



24° ENANCIB
Encontro Nacional de Pesquisa em Ciência da Informação
Perspectivas Contemporâneas na Ciência da Informação
• Vitória - ES • Ancib • PPGCI/UFES



XXIV ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – XXIV ENANCIB

ISSN 2177-3688

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

**CONTRIBUIÇÕES BRASILEIRAS PARA A BIODIVERSIDADE GLOBAL: ANÁLISE DE DATA PAPERS
UTILIZANDO DATASETS DO GBIF**

***BRAZILIAN CONTRIBUTIONS TO GLOBAL BIODIVERSITY: ANALYSIS OF DATA PAPERS USING
GBIF DATASETS***

Fabiano Couto Corrêa da Silva – Universidade Federal do Rio Grande do Sul (UFRGS)

Amanda Santos Witt – Universidade Federal do Rio Grande do Sul (UFRGS)

Bryan Nicollas Soares Costa – Universidade Federal do Rio Grande do Sul (UFRGS)

Laura Vilela Rodrigues Rezende – Universidade Federal de Goiás (UFG)

Modalidade: Resumo Expandido

Resumo: O Global Biodiversity Information Facility (GBIF) fornece acesso a dados sobre biodiversidade, o que inclui *datasets* e *data papers*. O estudo verifica o uso desses recursos por pesquisadores e universidades brasileiras, de 2010 a 2024. Foram encontrados 536 registros, de tipologias variadas. Com relação aos artigos, 27 são de pesquisadores brasileiros e desse total, 26 são de acesso aberto. Conclui-se que universidades brasileiras contribuem de modo significativo ao manter *datasets* sobre flora, fauna e ecossistemas marinhos atualizados. Os principais periódicos mapeados foram Biodiversity Data Journal e Gigabyte. Nesse sentido, a pesquisa destaca a contribuição das universidades brasileiras ao GBIF.

Palavras-chave: Data papers; Datasets; Biodiversidade.

Abstract: The Global Biodiversity Information Facility (GBIF) provides access to biodiversity data, including datasets and data papers. The study verifies the use of these resources by Brazilian researchers and universities from 2010 to 2024. A total of 536 records of various types were found. With regard to the articles, 27 are by Brazilian researchers and of this total, 26 are open access. It is concluded that Brazilian universities make a significant contribution to keeping datasets on marine flora, fauna and ecosystems up to date. The main journals mapped were Biodiversity Data Journal and Gigabyte. In this sense, the research highlights the contribution of Brazilian universities to GBIF.

Keywords: Data papers; Dataset; Biodiversity.

1 INTRODUÇÃO

A crescente relevância dos *data papers* na comunicação científica ressalta a importância do uso e compartilhamento de conjuntos de dados (*datasets*) para promover a transparência, reprodutibilidade e colaboração na pesquisa. Os *data papers* permitem que os pesquisadores publiquem descrições detalhadas dos seus conjuntos de dados, garantindo que outras pessoas possam acessar e utilizar essas informações de maneira eficiente e confiável. Nesse contexto, o Global Biodiversity Information Facility (GBIF) se destaca como uma plataforma internacional que disponibiliza vastos conjuntos de dados sobre biodiversidade, sendo essencial para pesquisas científicas que necessitam de informações nessa área (Pinheiro, 2003).

Na contemporaneidade, evidenciam-se novas formas de produzir ciência pautadas no compartilhamento e reuso de dados científicos. Estes deixam de ser entendidos apenas como insumos intermediários de uma pesquisa, o que exige a elaboração de métodos e processos que salientem o seu valor (Henning *et al.*, 2019). Práticas direcionadas à abertura de dados podem estar ligadas ao movimento mais amplo da Ciência Aberta (Rocha, 2019), caracterizado pelos fatores conhecimento científico aberto, tais como softwares abertos e dados científicos abertos; infraestruturas de pesquisa compartilhadas (virtuais ou físicas), que envolvem equipamentos ou conjuntos de instrumentos científicos e variados recursos como é o caso das plataformas e repositórios; engajamento aberto dos atores sociais com fins de ampliar a participação social em ciência aproximando cientistas e sociedade e, por fim, o diálogo aberto com outros sistemas de conhecimento (Unesco, 2022).

A análise do uso de *datasets* do GBIF por pesquisadores brasileiros pode trazer à tona diversas tendências, desafios e oportunidades. Identificar como esses *datasets* são utilizados pode revelar lacunas na pesquisa, áreas com maior atividade científica e demandas específicas de dados. É fundamental considerar a necessidade de implementação de políticas mais efetivas voltadas à disponibilização de dados de pesquisa, visando facilitar o uso e reuso dos dados por toda a comunidade científica, como destacado em estudo sobre compartilhamento de dados no contexto da ciência brasileira (Araújo; Dias; Autran, 2021).

Além disso, a pesquisa acadêmica brasileira tem demonstrado um crescente interesse pelo estudo do uso e compartilhamento de dados científicos. Isso pode ser observado na aproximação entre comunicação científica e divulgação científica, conforme identificado em

XXIV Encontro Nacional de Pesquisa em Ciência da Informação – XXIV ENANCIB
Vitória-ES – 04 a 08 de novembro de 2024

estudo sobre comunidades científicas e infraestrutura tecnológica no Brasil para uso de recursos eletrônicos de comunicação e informação na pesquisa (Lustosa, 2024).

Os dados ecológicos demonstram sua relevância e utilidade especialmente no que diz respeito ao monitoramento e à conservação da biodiversidade. Ademais, seu acesso na modalidade livre bem como sua reutilização permitem identificar locais em que determinadas espécies estão ameaçadas de extinção, contribuindo com o processo de tomada de decisões para reverter o quadro (Fox *et al.*, 2019). O crescente interesse pelo tema da biodiversidade e o desenvolvimento de várias iniciativas nesse sentido nos últimos anos, estão atrelados à proposição, no ano de 2015, pela Organização das Nações Unidas, de uma agenda voltada para o desenvolvimento sustentável, denominada Agenda 2030, o que culminou na criação de 17 Objetivos de Desenvolvimento Sustentável. Isto demonstra a crescente preocupação e os esforços globais direcionados para acabar com a pobreza, visando, dentre outros elementos, assegurar sustentabilidade, justiça social e melhores condições de vida para todas as pessoas de todos os lugares do mundo.

Este estudo é norteado pela seguinte problemática: a plataforma de relevância internacional GBIF, tem se destacado no Brasil, com relação a disponibilização de *data papers* e *datasets* no âmbito da biodiversidade, nos últimos dez anos, por pesquisadores e instituições brasileiras? Dessa forma, objetiva-se traçar um panorama geral da publicação de *data papers* correlacionados aos *datasets* brasileiros, com fins de compreender melhor como a disponibilização e utilização de conjuntos de dados podem impactar a pesquisa científica no Brasil, promovendo a transparência, reprodutibilidade e colaboração necessárias para avanços significativos na área da biodiversidade.

2 PROCEDIMENTOS METODOLÓGICOS

Utilizou-se o mapeamento de periódicos que publicam artigos de dados relacionados à biodiversidade, fornecido pelo GBIF em sua versão mais atualizada, por meio do qual verificou-se a existência de 36 periódicos e o passo seguinte foi analisar características dos referidos periódicos e suas publicações.

Para identificar e analisar os *datasets* publicados por instituições e centros de pesquisa do Brasil utilizamos a interface de busca do GBIF, acessível pelo site oficial da rede, aplicando filtros específicos para delimitar a busca pelos *datasets* publicados por universidades e

XXIV Encontro Nacional de Pesquisa em Ciência da Informação – XXIV ENANCIB
Vitória-ES – 04 a 08 de novembro de 2024

instituições de pesquisa brasileiras. Os filtros incluem parâmetros como o país de publicação (Brasil), tipo de instituição (universidades e centros de pesquisa) e categorias de dados (flora, fauna e ecossistemas marinhos).

O período considerado para esta pesquisa abrange os anos de 2014 a 2024. Este recorte temporal foi escolhido para capturar uma visão abrangente das tendências e padrões de publicação de *data papers* ao longo de mais de uma década, permitindo uma análise das mudanças e evoluções na contribuição brasileira para a base de dados do GBIF. Este intervalo de tempo também coincide com o aumento da digitalização e do acesso aberto a dados de biodiversidade, refletindo as práticas contemporâneas de compartilhamento de dados científicos.

Para esta pesquisa, é importante entender a diferença entre *data papers* e *datasets*. *Data papers* são artigos científicos que descrevem a coleta, o processamento e a qualidade dos dados, oferecendo uma visão detalhada e padronizada dos conjuntos de dados disponibilizados. Já os *datasets* são os conjuntos de dados propriamente ditos, que contêm as informações brutas coletadas durante as pesquisas.

Utilizamos filtros específicos na plataforma do GBIF para identificar *data papers* associados aos *datasets* brasileiros e verificamos a correspondência entre eles utilizando identificadores únicos como o Digital Object Identifier (DOI) e outras referências bibliográficas. Na sequência, analisamos a frequência de publicação e atualização dos *datasets* pelas instituições identificadas, e padrões e tendências na publicação de dados, identificando possíveis áreas de destaque ou necessidade de maior incentivo.

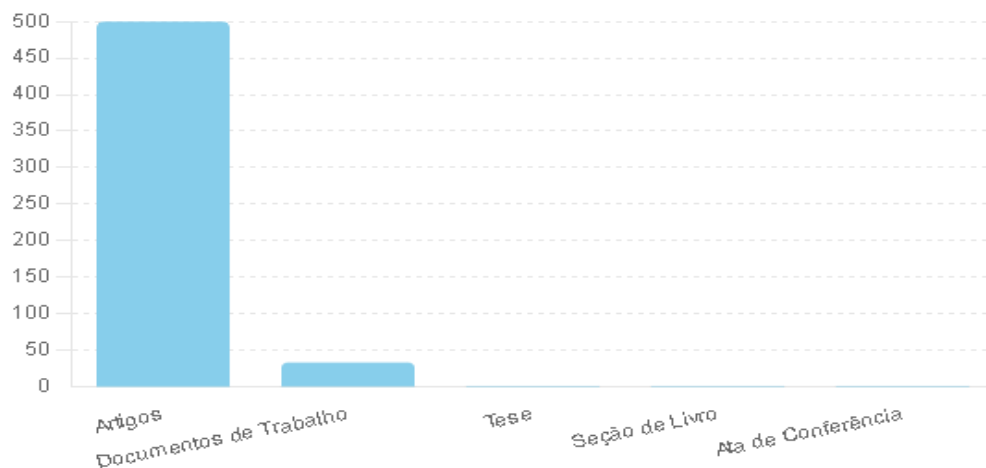
Extraímos informações detalhadas dos *datasets*, incluindo título, instituição publicadora, data de publicação e atualizações, tipo de dados (flora, fauna, ecossistemas), número de registros e DOI para referência. Verificamos a autenticidade e a atualização dos *datasets* utilizando ferramentas como o Darwin Core Archive (DwC-A) para garantir a padronização dos metadados. Focamos em identificar *datasets* publicados por instituições brasileiras de destaque, reconhecidas por sua vasta produção acadêmica e contribuição significativa em pesquisas de biodiversidade.

3 RESULTADOS

A pesquisa inicial revelou o total de 536 *data papers* publicados utilizando *datasets* do GBIF. Com relação ao tipo de trabalho científico produzido, dividem-se em 500 artigos, 33 documentos de trabalho (relatórios técnicos, *white papers* ou *working papers*, etc.), 1 tese, 1 seção de livro e 1 ata de conferência (ver Gráfico 1).

Ao filtrar os estudos por "país ou área do investigador" e selecionar o Brasil, foram localizados 27 estudos, estando todos publicados em forma de artigos. Destes, 26 são de acesso aberto, e 22 identificam o país ou área de cobertura dos dados, sendo 21 artigos cobrindo o Brasil e 1 o Chile.

Gráfico 1 - Distribuição de Tipos de Documentos



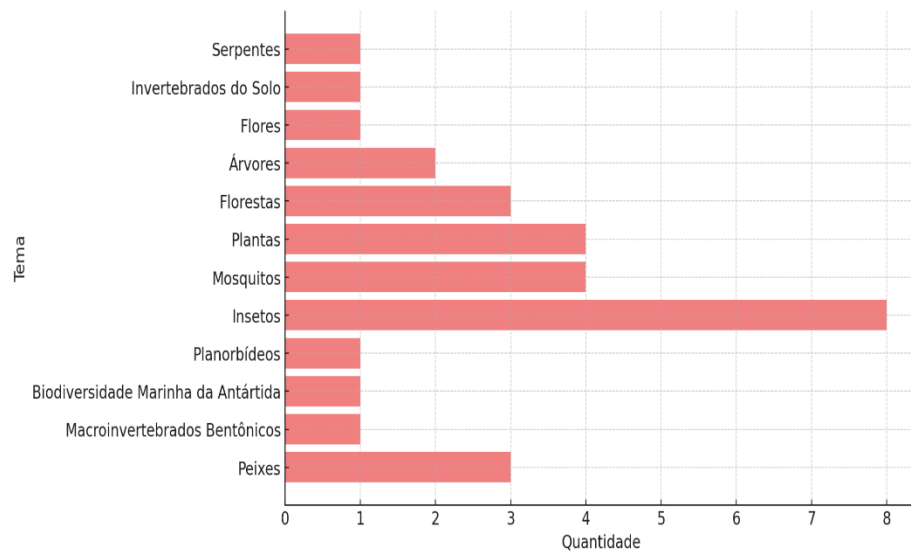
Fonte: Dados da pesquisa (2024)

No âmbito da Biodiversidade, foram identificadas temáticas variadas nos 27 artigos, as quais foram divididas em ambientes aquático e terrestre (ver Gráfico 2), dois importantes Objetivos do Desenvolvimento Sustentável (ODS), 14 e 15 respectivamente, definidos pela Organização das Nações Unidas:

- Ambiente aquático: Três artigos sobre a fauna ictiológica (peixes) incluindo os seguintes temas: macroinvertebrados bentônicos, produtores primários estuarinos e biodiversidade marinha da Antártida e planorbídeos (com foco nas espécies de moluscos *biomphalaria*, responsáveis pela disseminação da esquistossomose);
- Ambiente terrestre: Em maior número destacam-se oito estudos sobre insetos, abrangendo vagalumes, abelhas, besouros, e formigas, quatro estudos sobre

XXIV Encontro Nacional de Pesquisa em Ciência da Informação – XXIV ENANCIB
Vitória-ES – 04 a 08 de novembro de 2024

mosquitos, os flebotomíneos (incluem os principais vetores de transmissão da leishmaniose), um estudo sobre os *culicidae* (família de insetos habitualmente chamados de muriçocas, mosquitos ou pernilongos); quatro estudos sobre plantas; três sobre florestas; dois sobre árvores; aparecem com um estudo



cada: flores; invertebrados do solo e serpentes.

Gráfico 2 - Distribuição de temas

Fonte: Dados da pesquisa (2024)

Ademais, verificou-se uma variedade de biomas e regiões analisados nas pesquisas, tais como, Mata Atlântica, Cerrado, Pantanal, Caatinga e Amazônia, demonstrando a riqueza brasileira em termos de biodiversidade e o amplo potencial do uso de dados para a elaboração de *data papers* referentes à diversidade biológica do país. Tal cenário descrito vai ao encontro das proposições do ODS 14 - Vida na água cuja premissa é a conservação e uso de forma sustentável dos oceanos, mares e recursos marinhos e do ODS 15 - Vida terrestre, que visa a proteção, restauração e promoção do uso sustentável dos ecossistemas terrestres, combatendo a degradação dos solos e perda da biodiversidade (ONU Brasil, 2024).

Os periódicos que publicam esses *data papers* incluem:

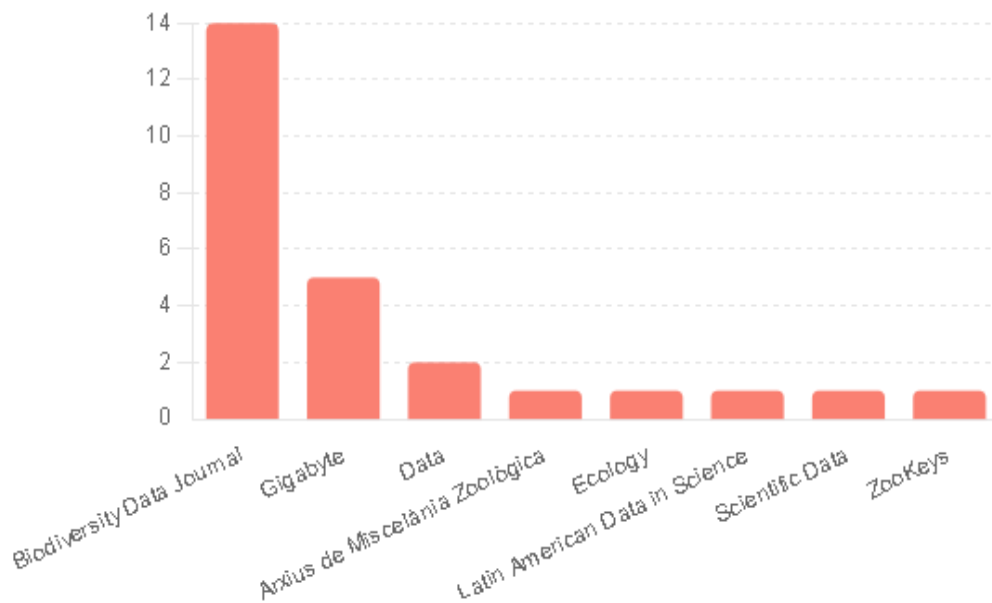
- Biodiversity Data Journal: 14 artigos;
- Gigabyte: 5 artigos;
- Data: 2 artigos;
- Arxius de Miscellània Zoològica: 1 artigo;
- Ecology: 1 artigo;
- Latin American Data in Science: 1 artigo;

XXIV Encontro Nacional de Pesquisa em Ciência da Informação – XXIV ENANCIB
Vitória-ES – 04 a 08 de novembro de 2024

- Scientific Data: 1 artigo;
- ZooKeys: 1 artigo

O gráfico a seguir desvela os periódicos que realizam publicação de *data papers*:

Gráfico 3 - Periódicos que publicam *data papers*



Fonte: Dados da pesquisa (2024)

A análise dos artigos identificou 51 conjuntos de dados citados, sendo necessário verificar aqueles que não mostram os DOIs para garantir a rastreabilidade. Apenas um artigo não apresentou o DOI na interface do GBIF, em dados referidos no estudo, que foi o artigo intitulado “Herbarium of the Pontifical Catholic University of Paraná (HUCP), Curitiba, Southern Brazil”, o qual fez uma análise sobre a coleção em si do conjunto de dados da instituição PUC, denominado “Herbário da Pontifícia Universidade Católica do Paraná - HUCP”.

Para uma análise mais aprofundada, exploramos os *datasets* publicados por instituições brasileiras, com ênfase nas universidades. Ao acessar o link para a busca de *datasets* publicados por instituições brasileiras no GBIF, observamos que existem diversos conjuntos de dados frequentemente atualizados e utilizados por pesquisadores no Brasil. A análise inicial indica uma forte contribuição das universidades brasileiras na publicação desses *datasets*, os quais abrangem uma variedade de dados de biodiversidade, incluindo flora, fauna e ecossistemas marinhos.

XXIV Encontro Nacional de Pesquisa em Ciência da Informação – XXIV ENANCIB
Vitória-ES – 04 a 08 de novembro de 2024

Um exemplo relevante encontrado no GBIF é o conjunto de dados intitulado "Meiobentos, National Monitoring Study 2019, EMBLAS Plus Project", publicado pelo Ukrainian Scientific Centre of Ecology of the Sea (UkrSCES). Este inclui dados sobre a composição de espécies, abundância e biomassa do meiobentos coletados durante o estudo de monitoramento nacional nas águas ucranianas do Mar Negro. Trata-se de um exemplo de como os dados de biodiversidade são coletados, processados e disponibilizados através do GBIF, permitindo a utilização por pesquisadores ao redor do mundo para estudos ecológicos, monitoramento ambiental e avaliações de impacto (GBIF, 2024).

O filtro aplicado para *data papers* revelou que, dos 536 registros identificados, uma quantidade significativa é de artigos científicos, indicando uma tendência crescente de publicação de *data papers* pelos pesquisadores brasileiros. Esses artigos são publicados em periódicos renomados e frequentemente acessados, demonstrando a relevância dos *datasets* do GBIF para a comunidade científica brasileira.

A análise sobre os periódicos que publicam *data papers* de biodiversidade indica que o Biodiversity Data Journal (BDJ) é um dos principais canais de divulgação desses artigos. O BDJ facilita a publicação de dados de biodiversidade de várias formas, incentivando o compartilhamento e a reutilização dos dados através de uma plataforma de acesso aberto. Este periódico é altamente reconhecido e frequentemente utilizado para a publicação de *data papers* relacionados a dados de biodiversidade, incluindo iniciativas recentes como a chamada para *papers* sobre biodiversidade do solo em colaboração com o GBIF e o Finnish Biodiversity Information Facility (FinBIF). O BDJ possui uma estrutura que acelera o processo de revisão e publicação, tornando-o uma escolha popular entre os pesquisadores. Além do BDJ, periódicos importantes na área incluem o Gigabyte e outros de alto impacto que também contribuem significativamente para a disseminação científica dos *data papers*.

4 CONSIDERAÇÕES FINAIS

Este estudo destacou a importância e a relevância dos *data papers* na comunicação científica, especialmente no contexto da biodiversidade. A utilização e o compartilhamento de *datasets* por meio do Global Biodiversity Information Facility (GBIF) promovem a transparência, reprodutibilidade e colaboração na pesquisa científica, elementos fundamentais para o avanço do conhecimento.

XXIV Encontro Nacional de Pesquisa em Ciência da Informação – XXIV ENANCIB
Vitória-ES – 04 a 08 de novembro de 2024

A análise dos *data papers* publicados por pesquisadores brasileiros revelou uma contribuição significativa das universidades brasileiras, particularmente a Universidade de São Paulo (USP), Universidade Federal de Minas Gerais (UFMG) e Universidade Estadual de Campinas (UNICAMP), para a base de dados do GBIF. Os *datasets* examinados englobam uma ampla variedade de dados de biodiversidade, incluindo flora, fauna e ecossistemas marinhos, evidenciando o papel crucial dessas instituições na coleta e na disponibilização de informações essenciais para a pesquisa científica. O acesso livre e a padronização desses dados, registrados pelas instituições brasileiras, incrementam o cenário da pesquisa no país e permitem análises em escala global, fortalecendo redes de cooperação internacional.

A identificação de 536 registros, incluindo 500 artigos, reforça a tendência crescente de publicação de *data papers* por pesquisadores brasileiros. A predominância de periódicos como o Biodiversity Data Journal (BDJ) como canal de divulgação sublinha a importância deste meio na disseminação científica, facilitando o compartilhamento e a reutilização dos dados. Os resultados obtidos demonstram a necessidade de políticas mais efetivas para a disponibilização de dados de pesquisa, visando facilitar o uso e reuso desses dados por toda a comunidade científica. A implementação dessas políticas pode fortalecer a infraestrutura de pesquisa, aumentando a visibilidade e o impacto das instituições brasileiras no cenário científico global.

O Brasil em função de sua extensão territorial, diversidade geográfica e climática, apresenta vasta diversidade biológica, o que amplifica o seu papel na produção, armazenamento e disseminação de dados sobre biodiversidade. Em conclusão, este estudo não apenas sublinha a importância do GBIF como uma plataforma vital para a biodiversidade global, mas também ressalta a contribuição brasileira para essa iniciativa. A continuidade e expansão dessas práticas são essenciais para promover avanços científicos e enfrentar os desafios relacionados à biodiversidade de maneira colaborativa e inclusiva.

REFERÊNCIAS

ARAÚJO, Débora Gomes de; DIAS, Guilherme Ataíde; AUTRAN, Marynice de Medeiros Matos. Compartilhamento de dados no contexto da ciência brasileira: um estudo integrativo. **Informação & Informação**, Londrina, v. 26, n. 3, p. 100, 2021. Disponível em: <https://doi.org/10.5433/1981-8920.2021v26n3p100>.

XXIV Encontro Nacional de Pesquisa em Ciência da Informação – XXIV ENANCIB
Vitória-ES – 04 a 08 de novembro de 2024

FOX, Richard; BOURN, Nigel A. D.; DENNIS, Emily B.; HEAFIELD, Richard T.; MACLEAN, Ilya M. D.; WILSON, Robert J. Opinions of citizen scientists on open access to UK butterfly and moth occurrence data. **Biodiversity and Conservation** v. 28, p. 3321-3341, 2019. Disponível em: <https://link.springer.com/article/10.1007/s10531-019-01824-6#citeas>. Acesso em: 10 set. 2024.

GLOBAL BIODIVERSITY INFORMATION FACILITY - GBIF. **Artigos de dados**. Disponível em: <https://www.gbif.org/pt/data-papers>. Acesso em: 08 jun. 2024.

GLOBAL BIODIVERSITY INFORMATION FACILITY - GBIF. **Meiobenthos, National Monitoring Study 2019, EMBLAS Plus Project**. Ukrainian Scientific Centre of Ecology of the Sea (UkrSCES). Disponível em: <https://doi.org/10.15468/gxp27f>. Acesso em: 27 jun. 2024.

GLOBAL BIODIVERSITY INFORMATION FACILITY - GBIF. **Site institucional** [artigos de dados]. Disponível em: https://www.gbif.org/pt/resource/search?contentType=literature&topics=DATA_PAPER&relevance=GBIF_PUBLISHED. Acesso em: 08 jun. 2024.

GLOBAL BIODIVERSITY INFORMATION FACILITY - GBIF. **Site institucional** [instituições publicadoras]. Disponível em: <https://www.gbif.org/pt/publisher/search?q=brasil>. Acesso em: 08 jun. 2024.

HENNING, Patricia; RIBEIRO, Claudio José Silva; SANTOS, Luiz Olavo Bonino da Silva; SANTOS, Paula Xavier dos GO FAIR e os princípios FAIR: o que representam para a expansão dos dados de pesquisa no âmbito da Ciência Aberta. **Em questão**. Porto Alegre, v. 25, n. 2, maio/ago. 2019. Disponível em: <https://seer.ufrgs.br/EmQuestao/article/view/84753/0>. Acesso em: 10 set. 2024.

LUSTOSA, Marllus de Melo; FARIAS, Maria Giovanna Guedes; FARIAS, Gabriela Belmont de. Inteligência artificial e comunicação científica. **Brazilian Journal of Information Science: Research Trends**, Marília, v. 18, e024004, 2024. Disponível em: <https://doi.org/10.36311/1981-1640.2024.v18.e024004>. Acesso em: 27 jan. 2025.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS - BRASIL. **Os Objetivos de Desenvolvimento Sustentável no Brasil**. 2024. Disponível em: <https://brasil.un.org/pt-br/sdgs>. Acesso em: 15 jun. 2024.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS PARA A EDUCAÇÃO, A CIÊNCIA E A CULTURA - UNESCO. **Recomendação da UNESCO sobre Ciência Aberta**. Distrito Federal: Representação da UNESCO no Brasil, 2022.

PINHEIRO, Lena Vania Ribeiro. Comunidades científicas e infra-estrutura tecnológica no Brasil para uso de recursos eletrônicos de comunicação e informação na pesquisa. **Ciência Da Informação**, Brasília, v. 32, n. 3, p. 62-73, set. 2003. Disponível em: <https://doi.org/10.1590/s0100-19652003000300008>.

XXIV Encontro Nacional de Pesquisa em Ciência da Informação – XXIV ENANCIB
Vitória-ES – 04 a 08 de novembro de 2024

ROCHA, Luana M. Pinto. **Os cientistas e a ciência cidadã: um estudo exploratório sobre a visão dos pesquisadores profissionais na experiência brasileira**. 2019. Dissertação (Mestrado em Ciência da Informação) - Escola de Comunicação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2019. Disponível em:
<https://ridi.ibict.br/bitstream/123456789/1053/1/dissertacao-final-LuanaRocha-Ciencia%20cidadada%20e%20cientistas%20profissionais.pdf>. Acesso em: 10 set. 2024.