÇÃO. CAPES. Portal de Periódicos da Capes. **Institucional**, 2019. Disponível em: <<https://www.periodicos.capes.gov.br>>. Acesso em: 21 set. 2019.

CAFÉ, L. M. A.; SALES, R. D. Organização da informação: Conceitos básicos e breve fundamentação teórica. In: ROBREDO, J.; BRÄSCHER, M. **Passeios pelo bosque da informação: estudos sobre representação e organização da informação e do conhecimento** - EROIC. Brasília: IBICT, 2010. Cap. 6. 329 p, p. 115-129. ISBN 978-85-7013-072-3.

CAMPOS, M. L. M.; CAMPOS, M. L. D. A.; CAMPOS, L. M. Web semântica e a gestão de conteúdos informacionais. In: SAYÃO, L. **Bibliotecas digitais: saberes e práticas**. Salvador, Brasília: EDUFBA, IBICT, 2005. p. 55-75. ISBN ISBN 8523203508. Disponível em: <<http://livroaberto.ibict.br/handle/1/1013>>.

CHAVAN, V.; PENEV, L.; HOBERN, D. Cultural Change in Data Publishing Is Essential. **BioScience**, v. 63, n. 6, p. 419-420, 2013. ISSN <https://doi.org/10.1525/bio.2013.63.6.3>.

COSTA, S. M. D. S.; LEITE, F. C. L. L. Imbricações teóricas entre comunicação e gestão da informação e do conhecimento na Ciência da Informação. In: COSTA, S. M. D. S.; LEITE, F. C. L.; TAVARES, R. B. **Comunicação da informação, gestão da informação e gestão do conhecimento**. Brasília: IBICT, 2018. p. 17-44. ISBN ISBN 978-85-7013-147-8. Disponível em: <<http://livroaberto.ibict.br/handle/123456789/1071>>. Acesso em: 30 abr. 2019.

COSTA, S. M. D. S.; LEITE, F. C. L.; TAVARES, R. B. **Comunicação da informação, gestão da informação e gestão do conhecimento**. 1ª. ed. Brasília: Ibict, 2018. 415 p. ISBN 978-85-7013-147-8. Disponível em: <<http://livroaberto.ibict.br/handle/123456789/1071>>. Acesso em: 2019 jan. 11.

COSTA, S. M. S. O novo papel das tecnologias digitais na comunicação científica. In: SAYÃO, L. **Bibliotecas Digitais**. Salvador. Brasília: UFBA, IBICT, v. 1, 2005. p. 165-183. Disponível em: <<http://repositorio.unb.br/handle/10482/1437>>. Acesso em: 11 março 2019.

COSTELLO, M. J.; WIECZOREK, J. Best practice for biodiversity data management and publication. **Biological Conservation**, v. 173, p. 68-73, 2014.

CRESWELL, J. W. **PROJETO DE PESQUISA: MÉTODOS QUALITATIVO, QUANTITATIVO E MISTO; TRADUÇÃO**. Tradução de Magda Lopes. 3ª. ed. Porto Alegre: ARTMED, 2010. 296 p.

DCC. Metadata Standards. **Digital Curation Center - DCC**, 2020. Disponível em: <<http://www.dcc.ac.uk/resources/metadata-standards>>. Acesso em: 05 mar. 2020.

ECOINFORMATICS. Tools - Metadata Specification. **Ecoinformatics**, 2019. Disponível em: <<http://ecoinformatics.org/tools.html>>. Acesso em: 10 dez. 2019.

GARVEY, W. D. **Communication: the essence of science - facilitating information among Librarians, Scientists, Engineers and Students**. 1ª. ed. Oxford: Pegamon, 1979. 344 p. ISBN ISBN 9781483182070.

GBIF. GBIF Metadata Profile – How-to Guide. **Global Biodiversity Information Facility (GBIF)**, 2017. Disponível em: <<https://github.com/gbif/ipt/wiki/GMPHowToGuide>>. Acesso em: 19 dez. 2019.

GILLILAND, A. J. Setting the stage, Los Angeles, 2016. Disponível em: <<http://www.getty.edu/publications/intrometadata/>>. Acesso em: 03 maio 2019.

HARPRING, P. **Introdução aos vocabulários controlados**: terminologia para arte, arquitetura e outras obras culturais. São Paulo: Se. Tradução de Christina Maria Müller. São Paulo: Secretaria da Cultura do Estado: Pinacoteca de São Paulo: ACAM Portinari, v. 4, 2016. 288 p. ISBN ISBN 978-85-63566-19-5.

ISOTANI, S.; BITTENCOURT, I. I. **Dados abertos conectados**. [S.l.]: Novatec, 2015. Disponível em: <<http://ceweb.br/livros/dados-abertos-conectados/>>. Acesso em: 04 jun. 2018. versão html.

KAUARK, F.; MANHÃES, F. C.; MEDEIROS, C. H. **Metodologia da pesquisa**: guia prático. Itauna: VIA LITTERARUM, 2010. 86 p.

KOTHARI, C. R. **Research Methodology**: Methods and Techniques. 2<sup>a</sup>. ed. Nova Deli: NEW AGE INTERNATIONAL (P) LIMITED, PUBLISHERS, 2004. 401 p.

LANCASTER, F. W.; SMITH, L. C. Science, scholarship and the communication of knowledge. **Library Trends**, v. 27, n. 3, p. 367-387, 1978.

LEITE, F. C. L. Modelo genérico de gestão da informação científica para instituições de pesquisa na perspectiva da comunicação científica e do acesso aberto - Tese (Doutorado em Ciência da Informação). **Repositório Institucional da Universidade de Brasília**, Brasília, p. 262, 2011. Disponível em: <<http://repositorio.unb.br/handle/10482/9753>>. Acesso em: 10 agosto 2018.

LIMA, J. A. D. O.; CUNHA, M. B. D. **Modelo genérico de relacionamento e a organização da informação**. CONFERÊNCIA IBEROAMERICANA DE PUBLICAÇÕES ELETRÔNICAS NO CONTEXTO DA COMUNICAÇÃO CIENTÍFICA. Anais. Campo Grande: UNIDERP. 2006. p. 67-74.

MENDES, K. D. S.; SILVEIRA, R. C. D. C. P.; GALVAO, C. M. Revisão integrativa: método de pesquisa para a incorporação de evidências na saúde e na enfermagem, Florianópolis, 2008. ISSN <http://dx.doi.org/10.1590/S0104-07072008000400018>. Acesso em: 12 ago. 2019.

MORITZ, T. et al. Towards mainstreaming of biodiversity data publishing: recommendations of the GBIF Data Publishing Framework Task Group. **BMC Bioinformatics**, n. 12, S1, 2011. ISSN <https://doi.org/10.1186/1471-2105-12-S15-S1>.

MUELLER, S. P. M.; PASSOS, E. J. L. As questões da comunicação científica e a ciência da informação. In: MUELLER, S. P. M.; PASSOS, E. J. L.; (ORGS) **Comunicação científica**. Brasília: Ciência da Informação, 2000. p. 13-22. Disponível em: <<http://repositorio.unb.br/handle/10482/1444>>.

NATIONAL SCIENCE FOUNDATION - NFS. Cyberinfrastructure Vision for 21st Century Discovery, Arlington, 2007. Disponível em: <<https://www.nsf.gov/pubs/2007/nsf0728/>>. Acesso em: 11 abr. 2018.

NISO - NATIONAL INFORMATION STANDARD ORGANIZATION. **Understanding metadata what is metadata, and what is it for?** National Information Standards Organization - NISO. Baltimore, p. 45. 2017.

PETERS, D. P. C. Accessible ecology: synthesis of the long, deep, and broad. **Trends in Ecology & Evolution**, v. 25, n. 10 - Special Issue: Long-term ecological research, 2010. ISSN <https://doi.org/10.1016/j.tree.2010.07.005>.

SAYÃO, L. F. Uma outra face dos metadados: informações para a gestão da preservação digital. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, v. 15, n. 30, p. 1-31, 2010.

SAYÃO, L. F.; SALES, L. F. **Guia de Gestão de Dados de Pesquisa para Bibliotecários e Pesquisadores**. CNEN/IEN. Rio de Janeiro, p. 90. 2015. (978-85-61905-03-3).

SILVA, T. E. D.; TOMAÉL, M. I. **Política de Informação: tendências internacionais**. Anais - GT5 - X Encontro Nacional de Pesquisa em Ciência da Informação. João Pessoa: [s.n.]. 2009.

SMIT, J. W.; KOBASHI, N. Y. **Como elaborar vocabulário controlado para aplicação em arquivos**. Como Fazer. ed. São Paulo: [s.n.], v. 10, 2002. 56 p. ISBN 85-86726-52-4 e 85-7060-216-2.

TDWG. Simple Darwin Core. **Biodiversity Information Standards (TDWG)**, 2015. Disponível em: <<http://rs.tdwg.org/dwc/simple/>>. Acesso em: 17 dez. 2019.

TDWG, B. I. S. ABCD 3.0. **Access to Biological Collection Data**, 2019. Disponível em: <<https://abcd.tdwg.org/3.0/>>. Acesso em: 16 out 2019.