

# Infraestruturas de Ciência e de Acesso Aberto no Brasil:

iniciativas do Instituto Brasileiro de Informação em Ciência e Tecnologia

## Organizadores

Bianca Amaro

Phillipe Campos

Janinne Barcelos



Editora  
Ibict

# **INFRAESTRUTURAS DE CIÊNCIA E DE ACESSO ABERTO NO BRASIL**

## INICIATIVAS DO INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA



**PRESIDÊNCIA DA REPÚBLICA**

Luís Inácio Lula da Silva

**Presidente da República**

Geraldo José Rodrigues Alckmin Filho

**Vice-Presidente da República**

**MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÕES**

Luciana Barbosa de Oliveira Santos

**Ministra da Ciência, Tecnologia e Inovações**

Luis Manuel Rebelo Fernandes

**Secretário Executivo**

**INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA**

Tiago Emmanuel Nunes Braga

**Diretor**

Carlos André Amaral de Freitas

**Coordenador de Administração – COADM**

Ricardo Medeiros Pimenta

**Coordenador de Ensino e Pesquisa, Ciência e Tecnologia da Informação – COEPI**

Cecília Leite Oliveira

**Coordenadora-Geral de Informação Tecnológica e Informação para a Sociedade – CGIT**

Alexandre Faria de Oliveira

**Coordenador-Geral de Tecnologias de Informação e Informática – CGTI**

Washington Luís Ribeiro de Carvalho Segundo

**Coordenador-Geral de Informação Científica e Técnica - CGIC**

Marcel Garcia de Souza

**Coordenador de Tratamento, Análise e Disseminação da Informação Científica – CODIC**

Bianca Amaro  
Fhillipe de Freitas Campos  
Janinne Barcelos  
Organização

**INFRAESTRUTURAS DE CIÊNCIA E  
DE ACESSO ABERTO NO BRASIL**  
**INICIATIVAS DO INSTITUTO  
BRASILEIRO DE INFORMAÇÃO  
EM CIÊNCIA E TECNOLOGIA**





Esta obra é licenciada sob uma licença *Creative Commons* - Atribuição CC BY-NC-ND 4.0, sendo permitida a reprodução parcial ou total, desde que mencionada a fonte, de uso não comercial e sem derivações.

#### CONSELHO EDITORIAL

Gustavo Silva Saldanha | Milton Shintaku | Luana Sales  
Franciéle Garcês | Leyde Klébia Rodrigues da Silva | Stella Moreira Dourado |  
Daniel Strauch | Walisson Oliveira

#### COMITÊ EDITORIAL

Tiago Braga	Milton Shintaku
Henrique Denes	Cecília Leite Oliveira
Ricardo Pimenta	Leda Cardoso Sampson Pinto
Carlos André Amaral de Freitas	Marcel Souza
Alexandre Oliveira	Washington Segundo
Emanuelle Torino	Alexandre Faria de Oliveira

#### COMITÊ CIENTÍFICO

Ania Rosa Hernández Quintana – Universidad de La Habana, Cuba  
Fernanda do Valle – Universidade Federal do Estado do Rio de Janeiro, Unirio, Brasil  
María Arminda Damus – Universidad Nacional de Misiones, Argentina  
Martha Sabelli – Universidad de La República - Uruguay  
Natalia Duque Cardona – Universidad de Antioquia, Colombia  
Vinicio Menezes – Universidade Federal de Sergipe, UFS, Brasil  
Carlos Alberto Ávila Araújo – Universidade Federal de Minas Gerais, UFMG, Brasil

#### EQUIPE TÉCNICA

Revisão linguística	Walisson Oliveira
Diagramação	Franciéle Garcês; Walisson Oliveira
Normalização	Walisson Oliveira
Revisão	Stella Dourado; Franciéle Garcês; Walisson Oliveira
Capa	Andrea Castro Fleury Curado

Dados Internacionais de Catalogação na Publicação (CIP)

143

Infraestruturas de Ciência e de Acesso Aberto no Brasil: iniciativas do Instituto Brasileiro de Informação em Ciência e Tecnologia / Bianca Amaro; Phillippe de Freitas Campos; Janinne Barcelos (Organização).  
Brasília: Editora Ibitc, 2025.  
292 p.

Inclui Bibliografias.  
Disponível em: <https://editora.ibict.br>.  
ISBN Digital: 978-85-7013-254-3 (digital).  
ISBN Físico: 978-85-7013-252-9 (impresso).

1. Ciência Aberta. 2. Acesso Aberto. 3. Informação científica. 4. Instituto Brasileiro de Informação em Ciência e Tecnologia. I. Amaro, Bianca. II. Campos, Phillippe de Freitas. III. Barcelos, Janinne. IV. Título.

CDU 025.4:001.891(81)

Bibliotecário: Walisson Oliveira - CRB 1/3477

Como citar:

AMARO, Bianca; CAMPOS, Phillippe de Freitas; BARCELOS, Janinne. (org.). **Infraestruturas de ciência e de acesso aberto no Brasil**: iniciativas do Instituto Brasileiro de Informação em Ciência e Tecnologia. Brasília, DF: Editora Ibitc, 2025. 292 p.

As opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade dos autores, não exprimindo, necessariamente, o ponto de vista do Instituto Brasileiro de Informação em Ciência e Tecnologia ou do Ministério da Ciência, Tecnologia e Inovação.

**Endereço: IbiCT – Instituto Brasileiro de Informação em Ciência e Tecnologia**

Setor de Autarquias Sul (SAUS), Quadra 05, Lote 06, Bloco H – 5º andar

CEP: 70.070-912 - Brasília, DF



## **AGRADECIMENTOS**

Nenhuma grande obra é fruto de esforços isolados. A criação deste livro é um exemplo brilhante de como a união de muitas mentes pode contribuir para a produção e disseminação do conhecimento.

Queremos expressar nossos agradecimentos a todos os autores que aceitaram o desafio e contribuíram para esta obra. Foi a participação dedicada de cada um de vocês que tornou este livro possível. Cada capítulo, cada ideia compartilhada e cada página escrita é um reflexo do trabalho colaborativo e da paixão coletiva pelo avanço do Acesso Aberto e da Ciência Aberta no Brasil.

Obrigado por serem parte essencial desta jornada e por ajudarem a construir um futuro mais acessível e inclusivo para a ciência e a informação.

Nossos agradecimentos,

Bianca Amaro  
Fhillipe de Freitas Campos  
Janinne Barcelos  
Organização



# SUMÁRIO

<b>PREFÁCIO .....</b>	<b>13</b>
Eloy Rodrigues	
<b>APRESENTAÇÃO .....</b>	<b>21</b>
Bianca Amaro	
Phillipe de Freitas Campos	
Janinne Barcelos	
<b>LINHA DO TEMPO .....</b>	<b>27</b>
<b>BIBLIOTECA DIGITAL BRASILEIRA DE TESES E DISSERTAÇÕES (BDTD) .....</b>	<b>33</b>
Tainá Batista de Assis	
Juliana Araujo Gomes de Sousa	
Cássio Teixeira de Moraes	
<b>PORTAL BRASILEIRO DE PUBLICAÇÕES E DADOS CIENTÍFICOS EM ACESSO ABERTO (OASISBR) .....</b>	<b>55</b>
Gabriel Silveira Marques	
<b>REDE CARINIANA: UM ESFORÇO COLABORATIVO DE PRESERVAÇÃO DIGITAL DO IBICT .....</b>	<b>77</b>
Miguel Ángel Márdero Arellano	
<b>REPOSITÓRIO INSTITUCIONAL DO IBICT (RID) .....</b>	<b>95</b>
Cássio Teixeira de Moraes	
<b>REPOSITÓRIO COMUM DO BRASIL (DEPOSITA) .....</b>	<b>107</b>
Blena Estevam dos Santos	
Millena Cordeiro Matos de Lima	
Cássio Teixeira de Moraes	
<b>ALEIA: REPOSITÓRIO DE DADOS DE PESQUISA DO IBICT .....</b>	<b>119</b>
Tatyane Guedes Martins da Silva	
Leticia Guarany Bonetti	

<b>REPOSITÓRIO COMUM DE DADOS DE PESQUISA: DEPOSITA DADOS .....</b>	<b>131</b>
Tatyane Guedes Martins da Silva	
Leticia Guarany Bonetti	
<b>DIRETÓRIO DE POLÍTICAS EDITORIAIS DAS REVISTAS CIENTÍFICAS BRASILEIRAS (DIADORIM) .....</b>	<b>143</b>
Hillipe de Freitas Campos	
Bianca Amaro	
Denise Aparecida Freitas de Andrade	
<b>PORTAL BRASILEIRO PARA AS REVISTAS CIENTÍFICAS (MANUELZÃO) .....</b>	<b>153</b>
Denise Aparecida Freitas de Andrade	
Hillipe de Freitas Campos	
Raphael Faria Vilas Boas	
<b>DIRETÓRIO DAS REVISTAS CIENTÍFICAS ELETRÔNICAS BRASILEIRAS (MIGUILIM) .....</b>	<b>163</b>
Denise Aparecida Freitas de Andrade	
Hillipe de Freitas Campos	
Bianca Amaro	
<b>REDE MOARA: PLATAFORMA PARA COMPARTILHAMENTO DE CÓDIGOS-FONTE PARA PESQUISA .....</b>	<b>173</b>
Milton Shintaku	
Bernardo Dionízio Vechi	
Rebeca dos Santos de Moura	
Lucas Rodrigues Costa	
Diego José Macedo	

<b>O PGD BR COMO FERRAMENTA PARA A PROMOÇÃO DA CIÊNCIA ABERTA .....</b>	<b>189</b>
Fabiano Couto Corrêa da Silva	
Sandra de Albuquerque Siebra	
Alexandre Faria de Oliveira	
Laura Vilela Rodrigues Rezende	
<b>INFRAESTRUTURAS ABERTAS PARA CIÊNCIA CIDADÃ: A EXPERIÊNCIA DA CÍVIS .....</b>	<b>213</b>
Sarita Albagli	
Luana Rocha	
Josir Cardoso Gomes	
Matheus Dantas	
<b>ECOSSISTEMA DE INFORMAÇÃO DA PESQUISA CIENTÍFICA BRASILEIRA (BRCRIS) .....</b>	<b>233</b>
Thiago Magela Rodrigues Dias	
Washington Luís R. de Carvalho Segundo	
Marcel Garcia de Souza	
<b>REDE BRASILEIRA DE REPOSITÓRIOS DIGITAIS (RBRD)</b>	<b>249</b>
Juliana Araujo Gomes de Sousa	
Bianca Amaro	
<b>LAGUNA: INFRAESTRUTURA INFORMATACIONAL ABERTA E LAGO DE DADOS .....</b>	<b>263</b>
Patricia da Silva Neubert	
Janinne Barcelos	
Fabio Lorensi do Canto	
Priscila Machado Borges Sena	
<b>SOBRE OS ORGANIZADORES .....</b>	<b>279</b>
<b>SOBRE OS AUTORES .....</b>	<b>281</b>



# PREFÁCIO

Eloy Rodrigues

Diretor dos Serviços de Documentação e Bibliotecas da Universidade do Minho



Desde o final do século XX, o Brasil tem desempenhado um papel de grande relevância na promoção da abertura e democratização do acesso ao conhecimento científico e académico — movimento que, já no início do século XXI, passou a ser conhecido como Acesso Livre ou Acesso Aberto (sendo esta última a designação que veio a prevalecer). Este movimento global foi formalizado em 2002, com a Declaração de Budapeste, e constitui um dos alicerces do conceito mais amplo de Ciência Aberta, que se vem consolidando na última década.

Ao longo de mais de vinte anos, o Brasil, e em particular o Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), assumiu um papel de liderança neste percurso. Este livro, "Infraestruturas de Ciência e Acesso Aberto no Brasil: iniciativas do Instituto Brasileiro de Informação em Ciência e Tecnologia", é um testemunho eloquente dos contributos do Ibict, por via do trabalho dedicado de dezenas dos seus profissionais, para o avanço da Ciência Aberta no Brasil e no mundo.

As contribuições do Ibict abrangeram múltiplas áreas — da sensibilização e capacitação até à cooperação internacional —, mas esta obra centra-se numa dimensão essencial: as infraestruturas e serviços. Os diversos capítulos descrevem as principais iniciativas nesse domínio, permitindo reconhecer um percurso orientado por um princípio central: o uso de tecnologias e protocolos abertos para facilitar a interoperabilidade e a comunicação eficiente entre sistemas.

O livro inicia-se com um capítulo sobre uma das iniciativas pioneiras, a **Biblioteca Digital Brasileira de Teses e Dissertações (BDTD)**, lançada em 2002, que constitui um marco no acesso, disseminação e preservação e acesso a trabalhos acadêmicos produzidos no Brasil. Seguem-se quinze capítulos descrevendo outras tantas infraestruturas desenvolvidas pelo Ibict nas duas últimas décadas.

Desde logo, o **Oasisbr**, inaugurado em 2006 como “Portal Brasileiro de Repositórios e Periódicos de Acesso Livre”, que se

consolidou e alargou na década seguinte, e hoje se designa “Portal brasileiro de publicações e dados científicos em acesso aberto”, reunindo numa plataforma única e integrada, as publicações científicas e dados de pesquisa do Brasil.

A segunda década do presente século foi fértil em infraestruturas lançadas pelo Ibict: em 2011, o **Repositório Institucional do Ibict (RIDi)**, criado para preservar e disseminar o conhecimento científico produzido pelo corpo técnico do Instituto, servindo simultaneamente como instrumento estratégico de promoção do acesso aberto e como memória institucional do conhecimento gerado no Ibict; em 2012, o **Diretório de políticas editoriais das revistas científicas Brasileiras (Diadorim)**, criado com o objetivo de auxiliar pesquisadores, editores e gestores de repositórios institucionais a compreender as políticas de acesso aberto das revistas científicas brasileiras, fornecendo informações detalhadas sobre as condições de armazenamento e acesso de artigos de revistas científicas brasileiras em repositórios digitais; em 2013, a **Rede Cariniana**, criada para viabilizar a preservação digital de acervos científicos e culturais no Brasil, utilizando tecnologias avançadas e práticas colaborativas para assegurar que documentos valiosos sejam preservados e estejam disponíveis para as gerações futuras; em 2015, o **Repositório Comum do Brasil (Deposita)**, criado para armazenar e partilhar as produções científicas e académicas de instituições que não possuam repositórios próprios, permitindo que pesquisadores e instituições de todo o país possam arquivar e divulgar os seus trabalhos.

Nos últimos anos, já na década de 2020, o trabalho do Ibict na criação de infraestruturas continuou intenso. Em 2022, foram disponibilizadas três novas plataformas que ampliaram o escopo de atuação do Ibict para outras importantes áreas da Ciência Aberta: a Plataforma para compartilhamento de códigos-fonte para pesquisa (**Rede Moara**) para fomentar a

reprodutibilidade e a colaboração na pesquisa científica ao disponibilizar um ambiente seguro e acessível para o compartilhamento de códigos-fonte, facilitando a verificação de resultados, a replicação de experiências; o **Cadastro Nacional de Planos de Gestão de Dados (PGD-BR)**, criado para responder à necessidade de desenvolver Planos de Gestão de Dados, uma vez que diversas agências de fomento no Brasil têm exigido a elaboração de Planos de Gestão de Dados, oferecendo uma plataforma intuitiva e completa para que os pesquisadores elaborem seus planos; a **Plataforma de Ciência Cidadã (CIVIS)**, uma ferramenta para o desenvolvimento de projetos de ciência cidadã, oferecendo diversas funcionalidades, e que, ao conectar pesquisadores e cidadãos em torno de projetos colaborativos, democratiza o acesso ao conhecimento científico e estimula a participação da sociedade na solução de problemas complexos.

O ano de 2022 assinala também a realização do I Encontro da Rede Brasileira de Repositórios Digitais: Repositórios Digitais e Ciência Aberta. A **Rede Brasileira de Repositórios Digitais (RBRD)**, que é descrita num dos últimos capítulos do livro, é uma iniciativa colaborativa, agregando cinco sub-redes regionais (Norte, Nordeste, Centro-Oeste, Sudeste e Sul), visando fortalecer a cooperação entre instituições, facilitar a partilha de experiências e boas práticas, para promover o acesso aberto ao conhecimento científico.

O ano seguinte, 2023, foi ainda mais frutífero no lançamento de infraestruturas e serviços pelo Ibict em diferentes áreas. Na área da gestão e partilha de dados, foram apresentados o Aleia e o Deposita Dados. O **Aleia, Repositório de dados de pesquisa do Ibict**, foi criado para armazenar, preservar e disponibilizar conjuntos de dados de pesquisa produzidos por pesquisadores do Instituto. Por sua vez, o **Repositório Comum de Dados de Pesquisa (Deposita Dados)**, também lançado em 2023, foi concebido como um espaço virtual destinado ao armazenamento e compartilhamento de conjuntos

de dados de pesquisa de pesquisadores vinculados a instituições que ainda não possuem os seus próprios repositórios de dados.

Na área da gestão de informação de ciência e tecnologia, foi lançado o ***Ecossistema de Informação da Pesquisa Científica Brasileira (BrCris)***, uma plataforma que integra diversas fontes de dados sobre a pesquisa científica brasileira, proporcionando uma visão mais completa e detalhada do ecossistema científico nacional. Ao conectar informações sobre publicações, pesquisadores, instituições e projetos, o BrCris permite identificar tendências, mapear colaborações e avaliar o impacto da pesquisa em diferentes áreas do conhecimento.

Finalmente, na área das revistas científicas foram apresentados o ***Portal brasileiro para as revistas científicas (Manuelzão)*** e o ***Diretório das revistas científicas eletrônicas brasileiras (Miguilim)***, que formam um conjunto de ferramentas complementares desenvolvidas para dar resposta às necessidades e capacitar a comunidade científica brasileira em boas práticas de comunicação científica. O Manuelzão, disponibiliza um conjunto de alargado de informações e funcionalidades para auxiliar os editores na criação, gestão e melhoria das suas revistas. Por seu lado, o Miguilim, reúne informações detalhadas sobre as revistas científicas editadas e publicadas no Brasil, oferecendo um panorama completo do cenário das revistas científicas brasileiras.

O livro encerra com uma das últimas iniciativas promovidas pelo Ibitc: o ***Projeto Laguna***, ainda em desenvolvimento, visando criar um “lago de dados” e uma Infraestrutura Informacional Aberta para proporcionar um ambiente seguro e eficiente para a criação de repositórios de dados certificados, facilitando a análise, o compartilhamento e a reutilização de informações.

Em suma, este livro documenta o notável percurso do Ibitc na construção de um ecossistema robusto para a Ciência Aberta no Brasil. Desde a criação da BDTD até os projetos mais

recentes, o Ibict tem desempenhado um papel crucial na promoção do Acesso Aberto e da Ciência Aberta, não só no Brasil, mas também a nível internacional.

Permito-me concluir com uma nota pessoal. Tenho acompanhado de perto esta caminhada do Ibict, com o qual o meu próprio percurso pessoal e institucional, na Universidade do Minho, se tem cruzado nas últimas duas décadas. Neste período tive o privilégio de conhecer e trabalhar com vários profissionais de excelência do Ibict, muitos dos quais mais que colegas se tornaram amigos. E quero destacar duas dessas pessoas.

A primeira é o Hélio Kuramoto, que conheci em Guimarães em junho de 2003, na conferência Elpub, durante um Workshop sobre Teses e Dissertações Digitais no qual ambos fomos oradores (ele apresentando já a BDTD, e eu o trabalho que estávamos a desenvolver para a criação do RepositóriUM, que seria lançado alguns meses depois). Nos anos seguintes tive oportunidade de acompanhar o notável trabalho que desenvolveu na criação do Oasisbr, no apoio à criação de repositórios no Brasil, nas tentativas para o estabelecimento de políticas de acesso aberto, e na cooperação internacional, nomeadamente na operacionalização do “Memorando de Entendimento entre o Ministério da Ciência e Tecnologia de Brasil e Portugal para a promoção do acesso aberto às publicações científicas” e na realização da 1<sup>a</sup> ConfOA, originalmente “Conferência luso-brasileira sobre acesso aberto”.

A segunda é a Bianca Amaro, que conheci em 2010, no Ibict em Brasília, na reunião para iniciar o processo de implementação do Memorando de Entendimento e de preparação da ConfOA. Não saberia enumerar e fazer justiça a todas as importantes contribuições pelas quais a Bianca Amaro, atuando pelo Ibict, foi responsável nos últimos quinze anos. A sua visão, liderança, determinação e energia são certamente reconhecidos por todos os que vêm trabalhando pelo Acesso Aberto e a Ciência Aberta não apenas no Brasil e na América

Latina, mas em todo o mundo. Da minha parte quero apenas destacar o trabalho conjunto que permitiu consolidar a ConfOA como o evento de referência no espaço lusófono, e agradecer a sorte e o privilégio de termos partilhado algo que foi e é mais que um trabalho, constituindo-se como uma missão das nossas vidas: a promoção do Acesso Aberto e da Ciência Aberta.

Estou certo de que o rico percurso do Ibict, e o exemplo dos que o construíram até agora, serão fonte de inspiração para que o Instituto, com novos protagonistas, renove o compromisso com a transparência, a democratização e o acesso à informação como bem público e comum.

E para que continue a ser uma referência e um exemplo, no Brasil e no mundo, no prosseguimento do caminho rumo a uma ciência mais responsável, inclusiva, transparente e reproduzível, cujos resultados, acessíveis e abertos, beneficiem o conjunto da sociedade.

Guimarães, julho de 2025.



# APRESENTAÇÃO

Bianca Amaro  
Phillipe de Freitas Campos  
Janinne Barcelos



A partir do final da década de 1990, a comunicação científica mundial começou a passar por mudanças disruptivas, impulsionadas por dois fatores principais: o surgimento de novas tecnologias de comunicação e informação e o Movimento de Acesso Aberto à Informação Científica. Juntos, esses fatores transformaram a maneira de produzir e disseminar ciência, com impacto em todas as áreas do conhecimento.

Nesse cenário, o Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), unidade de pesquisa vinculada ao Ministério da Ciência, Tecnologia e Inovação (MCTI), reconheceu a necessidade de dotar o Brasil com infraestruturas nacionais alinhadas aos novos paradigmas da comunidade científica global. Tal iniciativa encontra-se ancorada na própria missão institucional do Instituto: “Promover a competência, o desenvolvimento de recursos e a infraestrutura de informação em ciência e tecnologia para a produção, socialização e integração do conhecimento científico e tecnológico”.

Guiados por uma visão pioneira e cientes das transformações em curso na comunicação científica mundial, os pesquisadores Hélio Kuramoto e Silvia Barcelos desempenharam um papel central ao trazer para o Ibict os fundamentos do Movimento de Acesso Aberto. Em uma visita técnica à Alemanha, Kuramoto e Barcelos tiveram contato com a tecnologia dos arquivos abertos (*open archives*) e, ao retornarem, introduziram no Instituto discussões que estavam ganhando força globalmente. Esse marco foi o ponto de partida para as infraestruturas de acesso aberto desenvolvidas pelo Ibict, detalhadas nas páginas seguintes.

Embora esta obra tenha como foco o resgate histórico das principais iniciativas do Ibict, é inegável que ações nacionais e internacionais desempenharam um papel fundamental na consolidação do Acesso Aberto e da Ciência Aberta. Por essa razão, o livro se inicia com uma linha do tempo que destaca os principais acontecimentos externos ao Ibict, mas que foram essenciais para a construção das infraestruturas existentes no Instituto.

Em 2001, o Movimento de Acesso Aberto consolidou-se de maneira significativa por meio da *Budapest Open Access Initiative* (BOAI). Ao qualificar a informação científica como bem público, a BOAI advogou pela eliminação das barreiras de acesso que até então vigoravam, visando sua democratização. Para alcançar esse objetivo, a Iniciativa preconizou a Via Dourada (publicação em revistas científicas de acesso aberto) e Via Verde (criação de repositórios institucionais).

Diante dessas prerrogativas internacionais, o Ibict iniciou a criação de seus produtos e serviços voltados à estruturação das duas Vias do Acesso Aberto no Brasil. Internamente foram criadas infraestruturas nacionais para atender instituições de ensino e pesquisa, além da oferta de oficinas e treinamentos que capacitaram as instituições a implementarem suas próprias infraestruturas.

Este livro tem como objetivo resgatar e registrar historicamente essas ações, produtos e serviços criados e mantidos pelo Instituto para consolidar o Acesso Aberto e a Ciência Aberta no Brasil. A pesquisa foi conduzida principalmente por meio de análise documental e consulta direta às políticas, diretrizes e websites relacionados aos produtos e serviços do Instituto. Além disso, depoimentos

de técnicos e pesquisadores envolvidos na construção e manutenção dessas iniciativas - muitos deles também autores de capítulos desta obra - foram fundamentais para a reconstrução dessa trajetória.

Nesse processo de consolidação das infraestruturas e políticas de Acesso Aberto e Ciência Aberta no contexto do Ibict, destaca-se também a importância da Conferência Lusófona de Ciência Aberta (ConfOA), que embora não se configure como uma infraestrutura tecnológica tornou-se um espaço estratégico de articulação internacional e intercâmbio de experiências. Em parceria com as instituições portuguesas Universidade do Minho (Uminho) e Fundação para a Ciência e a Tecnologia (FCT), há 16 anos o Ibict tem desempenhado papel ativo na organização e participação da ConfOA, contribuindo para o alinhamento das iniciativas brasileiras com práticas adotadas em outros países lusófonos.

A história da ciência é marcada por eventos que definem seus rumos, e a trajetória do Ibict não é exceção. Como disse Isaac Newton: “Se enxerguei mais longe, foi por estar sobre os ombros de gigantes”. Essa frase emblemática reflete o espírito desta obra, que reconhece o legado daqueles que contribuíram para o avanço da comunicação científica no Brasil.

Ao longo de seus 70 anos, o Ibict foi enriquecido por novos nomes, ideias e tecnologias, cujos impactos são apresentados neste livro. Convidamos o leitor a conhecer e valorizar os esforços do Ibict na promoção do Acesso Aberto e da Ciência Aberta no Brasil. Que estas páginas inspirem a todos a participar dessa transformação

essencial para o avanço do conhecimento científico em nosso país.

**Boa leitura!**



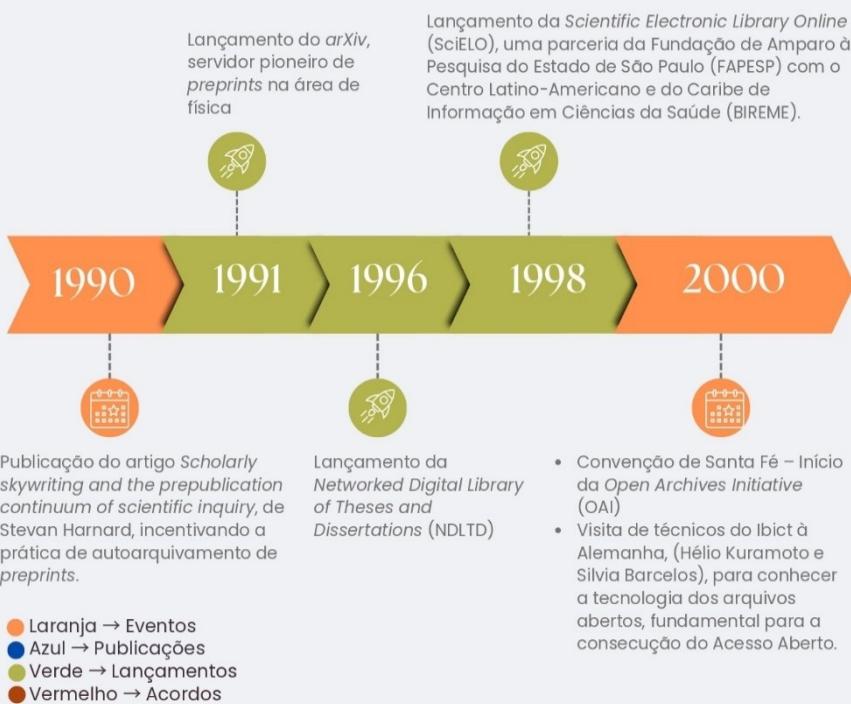
# LINHA DO TEMPO



# Linha do Tempo

O desenvolvimento das atuais infraestruturas de Acesso Aberto e de Ciência Aberta no Brasil foi influenciado por declarações, acordos e iniciativas internacionais que, ao longo do tempo, criaram condições favoráveis para sua implementação. Nesse cenário, o Ibict e outras instituições pioneiras desempenharam um papel central na consolidação de produtos e serviços que hoje se tornaram referência no país e no exterior.

A linha do tempo apresentada a seguir reúne marcos que ajudam a compreender esse percurso, destacando a articulação de diferentes esforços que resultaram no ecossistema de informação científica atualmente disponível no Brasil. Não se trata de uma compilação exaustiva: alguns eventos e iniciativas podem não estar contemplados, mas buscou-se reunir marcos representativos que auxiliam na compreensão dessa trajetória.



# Linha do Tempo

- Laranja → Eventos
- Azul → Publicações
- Verde → Lançamentos
- Vermelho → Acordos

Publicação da *Budapest Open Access Initiative*, fruto da reunião sediada pelo então *Open Society Institute*, defendendo a disponibilidade da literatura científica "sem expectativa de pagamento"

- Publicação da "Declaração de Bethesda sobre publicação em acesso aberto"
- Publicação da "Declaração de Berlim sobre o acesso aberto ao conhecimento nas ciências e humanidades",

Fundação da *Open Knowledge Foundation*, organização não governamental voltada à promoção do acesso aberto ao conhecimento

Publicação da Declaração de Florianópolis, assinada por pesquisadores da área de psicologia, em apoio ao movimento mundial pelo acesso aberto.



2002

2002

2003

2003

2004

2005

2006



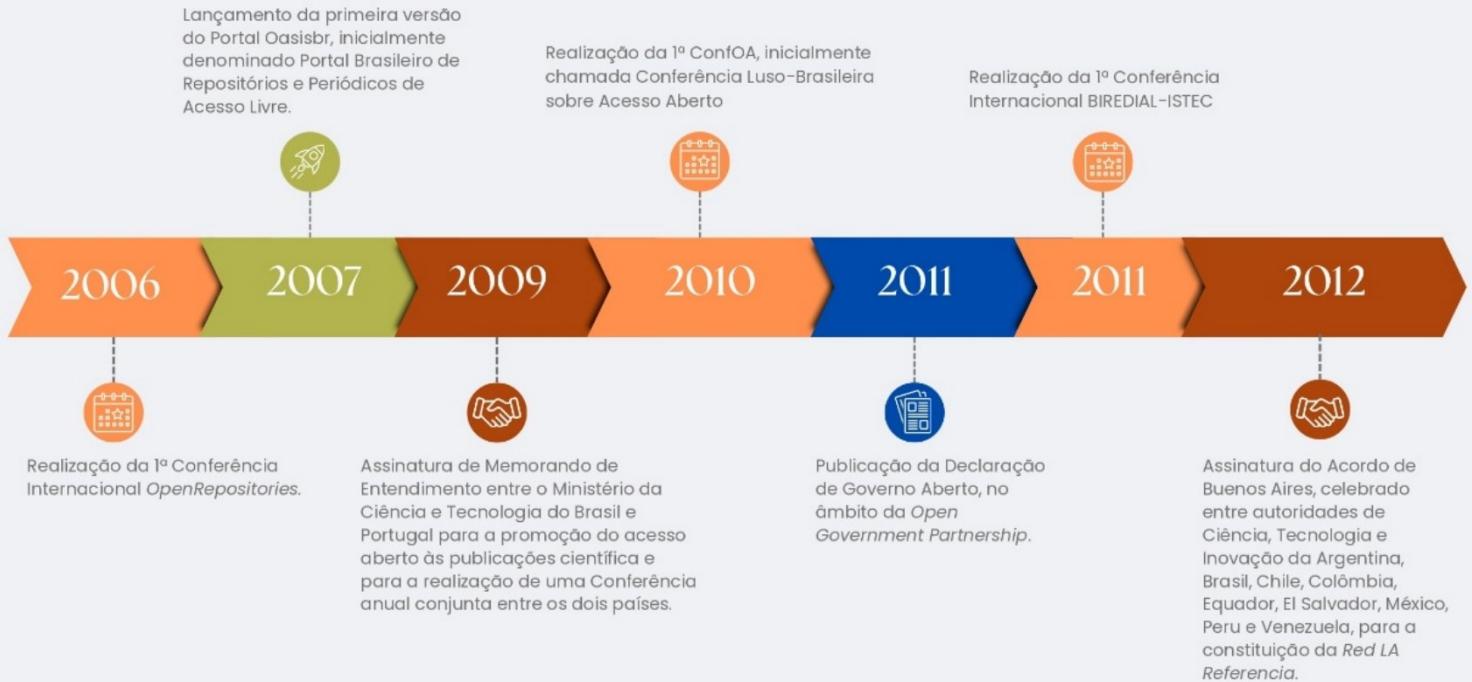
Lançamento oficial da primeira versão da Biblioteca Digital Brasileira de Teses e Dissertações (BDTD). Criação do Comitê Técnico-Consultivo (CTC) e desenvolvimento do Padrão Brasileiro de Metadados para Teses e Dissertações (MTD-BR).

Lançamento do *Directory of Open Access Journals* (DOAJ), pela Universidade de Lund

- Publicação, pelo Ibict, do Manifesto Brasileiro de Apoio ao Acesso Aberto à Informação Científica.
- Publicação da Carta de São Paulo, assinada por professores, pesquisadores, bibliotecários e representantes da sociedade civil em defesa do acesso aberto.
- Publicação da Declaração de Salvador sobre Acesso Aberto: a perspectiva dos países em desenvolvimento.

# Linha do Tempo

- Laranja → Eventos
- Azul → Publicações
- Verde → Lançamentos
- Vermelho → Acordos



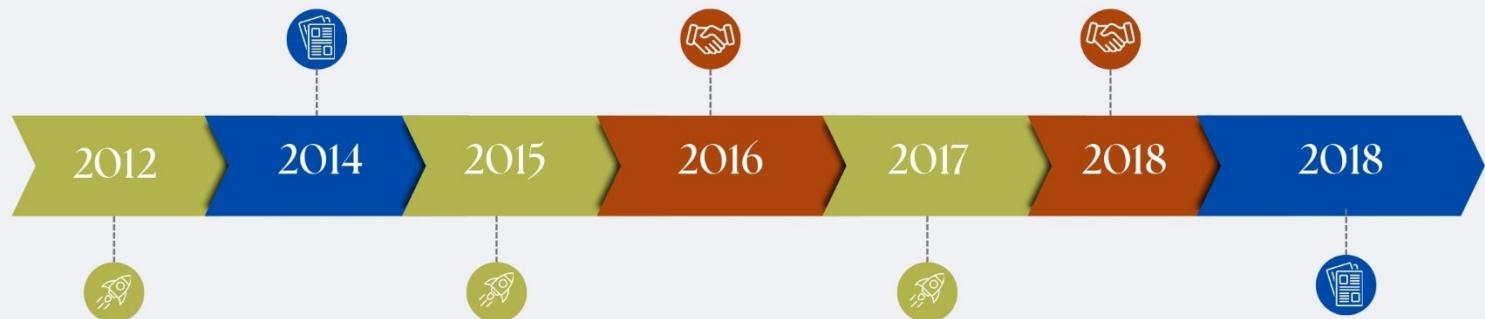
# Linha do Tempo

- Laranja → Eventos
- Azul → Publicações
- Verde → Lançamentos
- Vermelho → Acordos

- Publicação da Carta de Belém, que definiu como estratégia central para a operacionalização do acesso aberto o depósito da produção científica em repositórios institucionais
- Publicação da Portaria nº 043 de 10 de setembro, oficializando a política do Repositório Institucional do Ibiti (RIDi)

Assinatura e lançamento do Manifesto de Acesso Aberto a Dados da Pesquisa Brasileira para Ciência Cidadã

Assinatura do segundo Memorando de Entendimento entre o Ministério da Ciéncia e Tecnologia de Brasil e Portugal, agora já em matéria de Ciéncia Aberta, ratificando a colaboração dos dois países e de suas infraestruturas



- Lançamento do Diadorim, então Diretório de políticas de acesso aberto das revistas científicas brasileiras
- Lançamento da Red LA Referencia, então Red Federada Latinoamericana de Re却ritórios de Documentación Científica

Lançamento de nova versão do Portal Oasisbr, agora denominado Portal de Publicações Científicas em Acesso Aberto

Lançamento do Re却ritório Comum do Brasil (Deposita).

Publicação da Declaraciéon de Panamá sobre Ciéncia Abierta.

# Linha do Tempo

● Laranja → Eventos  
● Azul → Publicações  
● Verde → Lançamentos  
● Vermelho → Acordos



# BIBLIOTECA DIGITAL BRASILEIRA DE TESES E DISSERTAÇÕES (BDTD)

Tainá Batista de Assis

Juliana Araujo Gomes de Sousa

Cássio Teixeira de Moraes



## APRESENTAÇÃO

As teses e dissertações são caracterizadas por serem publicações científicas extensas que formalizam o processo para a obtenção de graus acadêmicos, como doutorado e mestrado, respectivamente, em todo o mundo. Esses documentos são resultantes de pesquisas originais conduzidas por estudantes sob uma orientação e representam uma contribuição significativa para o corpo de conhecimento em diversas disciplinas.

Em uma perspectiva ampliada, pode-se citar o papel crucial que essas produções científicas desempenham para além da outorga dos graus: (i) exploram novos territórios de conhecimento; (ii) contribuem diretamente para a descoberta de inovações; (iii) promovem o desenvolvimento de habilidades críticas da pesquisa, como a análise, a escrita acadêmica e a lógica argumentativa; (iv) contribuem de forma direta para o avanço do conhecimento, uma vez que essas produções podem resultar em outras publicações científicas (ex.: artigos científicos, capítulos de livros, trabalhos apresentados em eventos científicos, patentes e desenvolvimentos tecnológicos); (v) formam especialistas em diversas áreas do conhecimento, fortalecendo o meio acadêmico; (vi) proporcionam o aumento da visibilidade e da credibilidade das instituições científicas.

Segundo Campello (2000), a apresentação das teses e dissertações como produtos científicos para a conferência de graus acadêmicos remonta do século XII com as universidades medievais. Com o surgimento da imprensa e a proliferação das universidades, o

desenvolvimento da escrita foi considerado um componente essencial do processo educativo, o que impulsionou também a exposição oral dos estudantes perante a uma banca de examinadores em defesa de suas produções. Moxley (2001, p. 62) apontou que a primeira tese submetida, em 1860, compreendia “*a six-page, hand written thesis at Yale University*” e que, ao longo dos séculos, as teses e dissertações se consolidaram como instrumentos capazes de refletir a qualidade das instituições de ensino e pesquisa nas produções intelectuais de seus alunos. Nesse contexto também, Orera Orera (2003) apresenta uma reflexão sobre a dualidade da função das teses e dissertações frente à disponibilização pública. A autora elenca como primeira delas o entendimento de sua função administrativa para a obtenção do grau acadêmico, como mencionado anteriormente, com sua defesa pública. Por outro lado, a autora considera que essas produções devem ser submetidas também às regras de proteção dos direitos de autor. Tendo em vista esses dois aspectos, considera-se que a disponibilização e o acesso às teses e dissertações constituem em preocupações primordiais à comunidade científica e de seus produtores, de acordo com todas as devidas proteções ao autor, quando o caso.

Assim sendo, para que as teses e dissertações possam alcançar todos os seus benefícios é imprescindível que estejam facilmente acessíveis e não apenas presentes nos acervos de bibliotecas de suas instituições de defesa.

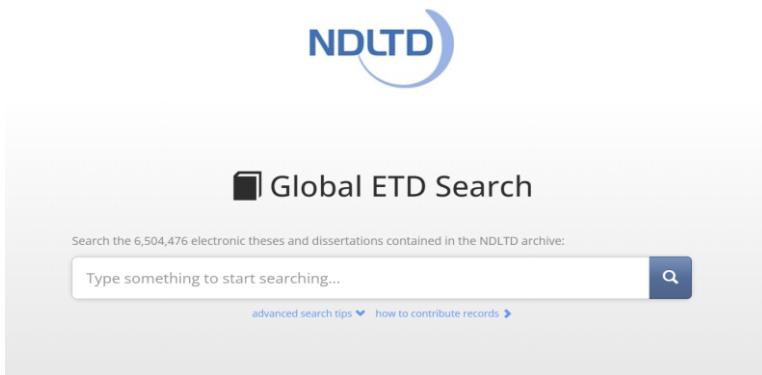
## O MOVIMENTO DE “DIGITALIZAÇÃO” DAS TESES E DISSERTAÇÕES

A evidente importância do acesso às teses e dissertações, alinhada aos avanços das tecnologias e discussões em torno das construções de sistemas para o armazenamento da informação científica, surge o movimento de “digitalização” desses documentos. A literatura passa a referenciar como teses e dissertações eletrônicas (ETD), do inglês, *electronic thesis and dissertations*, os documentos que foram digitalizados (quando seus originais são impressos) e aqueles que já nascem digitais.

O termo ETD foi cunhado em 1987 apresentado pela *Universal Microfilms International* (UMI) (Al Salmi, 2014; Pereira, 2007) que, à época teve seu serviço de gestão de teses e dissertações migrado para a base da ProQuest, mas que também a UMI estabeleceu contatos com a *Virginia Tech* sobre a padronização de metadados e estruturas das ETDs. Dessas tratativas, surge uma das principais iniciativas de gestão, visibilidade e acesso às teses e dissertações: a *Networked Digital Library of Theses and Dissertations* (NDLTD, 2025).

A NDLTD é uma rede que reúne teses e dissertações eletrônicas de diferentes instituições e países. Em números, são mais de 205 instituições participantes e mais de 6.504.470 documentos disseminados, conforme dados coletados em 2024. A Figura 1 a seguir apresenta a tela inicial do sistema de busca da NDLTD que qualquer usuário pode ter acesso.

Figura 1 - Sistema de busca por teses e dissertações da *Networked Digital Library of Theses and Dissertations* (NDLTD)



Fonte: Captura de tela da NDLTD (2024).

Ao observar analiticamente a base da NDLTD tem-se que a maioria das teses e dissertações estão redigidas sob o idioma inglês com 1.681.753 documentos, seguido do idioma chinês com 1.005.641 documentos e, em terceiro, o português com 807.651 teses e dissertações. Esses números são impulsionados em função das fontes de coletas das teses e dissertações que, se ao pensar em um ranking, destacam-se cinco iniciativas, apresentadas assim como estão descritas no sistema da NDLTD:

- *National Digital Library of Theses and Dissertations in Taiwan*
- *OCLC*
- *Ibict Brazilian ETDs*
- *Ethos UK*
- *DiVA Archive at Upsalla University*

Esses dados que compõem a base da NDLTD refletem as informações dos repositórios locais de teses e dissertações. Todavia, um aspecto a ser considerado é a diferença entre o total de documentos contabilizados pela NDLTD e o número de documentos que realmente estão em acesso aberto, ou seja, disponíveis em texto completo. A *National Digital Library of Theses and Dissertations* em Taiwan, por exemplo, registra na NDLTD um total de 1.095.286 teses e dissertações. No entanto, ao final de 2022, os dados extraídos diretamente do sistema de Taiwan indicavam um total de 1.345.991 documentos, dos quais apenas 658.379 tinham texto completo disponível. Para comparação, a BD TD, no mesmo ano, registrava 876.946 teses e dissertações com textos completos, posicionando-se como a principal base nacional de teses e dissertações em termos de quantidade de documentos disponíveis.

Outro aspecto a ser observado é que ocorrem problemas na coleta, resultando em uma discrepância nos números em comparação com a base real/local. Um exemplo disso é o caso da base brasileira, onde a NDLTD registra um total de 626.594 documentos. Entretanto, ao consultar a base do Brasil coordenada pelo Ibict, em 2024, o sistema indica, aproximadamente, 910.315 teses e dissertações, um número significativamente superior ao presente na NDLTD.

Além da NDLTD, outra iniciativa que também objetiva integrar e disseminar as teses e dissertações eletrônicas em todo o mundo é a *Open Access Theses and Dissertations* (OATD, 2025). Segundo informações da própria OATD, o sistema coleta informações das teses e dissertações de mais de 1.100 faculdades, universidades e

instituições de pesquisa, indexando mais de 7.230.360 documentos.

Em uma pesquisa conduzida na *North-West University*, que teve como objetivo analisar os impactos acadêmicos e sociais das teses e dissertações eletrônicas na área da Engenharia com base no número de citações, Bangani (2018) revelou que 612 teses e dissertações identificadas no estudo somaram um total de 931 citações. Além disso, o estudo registrou uma média de 323 visualizações por arquivo PDF. Esse dado ilustra o potencial ampliado pelo uso das tecnologias da informação e comunicação, que permite que as teses e dissertações eletrônicas aumentem seu alcance de maneira exponencial, superando barreiras geográficas e linguísticas, além de promover maior transparência e integridade da informação. A disseminação e o amplo acesso às teses e dissertações por meio de sua disponibilização nos ambientes digitais fortalecem a chamada *eScience* e a Ciência Aberta (Schöpfel *et al.*, 2014). Além da profundidade das discussões que esses documentos concentram, destaca-se o volume de dados de pesquisa que as teses e dissertações tornam visíveis à comunidade científica. Isto porque, é cada vez mais demandado a entrega dos conjuntos de dados científicos de forma integrada ao texto completo da tese ou dissertação.

Na esteira das iniciativas de sistemas que gerenciam teses e dissertações, o Brasil também criou e mantém sua base aberta de teses e dissertações seguindo os preceitos dos estudos da NDLTD. A Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), coordenada pelo Instituto

Brasileiro de Informação em Ciência e Tecnologia (Ibict), é fruto de vários estudos iniciados ainda na década de 1970, aprofundados nos anos de 1990 e que culminaram em seu lançamento no ano de 2002. A próxima seção deste capítulo apresenta a BD TD e a sua evolução técnica e conceitual.

## A BIBLIOTECA DIGITAL BRASILEIRA DE TESES E DISSERTAÇÕES (BD TD)

Como dito, lançada em 2002, a Biblioteca Digital Brasileira de Teses de Dissertações (BD TD) consiste em uma rede que integra em seu sistema mais de 150 instituições de todo o Brasil por meio da coleta de teses e dissertações organizadas em suas bibliotecas digitais e/ou repositórios institucionais de acesso aberto.

O acesso a essa produção científica é livre de quaisquer custos. Além disso, a BD TD também coleta as teses e dissertações defendidas por brasileiros em instituições de ensino estrangeiras. Neste caso, a coleta das teses e dissertações é realizada por meio do repositório comum do Brasil (Deposita) (Ibict, 2025a), em que o próprio pesquisador realiza o autodepósito de sua produção.

A Figura 2 apresenta a interface da página da Biblioteca Digital Brasileira de Teses e Dissertações (BD TD).

Figura 2 - Interface da Biblioteca Digital Brasileira de Teses e Dissertações (BDTD)



Fonte: Captura de tela da BDTD (2025).

Em 2006, Southwick já destacava o kit de ferramentas disponibilizado pelo Ibict às instituições interessadas em participar da BDTD. Segundo a autora (2006, tradução nossa), esse kit foi elaborado com o “objetivo de dar apoio técnico às organizações que o necessitam e como um meio para difundir métodos e padrões para a criação de bibliotecas digitais locais de ETD”. O kit de ferramentas referenciado por Southwick (2006) compreendia: (1) disponibilização de software para a gestão e disponibilização de ETD desenvolvido pelo Ibict; (2) programa para implementar o protocolo OAI para a coleta de dados; (3) capacitação dos profissionais; (4) equipamento (hardware). A disponibilização desse kit foi formalizado por intermédio do edital Funcate/Ibict nº

001/2005, o qual contemplou 13 instituições<sup>2</sup>. Para participar do edital as instituições deveriam ser públicas e ter programas de pós-graduação reconhecidos pelo Ministério da Educação (MEC).

Ainda no escopo de seu lançamento, a BD TD inovou também ao apresentar o Padrão Brasileiro de Metadados para Teses e Dissertações (MTD-BR), com base no padrão internacional descritivo de teses e dissertações proposto pela NDLTD, o chamado *Interoperability Metadata Standard for Theses and Dissertations* (ETD-MS) que é correspondente a uma extensão do padrão de metadados *Dublin Core*. A estruturação de um padrão brasileiro se faz necessário em razão das especificidades das estruturas organizacionais das instituições no Brasil que ofertam programas de pós-graduação *stricto sensu*. Alinhado a isso, o padrão de metadados proposto apresenta-se como um rico mecanismo capaz de coletar e organizar dados referentes à produção de teses e dissertações e fazer da BD TD uma base diferenciada das demais de mesma natureza. O Ibict justifica, assim, a proposição do padrão de metadados:

---

<sup>2</sup> Universidade estadual do Rio de Janeiro (UERJ); Universidade Federal do Espírito Santo (UFES); Universidade Federal de Goiás (UFG); Universidade Federal de Lavras (UFLA); Universidade Federal de Mato Grosso do Sul (UFMS); Universidade Federal do Pará (UFPA); Universidade Federal da Paraíba (UFPB); Universidade Federal de Pernambuco (UFPE); Universidade Federal do Rio Grande do Norte (UFRN); Universidade Federal de Santa Maria (UFSM); Universidade de Brasília (UnB); Universidade de Taubaté (Unitau); Universidade Federal de Viçosa (UFV).

[...] foi desenvolvido o Padrão Brasileiro de Metadados de Teses e Dissertações (MTD-BR) para atender aos objetivos da BDTD de oferta de produtos e serviços de informação com vistas à identificação e localização das teses e dissertações eletrônicas (TDEs), como também para a coleta de informação com vistas à geração de indicadores e integração com outros repositórios nacionais e internacionais (Ibict, 2003, web).

A partir da compreensão do histórico e do funcionamento da BDTD que serão apresentados a seguir, é possível observar que a estrutura pensada na constituição da BDTD a figura como uma base de alto valor e importância para a comunidade científica e sociedade em geral.

## BREVE HISTÓRICO DA BDTD

Antecedente à criação da BDTD, na década de 1980, o Ibict possuía acordo de cooperação com a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) e o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) para receber e manter em seu acervo uma cópia impressa da tese ou dissertação produzida no Brasil e/ou defendida por brasileiro no exterior. Essa ação culminou na formação de mais de 20 mil documentos que impulsionou a necessidade no Ibict de criação de uma base de dados para organizar e possibilitar a recuperação dessas teses e dissertações. A esse acervo, ao final dos anos de 1980, somaram os documentos acondicionados nas bibliotecas da Universidade Federal do Rio Grande do Sul (UFRGS), da Universidade de São Paulo (USP) e,

posteriormente, da Universidade Federal de São Carlos (UFSCar) e a Universidade Federal de Goiás (UFG).

Em 1996, assim, é apresentado pelo Ibict o projeto-piloto do Sistema de Informação sobre Teses e Dissertações (SITE). Segundo os autores Marcondes e Sayão (2003),

o Ibict chamou instituições brasileiras que tinham, até o momento, experiências de desenvolvimento de bancos de teses e dissertações eletrônicas, para se articularem no que vem sendo chamado de Consórcio Brasileiro de Teses Eletrônicas, com o objetivo de montar um catálogo coletivo de teses eletrônicas e torná-las disponíveis através de um único portal na internet. As instituições que inicialmente formaram o Consórcio foram a BIREME, que opera como experiência-piloto, o banco de teses eletrônicas da FIOCRUZ, a PUC-RIO, a UFSC, a USP e o CNPq (Marcondes; Sayão, 2003, p. 163).

Se integraram a essa iniciativa, em seu início, 13 instituições que formavam um acervo de 70.000 teses e dissertações. O Ibict era o órgão responsável pelo recebimento dos arquivos em disquetes e CD-ROM, o envio de cópias impressas foi descontinuado.

Com a experiência do SITE, no ano de 2001, o Ibict submeteu proposta de criação da Biblioteca Digital Brasileira de Teses e Dissertações (BDTD) no âmbito do Programa da Biblioteca Digital Brasileira (BDB). Em 2002, para o lançamento da BD TD, foi instaurado em abril, um comitê técnico-consultivo (CTC), constituído por

representantes do Ibict, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Ministério da Educação (MEC) - representado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) e Secretaria de Educação Superior (SESu), Finep e das três universidades que participaram do grupo de trabalho e do projeto-piloto: Universidade de São Paulo (USP), Pontifícia Universidade Católica do Rio de Janeiro (PUC-RIO) e a Universidade Federal de Santa Catarina (UFSC). Dentre as atribuições do grupo, o CTC apoiou o desenvolvimento e aprovou a primeira versão do Padrão Brasileiro de Metadados para Teses e Dissertações (MTD-BR).

A Biblioteca Digital Brasileira de Teses e Dissertações (BDTD) teve, então, as seguintes linhas principais de atuação: (i) estudar experiências existentes no Brasil e no exterior de desenvolvimento de bibliotecas digitais de teses e dissertações; (ii) desenvolver, em cooperação com membros da comunidade, um modelo para o sistema; (iii) definir padrões de metadados e tecnologias a serem utilizadas pelo sistema; (iv) absorver e adaptar as tecnologias a serem utilizadas na implementação do modelo; (v) desenvolver um sistema de publicação eletrônica de teses e dissertações para atender àquelas instituições de ensino e pesquisa que não possuíam sistemas automatizados para implantar suas bibliotecas digitais; (vi) difundir os padrões e tecnologias adotadas e dar assistência técnica aos potenciais parceiros na implantação das mesmas.

As linhas de atuação não se destacam apenas durante a implementação e o lançamento da BD TD, mas

também nos anos seguintes. Nesse contexto, a BDTD passou por várias atualizações, tanto na infraestrutura tecnológica quanto nas diretrizes informacionais, que abordam a padronização, a coleta e a exposição dos metadados no Portal. Destaca-se a atualização do MTD-BR que em sua primeira versão, que foi apresentada em 2002, era composto de 71 metadados, em que 38 eram de preenchimento obrigatório e 33 de preenchimento opcional. Em 2005 foi lançada a segunda versão do padrão de metadados, o qual foi denominado de MTD2-BR, que contava com 80 metadados dos quais 31 eram obrigatórios e 49 eram opcionais. O padrão de metadados atual é o MTD3-BR que começou a ser idealizado em meados de 2014 e teve seu lançamento em 2017. O MTD3-BR é composto por 54 metadados, em que 37 são de preenchimento opcional e 17 de preenchimento obrigatório. A atualização do MTD-BR envolveu não apenas a revisão da nomenclatura e da forma de preenchimento dos metadados, mas também uma análise estratégica para definir o número de metadados obrigatórios.

A disseminação e uso do MTD-BR por meio das instituições contou com o respaldo do desenvolvimento do Sistema de Publicação Eletrônica de Teses e Dissertações (TEDE) (Ibict, 2025b), que é uma ferramenta desenvolvida pelo Ibict no âmbito do projeto da BDTD e que foi repassado às instituições brasileiras que não tinham sistemas de informação para o armazenamento de suas teses e dissertações (Carvalho Segundo *et al.*, 2014). Em sua primeira versão, o TEDE foi desenvolvido com a programação PHP combinada com o banco de dados MySQL. A eficácia do TEDE no atendimento de sua proposta

teve um impacto positivo tanto no cenário nacional quanto internacional que foi identificado por Segundo Benítez de Vendrell (2011) ao detectarem o uso do TEDE pela *Universidad Nacional del Sur*, na Argentina.

Com os avanços tecnológicos ao longo dos anos percebeu-se a necessidade de atualizar o sistema. O processo de atualização teve início em 2012 e foi definido que o TEDE passaria a utilizar a versão 4.2 do software livre DSpace. O lançamento oficial do TEDE 2 aconteceu em 2014. As diferenças visuais entre as duas versões do TEDE podem ser vistas nas Figuras 3 e 4.

Figura 3 - Primeira versão do software TEDE distribuído pelo Ibitc



Fonte: Captura de tela (2025).

E a nova versão do TEDE (Figura 4).

Figura 4 - Segunda versão do TEDE, com a utilização do DSpace



Fonte: Carvalho Segundo et al. (2014).

Ao longo dos mais de 20 anos de atividade da BDTD, o Instituto tem firmado seu compromisso com a inovação tecnológica e a eficiência na atualização, representação e disseminação da produção científica proveniente dos programas de pós-graduação *stricto sensu*.

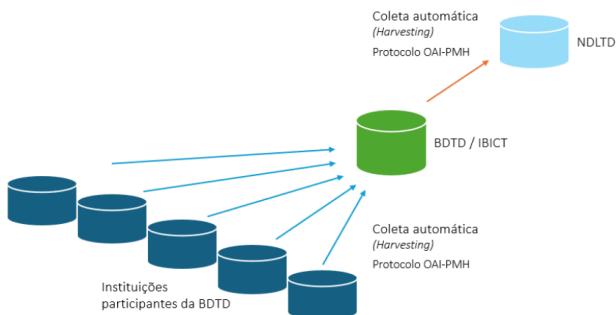
## **COMO FUNCIONA A BDTD?**

Desde o seu lançamento, a BDTD preconiza o modelo de Arquivos Abertos (*Open Archives*), o qual foi proposto na Convenção de Santa Fé (1999). A proposta da Convenção preconiza que os documentos digitais além de estar de acordo com os preceitos do Movimento do Acesso Aberto, também fossem gerenciados por softwares que permitissem a exposição dos dados daquele recurso por meio de uma interface de interoperabilidade que permita,

caso necessário, o intercâmbio de informações entre sistemas distintos.

Deste modo, o funcionamento da rede advém das tecnologias de agregação de bases de dados por meio da interoperabilidade entre sistemas. Nesse sentido, a BDTD desempenha o papel de um provedor de serviços enquanto que as bases de dados coletadas são chamadas de provedores de dados. A interoperabilidade entre a BDTD e as bases de dados locais decorrem por meio do protocolo de interoperabilidade *Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH). A Figura 6 apresenta o processo de coleta e apresentação dos metadados.

Figura 6 - Processo de coleta da BDTD e integração com a NDLD



Fonte: Ibict (2014), com adaptações.

Para que a coleta seja realizada, é necessário que a instituição cumpra com critérios de elegibilidade e de interoperabilidade entre metadados. Para que uma instituição participe da BDTD é necessário que a mesma possua programas de pós-graduação *stricto sensu* e disponibilize as teses e dissertações por meio de sistema que faça uso do protocolo de interoperabilidade OAI-PMH.

Além disso, a instituição deve fornecer, imprescindivelmente, os seguintes metadados: (i) Autor(a); (ii) Data da defesa; (iii) Orientador(a); (iv) Tipo de acesso; (v) Tipo de documento; (vi) URL persistente.

Nessa operacionalização da BDTD, dois softwares são fundamentais para a coleta, tratamento e apresentação dos dados: o coletador, um software desenvolvido pela Rede LA Referencia (2025), desenvolvido em linguagem Java, sob o *framework Spring* e o banco de dados PostgreSQL (Sousa et al., 2023), encarregado coletar, transformar e validar os dados provenientes das fontes de coleta da BDTD e o software livre VuFind, uma interface web que permite aos usuários visualizar, de forma agregada, todas as informações coletadas e validadas pelo software da plataforma LA Referencia.

Se adequar aos princípios obrigatórios de interoperabilidade requisitados pela BDTD permite a integração com a NDLTD e o Portal brasileiro de publicações e dados científicos em acesso aberto (Oasisbr), ampliando a visibilidade das teses e dissertações produzidas nas instituições de ensino e pesquisa brasileiras e por pesquisadores brasileiros.

## **CONSIDERAÇÕES FINAIS**

As teses e dissertações são documentos produzidos no contexto dos programas de pós-graduação que tem como objetivo principal conferir o título de doutor ou mestre. No entanto, seu impacto vai além da certificação acadêmica, desempenhando um papel essencial na disseminação do conhecimento e no avanço das ciências,

tanto sociais quanto naturais. Para maximizar os benefícios associados à publicização do conhecimento, é crucial adotar técnicas de disseminação, compartilhamento e acesso.

Apoiado pelas tecnologias de informação, o processo de disseminação das teses e dissertações em formato eletrônico demonstrou um potencial transformador na promoção do acesso aberto à informação científica. Sob essa perspectiva, a BD TD se destaca como uma iniciativa de referência no contexto de disseminação e visibilidade das teses e dissertações, contribuindo para a visibilidade do pesquisador brasileiro e das instituições de ensino e pesquisa nacionais.

Todavia, a reflexão presente também abrange as responsabilidades das instituições que participam da BD TD, uma vez que são essenciais para fornecer dados confiáveis e íntegros referentes às suas teses e dissertações. Essa responsabilidade compartilhada com o Ibict garante as possibilidades de estudos, análises e geração de subprodutos ou subserviços a partir do uso das informações presentes na base da BD TD.

## REFERÊNCIAS

- AL SALMI, Jamal Mattar Yousuf. **Electronic theses and dissertations programmes in the Arab Gulf States:** exploring factors affecting their adoption and development. 2014. 316 f. Tese (Doctor of Philosophy in Information Systems) - University of Wellington, Wellington, 2014. DOI: 10.26686/wgtn.17006701.v1.

BANGANI, Siviwe. The impact of electronic theses and dissertations: a study of the institutional repository of a university in South Africa. **Scientometrics**, Budapest, v. 115, n. 1, p. 131-151, 2018. DOI: 10.1007/s11192-018-2657-2.

BENÍTREZ DE VENDRELL, Belarmina. *Repositorios de tesis: capacidad del Sistema académico NEA para la generación de depósitos digitales de acceso libre*. In: JORNADAS VIRTUALES IBEROAMERICANAS DE CIENCIAS DE LA INFORMACIÓN Y LA DOCUMENTACIÓN. Buenos Aires, 2011.

CAMPELLO, Bernadete Santos. Teses e dissertações. In: CAMPELLO, Bernadete Santos; CENDÓN, Valadares; KREMER, Ette Marguerite (org.). **Fontes de informação para pesquisadores e profissionais**. Belo Horizonte: UFMG, 2000. p. 121-129.

CARVALHO SEGUNDO, Washington Luís Ribeiro de et al. The LA Referencia Software and the Brazilian Portal of Scientific Open Access Publications (oasisbr). In: INTERNATIONAL CONFERENCE ON OPEN REPOSITORIES, 12., 2017, Brisbane. *Anais* [...]. Brisbane: [s. n.], 2017.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Deposita**: Repositório comum do Brasil. Brasília: Ibict, 2025a. Disponível em: <https://deposita.ibict.br>. Acesso em: 11 ago. 2025.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Biblioteca Digital Brasileira de Teses e Dissertações (BDTD)**. Brasília: Ibict, 2003. Disponível em: [bdtd.ibict.br](http://bdtd.ibict.br). Acesso em: 21 fev. 2024.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Biblioteca Digital Brasileira de Teses e Dissertações (BDTD)**: revisão, atualização e modernização. Brasília: Ibict, 2014. (Apresentação em slide).

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Biblioteca Digital Brasileira de Teses de Dissertações (BDTD)**. Brasília: Ibitc, 2024. Disponível em: <https://bdtd.ibict.br/vufind>. Acesso em: 06 maio 2024.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Sistema de Publicação Eletrônica de Teses e Dissertações (TEDE)**. Brasília: Ibitc, 2025b. Disponível em: <http://tede.ibict.br>. Acesso em: 11 ago. 2025.

MARCONDES, Carlos Henrique; SAYÃO, Luís Fernando. Documentos digitais e novas formas de cooperação entre sistemas de informação em C&T. **Ciência da Informação**, Brasília, v. 31, n. 3, p. 42-54, set./dez. 2003. DOI: 10.18225/ci.inf.v31i3.947.

MOXLEY, Joseph M. American universities should require Electronic Theses and Dissertations. **Educause Quarterly**, Boulder, n. 3, 2001.

NETWORKED DIGITAL LIBRARY OF THESES AND DISSERTATIONS (NDLTD). **Global ETD Search**. [s.l.], 2024. Disponível em: <https://ndltd.org>. Acesso em: 21 fev. 2024.

NETWORKED DIGITAL LIBRARY OF THESES AND DISSERTATIONS (NDLTD). [s.l.], 2025. Disponível em: <https://ndltd.org>. Acesso em: 21 fev. 2025.

OPEN ACCESS THESES AND DISSERTATIONS (OATD). Disponível em: <https://oatd.org>. Acesso em: 11 ago. 2025.

ORERA ORERA, Luisa. Bibliotecas digitales de tesis doctorales: metodología para su planificación. **Boletín de la Asociación Andaluza de Bibliotecarios**, Málaga, n. 72, p. 55-72, set. 2003.

PEREIRA, Carla de Fátima Pires Saiago. **Teses e dissertações em estudos portugueses**: contributos para a organização de uma biblioteca digital. Dissertação (Mestre em Estudos de

Informação e Bibliotecas Digitais) - Departamento de Ciências e Tecnologias da Informação, Instituto Superior de Ciências do Trabalho e da Empresa, Lisboa, 2007.

REDE LATINO-AMERICANA DE CIÊNCIA ABERTA (LA REFERENCIA). [s.l.], 2025. Disponível em:  
[www.lareferencia.info/pt](http://www.lareferencia.info/pt). Acesso em: 11 ago. 2025.

SCHÖPFEL, Joachim *et al.* Open access to research data in electronic theses and dissertations: an overview. **Library Hi Tech**, Bingley, v. 32, n. 4, 2014. DOI: 10.1108/LHT-06-2014-0058.

SOUSA, Juliana Araujo Gomes de *et al.* Interface do gestor: uma dashboard para gestores de repositórios digitais. In: CONFERENCIA INTERNACIONAL DE BIBLIOTECAS Y REPOSITORIOS DIGITALES (BIREDIAL-ISTEC), 12., 2023. **Anais** [...]. Montevideo: Udelar, 2023.

SOUTHWICK, Silvia Barcellos. A biblioteca digital eletrônica brasileira de teses e dissertações: proporcionando acesso aberto à informação acadêmica. **Ciência da Informação**, Brasília, v. 35, n. 2, p. 103-110, maio/ago. 2006. DOI: 10.1590/S0100-19652006000200011.

### Como citar este capítulo

ASSIS, Tainá Batista de; SOUSA, Juliana Araujo Gomes de; MORAIS, Cássio Teixeira de. Biblioteca Digital Brasileira de Teses e Dissertações (BDTD). In: AMARO, Bianca; CAMPOS, Fhilipe de Freitas; BARCELOS, Janinne. (org.). **Infraestruturas de Ciência e de Acesso Aberto no Brasil**: iniciativas do Instituto Brasileiro de Informação em Ciência e Tecnologia. Brasília, DF: Editora Ibict, 2025. Cap. 1, p. 33-54. DOI: 10.22477/9788570132543.cap1

# PORTAL BRASILEIRO DE PUBLICAÇÕES E DADOS CIENTÍFICOS EM ACESSO ABERTO (OASISBR)

Gabriel Silveira Marques



## APRESENTAÇÃO

Nos últimos anos do século XX e início do XXI, a comunicação científica passou por transformações substanciais motivadas pela expansão das tecnologias de informação e comunicação (TIC) e pelo fortalecimento de redes colaborativas em nível global (Bandeira, 2017). Nessa conjuntura, firmaram-se as bases do que se convencionou chamar de Movimento de Acesso Aberto à Informação Científica (MAA). Esse movimento, de caráter internacional, ganhou impulso decisivo por meio de eventos e declarações que estabeleceram fundamentos e diretrizes para a disseminação livre de resultados de pesquisa. Entre os principais marcos, destacam-se a “Declaração de Santa Fé” (1999), a “Iniciativa de Budapeste para o Acesso Aberto” (2002), a “Declaração de Berlim” (2003) e a “Declaração de Bethesda” (2003), documentos que reforçaram a necessidade de abrir caminhos para a democratização do conhecimento científico (Rios; Lucas; Amorim, 2019; Rodrigues; Gugliotta, 2019).

As discussões internacionais pautadas nessas declarações estimularam países a desenvolverem políticas e ações concretas em favor do acesso aberto (Gomes, 2014). Com a popularização da internet no país entre a década de 1990 e o início de 2000, a demanda por serviços de informação científica mais organizados e acessíveis acelerou, levando agências governamentais e instituições de pesquisa a aderirem aos princípios do MAA. Nesse cenário, o Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) passou a desempenhar um importante papel na implementação de ações de incentivo à criação de

repositórios institucionais, revistas eletrônicas e outras plataformas de livre acesso (Bandeira, 2017).

A formalização do movimento brasileiro intensificou-se a partir de 2005, quando ocorreram diversos eventos e declarações nacionais convergentes ao acesso aberto. São exemplos o “Manifesto Brasileiro de Apoio ao Acesso Aberto à Informação Científica” (Ibict, 2025c), lançado oficialmente pelo Ibict em setembro de 2005, e a “Declaração de Salvador”, assinada durante o IX Congresso Mundial de Informação em Saúde e Bibliotecas (2005), que conclamava governos e instituições de pesquisa a aderirem às políticas de acesso livre. Também em dezembro de 2005, a “Carta de São Paulo” reforçava a importância do livre compartilhamento de artigos, teses, dissertações e demais documentos científicos, mobilizando a comunidade acadêmica na consolidação de práticas abertas de publicação (Ortellado, 2008; Rodrigues; Gugliotta, 2019). Posteriormente, em 2006, a “Declaração de Florianópolis” reiterou a necessidade de capacitação de gestores e profissionais que atuassem em repositórios digitais, enquanto o Manifesto e a Declaração ampliaram o diálogo e a visibilidade do tema (Ibict, 2016; Liinc em revista, 2012).

Nessa mesma conjuntura, diversos acordos e parcerias surgiram para fortalecer a infraestrutura de acesso aberto no Brasil. Em 2009, foi firmado o “Memorando de Entendimento entre Brasil e Portugal”, associando o Ministério da Ciência, Tecnologia e Ensino Superior da República Portuguesa e o Ministério da Ciência e Tecnologia da República Federativa do Brasil, iniciativa que consolidou a cooperação técnica entre o Repositório Científico de Acesso Aberto de Portugal (RCAAP) e o Portal

Brasileiro de Publicações e Dados Científicos em Acesso Aberto (Oasisbr), com o propósito de agregar em um só ambiente os repositórios e periódicos em acesso aberto (Shintaku, 2014).

O Oasisbr, portanto, emerge como uma resposta direta aos anseios do MAA no Brasil e como resultado das articulações conduzidas por órgãos governamentais, pesquisadores e gestores de informação. Sua concepção visava, sobretudo, reunir e dar visibilidade à produção acadêmica nacional, reforçando a convicção de que a informação científica, quando aberta, potencializa o desenvolvimento social, tecnológico e econômico (Bandeira, 2017; Rios; Lucas; Amorim, 2019). Com a adesão de múltiplas instituições e o apoio de diversos manifestos, o país vem consolidando uma cultura de acesso democrático ao conhecimento, oferecendo bases mais sólidas para a promoção da ciência aberta. Nesse sentido, a trajetória do movimento internacional de acesso aberto encontra no Brasil terreno fértil para a implementação de políticas e práticas que beneficiam pesquisadores, estudantes, agências de fomento e a sociedade em geral.

## **IBICT E O OASISBR**

Nesse sentido, o Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) desempenha papel estratégico na consolidação do acesso aberto no Brasil, atuando em múltiplas frentes para viabilizar a democratização do conhecimento científico. Entre suas principais iniciativas, destacam-se ações relacionadas à coordenação de repositórios, ao desenvolvimento de

tecnologias de gestão da informação, à capacitação de profissionais para trabalhar com sistemas de editoração e repositórios digitais, além de parcerias internacionais que ampliam o alcance das práticas de ciência aberta (Bandeira, 2017; Marques et al., 2025).

Assim, o Oasisbr constitui-se em uma das iniciativas centrais do Ibict para agregar e disponibilizar a produção científica nacional (Ibict, 2025a). Lançado em 2006, com apoio da Financiadora de Estudos e Projetos (Finep), o Portal teve como propósito inicial coletar, integrar e dar visibilidade a documentos científicos de diferentes fontes, tais como bibliotecas digitais de teses e dissertações, revistas científicas, repositórios institucionais e repositórios de dados de pesquisa (Marques et al., 2023; Silva; Alencar; Souza, 2014).

Ao longo de sua trajetória, o Oasisbr expandiu suas funcionalidades e consolidou-se como um portal agregador de metadados capaz de:

- Reunir e disseminar múltiplas tipologias documentais (artigos científicos, teses, dissertações, conjuntos de dados, *preprints*, entre outros): O Oasisbr apresenta seu papel de portal agregador ao contemplar diferentes tipologias documentais que vão além dos modelos tradicionais de publicação acadêmica. A presença de artigos científicos, teses, dissertações, relatórios técnicos, material de eventos, *preprints*, dados de pesquisa e outros objetos digitais de conhecimento permite uma visão plural do cenário científico nacional (Marques et al., 2025). Essa diversidade de formatos promove a interdisciplinaridade, favorece a

formação de redes de colaboração entre pesquisadores de distintas áreas e fortalece o movimento de Ciência Aberta, ao facilitar a disponibilização de resultados de pesquisa em diferentes formatos (Silva; Silveira, 2019).

- Fortalecer a visibilidade da produção científica brasileira: Ao agregar e difundir conteúdo proveniente de diferentes fontes nacionais, o Oasisbr promove a ciência brasileira também em âmbito internacional. A sistematização de metadados de suas fontes cadastradas, incrementa o alcance de estudos desenvolvidos em universidades, centros de pesquisa e instituições de fomento, promovendo maior impacto às publicações. Essa estratégia de visibilidade auxilia na consolidação de autores e grupos de pesquisa em agendas globais, ao mesmo tempo em que contribui para a transparência dos resultados científicos, estimulando a adoção de boas práticas de documentação e disponibilização de dados (Carvalho Segundo *et al.*, 2017).
- Integrar-se a redes internacionais: A presença do Oasisbr em redes com a *Red Latinoamericana para la Ciencia Abierta* (LA Referencia), a *Networked Digital Library of Theses and Dissertations* (NDLTD) e a *Open Access Infrastructure for Research in Europe* (OpenAIRE), configura um importante passo para o estabelecimento de padrões comuns de interoperabilidade e compartilhamento de dados. Essa integração amplia a difusão global da pesquisa brasileira por meio da adoção de protocolos

internacionais de metadados e do uso de softwares de coleta que possibilitam a sincronização contínua de registros (Ibict, 2024; Shintaku, 2014). Também favorece a convergência com iniciativas de repositórios em outros países, facilitando processos de localização, indexação e reuso de conteúdos acadêmicos produzidos por pesquisadores do Brasil (Gibbon *et al.*, 2023).

Figura 1 - Fluxo informacional do Oasisbr



Fonte: Ibict (2025).

- Atribuição do identificador persistente dARK: A participação do Oasisbr no projeto dARK relaciona-se à incorporação de novas estratégias de preservação digital e à adoção de tecnologias que asseguram a persistência de identificadores e o acesso de longo prazo a documentos científicos. O

dARK, baseado em redes *blockchain* institucionais, surge como uma alternativa descentralizada para a gestão de identificadores persistentes, reduzindo custos e aumentando a autonomia das instituições brasileiras. Ao envolver o Oasisbr como um dos componentes na criação, resolução e armazenamento desses identificadores, o portal se aprofunda com as diretrizes de interoperabilidade, já que o identificador pode ser integrado à sistemas distintos de identificação, como DOI, ARK e ORCID. Esse movimento alinha-se às tendências internacionais de fortalecimento da governança distribuída em soluções voltadas para a Ciência Aberta, agregando camadas adicionais de segurança, rastreabilidade e confiabilidade (Matas *et al.*, 2024).

Para aprimorar a experiência dos usuários, o Oasisbr adota ferramentas como o Protocolo OAI-PMH (*Open Archives Initiative Protocol for Metadata Harvesting*) e a plataforma VuFind, possibilitando a coleta de metadados e a implementação de um motor de busca responsivo que favorece a recuperação rápida e eficiente de documentos (Marques *et al.*, 2025; Shintaku, 2014). Além disso, diversas atualizações vêm sendo implementadas, tais como a melhoria da indexação pelos buscadores on-line, a exportação de dados para gerenciadores de referências (Zotero) e o acréscimo de novas tipologias documentais, incluindo patentes e *preprints*, o que consolida o Oasisbr como um dos maiores portais mundiais de acesso aberto (Ibict, 2023).

Figura 2 - Identificador dARK em um registro no Oasisbr

Busca / Mapa de Geodiversidade\_Rodovi... / Metadados do item

\* Citar | Imprimir | Exportar registro

Registros relacionados

- [Mapa de Geodiversidade\\_Rodovia BR-319](#)
- [Mapa de geodiversidade\\_Ferrovia Norte Sul](#)
- [Mapa de geodiversidade\\_Ferrovia Oeste - Leste](#)

Mapa de Geodiversidade. Rodovia BR-319

Autor(a) principal:	<a href="#">CPRM - SERVIÇO GEOLÓGICO DO BRASIL</a>
Data de Publicação:	2014
Tipo de documento:	Outros
Idioma:	por
Título da fonte:	<a href="#">Repositório Institucional de Geociências_Portal D</a>
dARK ID:	<a href="#">ark:/79692/001300000919</a>
Texto Completo:	<a href="https://rigeo.cprm.gov.br/handle/doc/14739">https://rigeo.cprm.gov.br/handle/doc/14739</a>

Metadados do item

Fonte: Elaborado pelo autor (2025).

A adoção do “Manifesto de Acesso Aberto a Dados da Pesquisa Brasileira para Ciência Cidadã”, em 2016, reforçou as iniciativas do Ibict para incorporar a coleta de bases de dados de pesquisa. Com isso, o Oasisbr passou a agregar conteúdo hospedado em plataformas como DSpace e Dataverse, ampliando o conjunto de documentos disponíveis para mais de 11 categorias, o que inclui artigos, dissertações, teses, trabalhos de conclusão de curso, capítulos de livros, relatórios, conjuntos de dados, *preprints*, patentes, entre outros (Bandeira, 2017; Ibict, 2025b).

Gradualmente, os dados de acessos ao Portal indicam um crescimento contínuo, atestando a crescente demanda por conteúdos de acesso aberto. Em novembro de 2023, por exemplo, o Oasisbr registrou recorde de mais de 90 mil visitantes únicos em apenas 15 dias, representando um aumento de 400% em relação ao mesmo período do ano anterior (Ibict, 2016) e em meados de 2025

já dispunha de mais de 5,4 milhões de registros. Tal incremento reflete, em grande medida, o aprimoramento contínuo da interface e das funcionalidades do Portal, bem como o fortalecimento de parcerias institucionais que promovem a difusão da produção científica nacional (Gibbon et al., 2023).

Assim, ao promover políticas e fornecer tecnologias de gestão da informação, o Oasisbr, se apresenta como agente importante para a democratização do conhecimento científico no Brasil, possibilitando maior alcance e impacto às pesquisas desenvolvidas no país e estimulando a adoção das práticas de ciência aberta em diversas comunidades acadêmicas e de pesquisa (Marques et al., 2025; Shintaku, 2014).

## **SOFTWARES UTILIZADOS ATUALMENTE NO OASISBR**

- OAI-PMH (*Open Archives Initiative Protocol for Metadata Harvesting*): O OasisBR adotou o OAI-PMH desde seu início para realizar a coleta de metadados de diferentes repositórios digitais brasileiros. Esse protocolo estabelece um padrão comum que permite a troca de informações entre sistemas, de modo que cada instituição siga regras de interoperabilidade e formatação que facilitam a integração desses metadados em um só portal (*Open Archives Initiative*, 2025). Na prática, significa que bibliotecas, repositórios institucionais e outros provedores de conteúdo podem disponibilizar os dados de suas coleções de forma organizada,

mantendo a qualidade e a consistência necessárias para a divulgação em larga escala.

- *Harversting* (Coletador) desenvolvido pela LA Referencia: Em 2015, a iniciativa Oasisbr passou por uma mudança significativa ao adotar a plataforma criada pela LA Referencia. Esse aperfeiçoamento resultou em um sistema voltado para coleta, validação e transformação de metadados, que estejam em conformidade com padrões reconhecidos internacionalmente, como o Dublin Core, e com normas nacionais de interoperabilidade. Para bibliotecários, essa estrutura facilita o monitoramento da forma como os registros são organizados, permitindo a identificação de inconsistências e o ajuste dessas informações para uso em diferentes contextos (La Referencia, 2025). Vale ressaltar que a adoção de tais padrões, reforça a capacidade de troca de dados entre distintas plataformas, permitindo que os registros coletados possam ser utilizados de maneira mais ampla e confiável, o que beneficia estudos, pesquisas e até mesmo a preservação de documentos em longo prazo. Dessa maneira, o *Harvesting* promovido pela LA Referencia agrega eficiência ao fluxo de metadados e fortalece o acesso democrático ao conhecimento.
- VuFind: O VuFind foi incorporado ao Oasisbr com o objetivo de oferecer uma busca unificada e intuitiva, além de contar com uma interface responsiva apoiada no framework CSS Bootstrap (VuFind, 2025). Assim, esse software livre atua como um

centro de referência, que possibilita aos gestores uma pesquisa simultânea em suas fontes. A facilidade de customização e a existência de uma comunidade ampla que utiliza a ferramenta são elementos que estimulam a troca de experiências e a construção de soluções. Dessa forma, o VuFind tende a se manter atualizado com novos recursos, alinhando-se às necessidades de quem trabalha diretamente com organização de coleções. Com uma visualização acessível, a localização de materiais se torna mais ágil, o que pode otimizar o atendimento ao público e o gerenciamento diário das coleções.

- Apache Solr: O Apache Solr complementa o VuFind como a tecnologia de indexação e busca, oferecendo funcionalidades que aprimoram a recuperação da informação. Essa ferramenta é reconhecida por permitir pesquisas em texto completo, bem como ressaltar partes relevantes dos documentos, o que colabora na identificação de trechos específicos dentro do conteúdo. Além disso, o Solr dispõe de recursos que organizam os resultados em categorias (facetas), auxiliando na filtragem dos materiais por autor, assunto e outros critérios de relevância. Outro ponto interessante é a possibilidade de agrupar resultados similares, o que ajuda a perceber conexões entre diferentes documentos e a encontrar conteúdos relacionados de modo mais integrado (Apache Solr, 2025). Para os gestores que atuam nos repositórios, esse conjunto de atributos pode dinamizar o trabalho de curadoria,

já que se torna mais simples explorar, classificar e compartilhar itens de interesse.

## ABRANGÊNCIA DE DOCUMENTOS COLETADOS E SEUS CRITÉRIOS PARA COLETA

Para ser cadastrada e seus documentos serem coletados, o sistema deve cumprir alguns critérios mínimos para tal. De maneira geral, todas as fontes devem:

- Fazer uso do protocolo OAI-PMH ou compatível (mediante avaliação específica).
- Gerenciar e depositar documentos de natureza científica e/ou tecnológica (ou acadêmica, dependendo da especificidade do tipo de documento).
- Adotar o esquema de metadados *Dublin Core* ou compatível (mediante avaliação específica).
- Descrever os documentos utilizando os campos descritivos/metadados básicos, sendo eles: Título, Autor, Resumo, Palavras-chave, Data de publicação ou depósito, Tipo de documento e Idioma.
- Disponibilizar os documentos na íntegra (texto completo).
- Manter conexão permanente com a Internet.
- Não solicitar login e senha para que o usuário possa acessar o documento na íntegra.
- Indicar, explicitamente, que os documentos depositados são de acesso aberto, ressalvados aqueles com indicação contrária.

- Encaminhar à equipe do Oasisbr ofício contendo os dados da fonte e do gestor responsável, conforme modelo.
- Todas as fontes de coleta devem ser mantidas por organizações brasileiras, salvo acordo prévio estabelecido.

## HISTÓRICO DE ATUALIZAÇÕES

Desde seu lançamento, o Oasisbr passou por algumas mudanças estruturais, modificando o layout, suas funcionalidades, softwares utilizados, parceiros (internos e externos) e nomenclatura. A figura a seguir apresenta a linha histórica do Portal:

Figura 3 - Histórico resumido do Oasisbr



Fonte: Elaborado pelo autor (2025).

O Oasisbr em seu lançamento foi lançado como “Portal Brasileiro de Repositórios e Periódicos de Acesso

Livre”, abrangendo alguns milhares de artigos científicos brasileiros de acesso livre, representado pela logo que lembrava um coqueiro (como observado acima). No ano de 2013 seu nome foi alterado para “Portal brasileiro de acesso aberto à informação científica”, por ter passado a reunir monografias, teses e dissertações, elevando significativamente o número de documentos reunidos para cerca de 140 mil documentos até o final deste período. O ano de 2016 foi um marco no Oasisbr, por firmar parcerias com a Rede La Referencia, adotando novas ferramentas de coleta e ganhando mais visibilidade e importância nacional, passando a disponibilizar cerca de 1.5 milhão de documentos.

A última atualização significativa foi em 2021, quando o Oasisbr passou a coletar novos tipos de documentos (*preprints*; conjunto de dados; patentes), adotando o nome de “Portal Brasileiro de Publicações e Dados Científicos em Acesso Aberto”, aumentando para mais de 5.4 milhões de documentos coletados distribuídos em mais de 1.600 fontes (Ibict, 2025d), atualizou sua interface gráfica inserindo visualização mais dinâmicas para seus usuários (lista de fontes coletadas de acordo com a região e tipo de documento e indicadores dos resultados de pesquisa) e adotou novas funções, como a exportação dos resultados da pesquisa e das fontes (por CSV RIS) e a certificação Lattes-Oasisbr (Marques *et al.*, 2023).

Cabe ressaltar que o selo de certificação Lattes-Oasisbr é exibido nas informações do currículo do usuário, nas seções de informações acadêmicas/titulação e orientações concluídas. Com o selo, é possível comprovar, de maneira rápida e simples, a autoria/orientação

concluída da tese ou dissertação e acessar o documento em acesso aberto no Oasisbr. O certificado pode ser emitido de forma automática pelo Currículo Lattes ou cadastrado manualmente pelo usuário, assim, o CV Lattes poderá verificar se a tese ou dissertação inserida no currículo do pesquisador está disponível em alguma base ou repositório acadêmico. Essa funcionalidade tem o intuito de ser o primeiro passo para outras certificações em planejamento (Ibict, 2025a; Marques *et al.*, 2025).

Atualmente, o Oasisbr se qualifica como uma das principais iniciativas do Ibict para a promoção do Acesso Aberto em território nacional. Exemplificando, a plataforma recebe cerca de 6 mil visitantes de todo o mundo por dia em média (dados de maio de 2024), alcançando cerca de 750 mil visitantes no ano de 2023 e elevando para mais de 1.8 milhões em 2024, segundo dados fornecidos pelo *Google Analytics*.

## **CONSIDERAÇÕES FINAIS**

O desenvolvimento e a consolidação do Oasisbr representam um marco significativo na democratização do acesso à informação científica no Brasil, pois desde sua criação o portal tem desempenhado um papel importante ao proporcionar visibilidade e acessibilidade à produção científica nacional e, mais recentemente, internacional, através de suas parcerias estratégicas e a adoção de tecnologias avançadas.

Ao longo dos anos, o Oasisbr evoluiu não apenas em termos de volume e diversidade de documentos coletados, mas também na sofisticação de suas ferramentas e na

abrangência de suas funcionalidades. A integração com plataformas como a LA Referencia e o RCAAP reforça a importância da cooperação internacional na promoção do acesso aberto, destacando o papel do Brasil como líder nesse movimento na América Latina.

O impacto do Oasisbr pode ser mensurado pela crescente visibilidade da produção científica brasileira, o que favorece o intercâmbio de conhecimento e a colaboração entre pesquisadores. Além disso, a plataforma tem contribuído para evitar a duplicação de esforços e recursos no desenvolvimento de pesquisas, ao fornecer dados e informações confiáveis que são amplamente acessíveis.

A introdução de funcionalidades como a certificação Lattes-Oasisbr e integração ao projeto dARK, exemplifica o compromisso contínuo do Ibict em aprimorar a qualidade e a credibilidade da informação científica disponibilizada. Essa funcionalidade não apenas facilita a verificação de autoria e orientação de teses e dissertações, mas também fortalece a integração entre diferentes sistemas de informação acadêmica.

No entanto, os desafios permanecem, sabendo que a manutenção e o aprimoramento contínuo do Oasisbr demandam investimentos constantes em tecnologia e capacitação profissional. Além disso, é crucial promover uma cultura de acesso aberto entre os pesquisadores e instituições, incentivando a adoção de práticas de ciência aberta e a submissão de trabalhos em repositórios de acesso livre. A expansão para a coleta de bases de dados de pesquisa e a integração com novas tipologias documentais são passos importantes para consolidar o Oasisbr como

uma referência na disseminação de informação científica em acesso aberto.

Por fim, o Oasisbr tem se estabelecido como um pilar no ecossistema de acesso aberto no Brasil, apresentando uma trajetória de evolução e impacto, demonstrando o potencial transformador das iniciativas de acesso livre para a democratização do conhecimento e a promoção da ciência aberta.

## REFERÊNCIAS

APACHE. **Apache Solr**. Disponível em: <https://solr.apache.org>. Acesso em: 10 abr. 2025.

BANDEIRA, Pablo Matias. **Movimento de acesso aberto no Brasil**: contribuição do Instituto Brasileiro de Informação em Ciência e Tecnologia a partir da implementação do Sistema Eletrônico de Editoração de Revistas. 2017. Dissertação (Mestrado em Ciência da Informação) - Universidade Federal da Paraíba, João Pessoa, 2017.

CARVALHO SEGUNDO, Washington Luís Ribeiro de et al. Portal Brasileiro de Acesso Aberto: ampliação e cooperação internacional. In: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS, 20., 2017, Salvador. **Anais** [...], Salvador: [s.n.], 2017.

DIAS, Thiago Magela Rodrigues; CARVALHO SEGUNDO, Washington Luís Ribeiro de; MATAS, Lautaro. Utilizando o framework LattesDataXplorer para Vincular Automaticamente os Currículos da Plataforma Lattes à Biblioteca Digital Brasileira de Teses e Dissertações (BDTD). **Ciência da Informação**, Brasília, v. 48, n. 3, p. 355-340, 2020. DOI: 10.18225/ci.inf.v48i3.5003.

GIBBON, Camila de Azevedo et al. Ciência Aberta brasileira: uma análise a partir do oasisbr. In: WORKSHOP DE INFORMAÇÃO, DADOS E TECNOLOGIA, 6., 2023. **Anais** [...], Brasília, p. 1-13, 2023. DOI: 10.22477/vi.widat.07.

GOMES, S. H. Acesso aberto: políticas e iniciativas para compartilhamento da informação científica. **Revista X**, [s.l.], v. 1, n. 1, p. 10-25, 2014.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Ibict lança Manifesto de Acesso Aberto a Dados da Pesquisa Brasileira para Ciência Cidadã**. Brasília, 2016. Disponível em: <https://antigo.ibict.br/sala-de-imprensa/noticias/item/1856-ibict-lanca-manifesto-de-acesso-aberto-a-dados-da-pesquisa-brasileira-para-ciencia-cidada>. Acesso em: 8 abr. 2025.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Oasisbr**: sobre. 2025a. Disponível em: <https://oasisbr.ibict.br/vufind/about/home>. Acesso em: 1 abr. 2025.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Portal Brasileiro de Publicações e Dados Científicos em Acesso Aberto** (Oasisbr). 2025b. Disponível em: <https://oasisbr.ibict.br/vufind>. Acesso em: 1 abr. 2025.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). Quantitativo de acessos ao Portal Oasisbr é o maior da história, segundo dados do Google Analytics. **Ibict Notícias**, 21 nov. 2023. Disponível em: [www.gov.br/ibict/pt-br/central-de-conteudos/noticias/2023/novembro/quantitativo-de-acessos-ao-portal-oasisbr-e-o-maior-da-historia-segundo-dados-do-google-analytics](http://www.gov.br/ibict/pt-br/central-de-conteudos/noticias/2023/novembro/quantitativo-de-acessos-ao-portal-oasisbr-e-o-maior-da-historia-segundo-dados-do-google-analytics). Acesso em: 8 abr. 2025.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Manifesto Brasileiro de Apoio ao Acesso Livre à Informação Científica.** Disponível em: <https://livroaberto.ibict.br/Manifesto.pdf>. Acesso em: 11 ago. 2025c.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Fontes coletadas:** conheça os tipos de fontes coletadas pelo Oasisbr. Disponível em: <https://oasisbr.ibict.br/vufind/datasources/home>. Acesso em: 11 ago. 2025d.

LA REFERENCIA. **Tecnologia.** Disponível em: [www.lareferencia.info/en/services/tecnologia](http://www.lareferencia.info/en/services/tecnologia). Acesso em: 10 abr. 2025.

LIINC EM REVISTA. Brasília: Ibict, v. 8 n. 2, 2012. Novos paradigmas da comunicação científica: ampliando o debate.

MARQUES, Gabriel Silveira et al. Padronização de metadados adotados no portal agregador de conteúdo científico Oasisbr. **Ciência da Informação**, Brasília, v. 54, n. 2, p. 1-7, 2025. DOI: 10.18225/ci.inf.v54i2.7223.

MARQUES, Gabriel Silveira et al. Relato de experiência do processo de desenvolvimento e planejamento do Novo Oasisbr. **Revista Científica da UEM: Série Letras e Ciências Sociais**, v. 4, p. 32-32, 2023.

MATAS, Lautaro et al. Bridging the Gap: Implementing the Decentralized Archival Resource Key (dARK) - The Evolution from Proof of Concept to Service Implementation. **PIDFest 2024**, Prague, 2024.

OPEN ARCHIVES INITIATIVE. OAI-PMH. **The Open Archives Initiative Protocol for Metadata Harvesting.** Disponível em: [www.openarchives.org](http://www.openarchives.org). Acesso em: 10 abr. 2025.

ORTELLADO, Pablo. O acesso aberto e a história de duas comunidades científicas: ciências naturais e ciências humanas. **Logos**, v. 15, n. 3, p. 185-197, 2008.

RIOS, Fahima Pinto; LUCAS, Elaine Rosangela de Oliveira; AMORIM, Igor Soares. Manifestos do movimento de acesso aberto: análise de domínio a partir de periódicos brasileiros.

**Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v. 15, n. 1, p. 148-169, 2019.

RODRIGUES, Flávio Salgueiro; GUGLIOTTA, Alexandre Carlos. A temática do acesso aberto na literatura científica da área da Ciência da Informação no Brasil entre 2002 e 2018. **Múltiplos Olhares em Ciência da Informação**, Belo Horizonte, v. 9, n. 1, 2019.

SHINTAKU, Milton. **Federação de Repositórios Científicos**: identificação, análise e proposta de modelo baseado nas tendências tecnológicas e da Ciência. 2014. Tese (Doutorado em Ciência da Informação) - Universidade de Brasília, Brasília, 2014.

SILVA, Fabiano Couto Corrêa da; SILVEIRA, Lúcia da. O ecossistema da ciência aberta. **Transinformação**, Campinas, v. 31, e190001, 2019. DOI: 10.1590/2318-0889201931e190001.

SILVA, M. F.; ALENCAR, M. F. F.; SOUZA, A. C. O Portal Oasisbr e a promoção do acesso aberto à informação científica. In: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS, 19., 2014. **Anais** [...], Belo Horizonte: UFMG, 2014.

VUFIND. **VuFind Project**. Disponível em: <https://vufind.org>. Acesso em: 10 abr. 2025.

## Como citar este capítulo

MARQUES, Gabriel Silveira. Portal Brasileiro de Publicações e Dados Científicos em Acesso Aberto (Oasisbr). In: AMARO, Bianca; CAMPOS, Fhillipe de Freitas; BARCELOS, Janinne. (org.). **Infraestruturas de Ciência e de Acesso Aberto no Brasil:** iniciativas do Instituto Brasileiro de Informação em Ciência e Tecnologia. Brasília, DF: Editora Ibict, 2025. Cap. 2, p. 55-76. DOI: 10.22477/9788570132543.cap2

# REDE CARINIANA: UM ESFORÇO COLABORATIVO DE PRESERVAÇÃO DIGITAL DO IBICT

Miguel Ángel Márdero Arellano



## APRESENTAÇÃO

A Ciência Aberta e a preservação digital vêm estabelecendo uma relação necessária na conscientização dos efeitos das crises econômicas, políticas e das alterações climáticas nos acervos científicos e culturais mantidos em bibliotecas, arquivos e museus. O trabalho constante de práticas sustentáveis de curadoria das memórias registradas digitalmente envolve a gestão da informação de longo prazo, expressada em programas, planos de ação, projetos e investimentos, que garantam condições concretas de acesso ao longo do tempo.

A concretização de soluções economicamente sustentáveis, necessárias para garantir uma vida longa aos objetos digitais e a sua prevenção da obsolescência tecnológica não é, ainda, uma realidade generalizada. As instituições precisam manter a qualidade original do patrimônio digitalizado e nativo digital, avaliando estratégias e cuidando dos processos de gestão e o impacto de seu desempenho, a fim de efetivarem as ações de preservação digital.

O acesso a uma cópia digital dos materiais derivados da pesquisa científica, disponível em um ou mais repositórios e bibliotecas digitais, é um dos objetivos do Movimento de Acesso Aberto desde seu início. O compartilhamento dessas coleções em grandes redes e agregadores nacionais e internacionais também colaborou para a consolidação das práticas de preservação digital em redes. Universidades e centros de pesquisa têm estabelecido acordos de cooperação técnica para o intercâmbio de experiências em projetos e pesquisas que

agregam conhecimento sobre processos e metodologias técnicas para tentar resolver algumas das principais ameaças aos acervos digitais. As instituições que estão comprometidas com a sustentabilidade dos serviços de informação que prestam, pela sua proteção e certificação, sabem que essas iniciativas não envolvem apenas a sua infraestrutura tecnológica, mas também a qualidade das suas práticas de gestão.

A preservação digital na gestão dos dados das pesquisas científicas presta um serviço de interesse vital à comunidade de investigação. Os dados devem ser armazenados e reaproveitados por outros pesquisadores, e não apenas mantidos até que não sejam mais necessários. Com os planos de gestão de dados começa a ser garantida sua preservação e acesso ao longo prazo. Também, o uso generalizado de repositórios colabora nesse sentido, com plataformas que podem ser integradas às redes de preservação digital, como é o caso do repositório Figshare (2025) que, além de facilitar a localização e citação dos dados, alimenta a rede Clockss (2025) para arquivar de forma sustentável todo o conteúdo da Web produzido na área científica, sem fins lucrativos e promovendo parcerias entre editores e bibliotecas acadêmicas (Silva, 2017).

As redes de preservação digital oferecem oportunidades para colaboração internacional, além de serem uma opção tecnologicamente viável para preservar conteúdo digital institucional. A preservação digital ainda é amplamente vista como um empreendimento complexo e caro, exigindo anos de planejamento e grandes aportes de dinheiro e outros recursos. Entretanto, as experiências internacionais das últimas décadas sugerem que é

perfeitamente possível construir soluções de preservação robustas, escaláveis e economicamente sustentáveis com recursos relativamente modestos, além de possibilitar sua extensão a diferentes tipos de instituições em vários países.

A necessidade decorrente dessa realidade é que existam programas de formação de profissionais de diferentes áreas do conhecimento, que saibam lidar com a complexidade dos objetos digitais, do seu registro nos metadados, das informações sensíveis e de intervenções subsequentes nesses acervos. No Brasil, o Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) vem pesquisando e disponibilizando conhecimento na área da preservação digital desde o início do século XXI, conscientizando as diferentes comunidades que lidam com registros de informação científica e cultural sobre as habilidades específicas que seus gestores precisam adquirir. Sua primeira preocupação foi o papel da preservação na gestão das publicações eletrônicas, principalmente dos periódicos científicos em formato digital, que após o período da pandemia ocasionada pela Covid, constatou um aumento no entendimento da importância do acesso e arquivamento seguro das revistas científicas por parte das equipes editoriais.

## **A REDE CARINIANA DO IBICT**

Desde 2010, equipes do Ibict iniciaram pesquisas e testes para implementação de uma rede nacional de preservação digital, propondo um projeto estruturante no seu Plano Diretor 2011-2015 (Ibict, 2010). A preservação digital foi, naquele momento, o único projeto estruturante

do instituto, tendo como metas: (63) desenvolver política de preservação digital da rede nacional; (64) criar a infraestrutura tecnológica necessária, (65) implantar a rede nacional e, (66) integrar a rede nacional à iniciativa LOCKSS.

No desenvolvimento do projeto, foi indispensável a realização de pesquisas em organizações internacionais que possuíam projetos de preservação digital. Durante o período de estudo inicial foi possível realizar visitas técnicas à Universidade do Minho, em Portugal; ao projeto Data-Pass, da Harvard University; e, à Aliança LOCKSS, da *Stanford University Libraries*. Em 2012, o trabalho de prospecção e de soluções tecnológicas e de iniciativas internacionais no Ibict concretizou a proposta de uma política de preservação digital e a criação de um protótipo do sistema na etapa de teste das funcionalidades desejadas.

A política de preservação digital para o gerenciamento de uma rede nacional de preservação digital foi elaborada tendo como base uma lista de requisitos funcionais para garantir a validade dos registros e a sua permanência, seguindo o modelo de referência OAIS (Márdero Arellano; Leite, 2009). A respectiva proposta descreveu as diretrizes que o Ibict adotou relacionadas à administração da rede. Outro resultado da prospecção foi a proposta de solução tecnológica e metodológica *Lots of Copies Keep Stuff Safe* (LOCKSS), de autoria da *Stanford University*, que o Instituto indicou às instituições parceiras do projeto.

A escolha da ferramenta LOCKSS foi realizada após análise de sua estrutura, requisitos e fluxos de dados. A definição das condições para a realização das parcerias

que possibilitaram a implementação da rede piloto, dirigida à preservação de periódicos que utilizam a ferramenta *Open Journal Systems/Sistema Eletrônico de Editoração de Revistas* (OJS/SEER), foi resultado das atividades da primeira equipe de bolsistas do Programa de Capacitação Institucional (PCI).

Os acordos elaborados definiram a interação entre as equipes que passariam a gerenciar essa e outras sub-redes. A composição inicial foi estruturada em sub-redes considerando a necessidade de preservação do tipo de documento e de seu valor quanto às ligações que ele possui, sejam eles periódicos científicos, livros eletrônicos, teses, dissertações, dados de pesquisa (independentemente do formato), ou base de dados e multimídia. A garantia da existência de um avançado conjunto de serviços de preservação foi considerada no momento da criação de espaços para o armazenamento de documentos digitais em várias instituições.

A Cariniana iniciou suas atividades como rede em 2012, a partir da realização de uma reunião com representantes de instituições de ensino superior convidadas a participar da proposta. Nesse evento foi apresentada a experiência da Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), que já havia efetuado a preservação de suas teses e dissertações na rede CLOCKSS. A constatação de que a gestão e a instalação da ferramenta no parque tecnológico do Ibict seria factível promoveu os contatos iniciais com os líderes e criadores da Aliança LOCKSS da *Stanford University*, com o objetivo de conhecer melhor a única solução de preservação digital distribuída no mundo.

Desde o início do século XXI, o LOCKSS foi se consolidando como um sistema de armazenamento seguro de longo prazo em vários países. Foram testados seus atributos de integração com sistemas de editoração (OJS), com repositórios (DSpace, Dataverse) e com outros sistemas de arquivamento (Archivematica, Heritrix). No Brasil, o Ibict foi a primeira instituição que distribuiu o software OJS como uma ferramenta que contempla ações primordiais para a automação de atividades de editoração de periódicos científicos, atuando em treinamentos e palestras para editores científicos em vários eventos (Márdero Arellano; Santos; Fonseca, 2005). A partir desta iniciativa, editores brasileiros começaram a compreender a preservação digital em consonância com o acesso aberto no país. Considerando este tipo de preservação como uma prática institucional, os gestores dos portais de revistas tiveram os primeiros contatos com iniciativas internacionais que envolvem a estruturação de redes de serviços de arquivos e bibliotecas que, por sua vez, compartilham acervos e possibilitam o intercâmbio entre comunidades de usuários de sistemas de preservação digital.

Em 2014, com apoio da Financiadora de Estudos e Projetos (Finep), a licença LOCKSS para uso na construção da Rede Cariniana foi adquirida, e a adesão à Aliança LOCKSS resultou na apreciação crítica da necessidade da aplicação da respectiva ferramenta no período relativo a 12 meses, com extensão para 24 meses (2014 a 2016). O Ministério de Ciência, Tecnologia e Inovações (MCTI) permitiu a extensão da adesão, seguindo um processo de inexigibilidade de licitação pois a *Stanford University*

*Libraries* era a única instituição a disponibilizar ferramenta de preservação digital distribuída que observava os requisitos funcionais para construção de redes colaborativas com nós de armazenamento nacional.

A adoção da ferramenta LOCKSS foi oficializada em 2014 e, desde então, ela tem sido a base das atividades da Rede Cariniana. Em 2019, o Ibict formalizou, regimentalmente, a estrutura da coordenação da Rede, em portaria do MCTI, definindo os comitês, os tipos de parcerias e a documentação oficial das suas atividades.

A operação de aquisição da licença LOCKSS permitiu a participação do Ibict nas atividades de desenvolvimento e transferência de informações dentro da Aliança LOCKSS Internacional, assim como o apoio técnico para o funcionamento da Rede Nacional. O LOCKSS é uma ferramenta que possui escalabilidade alta no armazenamento de objetos digitais, com capacidade de salvaguardar milhões de objetos digitais em nós de armazenamento distribuídos dentro do país. Os *plug-ins* acoplados ao sistema dão subsídio para armazenar e recuperar todo tipo de material digital existente.

Até o presente momento, o software LOCKSS continua estabelecendo a arquitetura de sub-redes colaborativas para tratamento e preservação digital de publicações técnico-científicas. No Ibict, o reconhecimento da importância da participação da Aliança LOCKSS se baseou no pressuposto de que era necessário desenvolver redes colaborativas nacionais que pudessem garantir não somente as migrações, mas também o contexto, a estrutura e a acessibilidade dos documentos digitais produzidos no país, contribuindo para salvaguardar

o patrimônio nacional. A aceitação de uma solução prática de preservação digital como o LOCKSS, adotada por instituições de investigação na Europa, Ásia e América do Norte, significou estar em harmonia com outras iniciativas que têm evidenciado o valor científico do sistema como padrão internacional de preservação digital. (Márdero Arellano, 2012).

Dentro do escopo da Rede Cariniana, a tipologia de documentos digitais preservados propiciou a formulação de atividades relacionadas aos arquivos, aos museus e às bibliotecas. Os acervos históricos, científicos e arquivísticos das instituições parceiras permitiram, ainda, que alguns formatos de arquivos fossem os mais estudados. Os formatos de fundos de memória audiovisual, a necessidade de metodologias de avaliação de níveis de preservação, o arquivamento de páginas web, de correio eletrônico e de um programa de formação em preservação digital concentraram os esforços da coordenação da Cariniana e de suas instituições parceiras nos seus primeiros anos de funcionamento.

Essa iniciativa levou o Ibict a se tornar um dos canais oficiais na identificação, avaliação e disseminação de estratégias de preservação digital que, há mais de uma década, vem desenvolvendo e implantando uma infraestrutura técnica e operacional com outras instituições parceiras. O foco da rede de preservação digital distribuída é o diagnóstico e preservação de importantes acervos digitais da memória científica e tecnológica do país. Com essa perspectiva, os serviços da rede começaram a ser organizados mesmo antes de 2015, sendo seguidas pelas atividades de ensino e pesquisa, a partir da

efetivação de parcerias e por meio de acordos de cooperação técnica e termos de adesão. Além disso, os planos de trabalho elaborados junto às parcerias contribuíram na definição do escopo dos projetos que a Rede mantém e apoia.

## A REALIDADE ATUAL

Oficialmente, o Ibict estabeleceu, desde 2020, metas voltadas para a preservação digital (Ibict, 2020): efetuar estudos comparativos de tecnologias abertas e da transferência de tecnologias voltadas à identificação de objetos digitais, e apoiar atividades de pesquisa em preservação digital e o trabalho prospectivo de colaboradores através de uma rede de pesquisa. Com vistas à criação de novos produtos e serviços, o grupo de pesquisa da Cariniana acompanha e divulga o desenvolvimento de ferramentas tecnológicas que estejam em consonância com as normas internacionais de gestão da preservação de acervos digitais (Ibict, 2025).

A relevância das atividades de pesquisa para o desenvolvimento da Rede Cariniana está associada à quantidade de informação digital produzida no Brasil pela área científica e pela área acadêmica, cujos critérios de validação exigem garantias de preservação a longo prazo. Estes critérios incluem autenticidade e confiabilidade dos objetos digitais e dos seus registros de identificação, em contraponto ao imensurável volume de informações, reprodução e compartilhamento em meios digitais não confiáveis e à própria condição de obsolescência tecnológica possível dos serviços de informação digital.

Considerando a dimensão de longa duração e o impacto da obsolescência tecnológica dos objetos digitais não preservados, o papel de uma rede de preservação digital coordenada pelo Ibict precisa manter uma infraestrutura de suporte nacional e canais de interlocução dedicados à pesquisa e à promoção da salvaguarda do patrimônio científico digital. Assim, no contexto atual, merece atenção a questão da gestão institucional da documentação científica, que, em conformidade com os ditames da dimensão pública da comunicação científica, inclui o acesso amplo e aberto à informação e ao conhecimento produzido associado, além da necessidade de criação de políticas e serviços de preservação digital.

O Ibict continua a realizar pesquisas nas iniciativas reconhecidas pelo uso de soluções tecnológicas de acesso aberto amplamente adotadas em projetos internacionais de preservação digital, identificando ferramentas que estão sujeitas à verificação do gerenciamento e controle de requisitos funcionais para o funcionamento em rede de preservação distribuída. O esforço levou a inclusão da rede Cariniana no registro de agências de arquivamento de periódicos eletrônicos, *The Keepers*, mantido desde 2008 pelo Centro Internacional do ISSN (*International Standard Serial Number*) em Paris, contando atualmente com um grupo de 19 iniciativas internacionais que preservam de forma segura e certificada publicações digitais (ISSN..., 2025) e que observam políticas de acesso e preservação e ações de auditoria. *The Keepers* é uma organização internacional que tem como missão o registro de periódicos eletrônicos que possuem um conteúdo preservado a longo prazo. Ela também oferece serviços de divulgação sobre as

melhores práticas de salvaguarda do acesso em longo prazo das publicações científicas e com destacado conteúdo cultural em formato digital.

O papel do *The Keepers* é servir como fonte de evidência para as bibliotecas e outros serviços de informação verificarem se o conteúdo digital que elas colecionam não tem alto risco de perda. A Cariniana é a rede que mais preserva periódicos que não estão em nenhuma outra agência, promovendo assim a visibilidade e a preservação que o *The Keepers* concede, uma vez que há rigorosos critérios e requisitos para garantir a participação de instituições<sup>3</sup>. Em 2023 a Coordenação da Cariniana disponibilizou seu Programa de Formação em Preservação Digital, contando com a colaboração de especialistas e pesquisadores nacionais e internacionais. O programa é o único em atividade em países da América Latina, reconhecido por instituições como a *International Federation of Library Associations and Institutions* (IFLA) e pela Associação Ibero-Americana de Preservação Digital (APREDIG, 2025).

A rede Cariniana continua desenvolvendo atividades práticas de preservação digital de documentos tendo como principal característica a colaboração entre instituições e especialistas. A experiência regional de preservação digital distribuída segue o modelo de uma aliança internacional para salvaguardar a produção intelectual de um país. O modelo envolve a gestão da informação em acesso aberto, onde a perpetuidade da recuperação dos dados é garantida

---

<sup>3</sup> A Cariniana arquiva 2.178 revistas com pelo menos um volume no sistema LOCKSS (julho 2024).

pela forma aberta de gestão das etapas de entrada, arquivamento e difusão em sistemas eletrônicos de informação confiáveis.

## **CONSIDERAÇÕES FINAIS**

Ainda há uma grande distância entre as propostas de iniciativas internacionais de preservação digital que surgem e o status atual no Brasil, dada a insuficiência de recursos e o planejamento inadequado das práticas para permitir o acesso em longo prazo. O esforço científico do Ibict tem sido apoiar a pesquisa de soluções tecnológicas de acesso aberto amplamente adotadas em projetos internacionais de preservação digital, identificando ferramentas que passaram pela verificação do gerenciamento e controle de requisitos funcionais para o funcionamento em rede de preservação distribuída.

É importante destacar que a estrutura da Rede Cariniana incentiva a cooperação, gerando uma comunidade dedicada à preservação digital. Para isso, é fundamental que cada instituição colaboradora tenha comprometimento na manutenção de sua infraestrutura local nos próximos anos e que o Ibict permita que sua equipe desenvolva habilidades necessárias para continuar coordenando o suporte a todos os parceiros da rede (Márdero Arellano, 2012).

Com a adoção do modelo distribuído, atividades de colaboração entre as instituições parceiras da Rede levaram à concretização de políticas de preservação em instituições de ensino e pesquisa no país, assim como a participação em projetos institucionais de preservação de

acervos analógicos desde a fase de digitalização até seu armazenamento seguro em repositórios digitais (Silva *et al.*, 2019). Entre eles estão o projeto de preservação dos arquivos da Comissão Nacional da Verdade no Arquivo Nacional, o Serviço de Preservação Digital da Câmara dos Deputados, a política de preservação digital da Universidade Estadual Paulista (Unesp), da Fundação Oswaldo Cruz (Fiocruz) e da Fundação Biblioteca Nacional, entre outras. A Cariniana desenvolve projetos em parceria com 28 instituições nacionais e internacionais, além da colaboração com redes de 11 países.

As atividades de preservação digital devem ser observadas na forma como são concebidas e na sua atuação, como finalidade ativa dentro da curadoria e gestão de acervos digitais, bem como de perfis específicos, de acordo com as necessidades da comunidade de usuários e do cenário local. Uma interpretação clara da preservação digital pode permitir a sua adoção na cultura organizacional com atividades de gestão e planejamento estratégico.

O registro da produção científica no Brasil, como processo, corre o risco de não garantir às gerações futuras o seu legado, que depende de práticas intensivas, sistemáticas e seletivas de arquivamento digital, tendo em conta o longo prazo. A preservação digital como prática institucional está, ainda, vinculada às experiências relativamente restritas no Brasil. A perspectiva de expansão de procedimentos e de institucionalização de políticas nesse campo envolve a estruturação em redes de arquivos ou de bibliotecas que contribuam para disseminar as tecnologias e os serviços de suporte, promover o compartilhamento de acervos e ampliar as possibilidades

de intercâmbio entre pesquisadores dedicados às mesmas temáticas. No país, as áreas que lidam com informação e com atividades de desenvolvimento científico não contam com canais oficiais para a identificação, comunicação e avaliação da preservação digital em longa duração. A organização desses canais não deve desconsiderar sua dimensão internacional, pois no momento em que se escolhe determinado conjunto de critérios de preservação, este deve ser integrado aos outros sistemas que adotem critérios semelhantes.

O Ibict é uma instituição com ampla experiência na prospecção, absorção e customização de sistemas baseados em software livre. Na última década, o Instituto tem reconhecido a importância do estudo de soluções apropriadas para a preservação das informações em C&T em formato digital. Além da criação de infraestrutura tecnológica, resta o desafio de investigação sobre a relação entre memória institucional e tecnologia, que continua sendo a preservação de todas as características de um objeto digital (Tavares; Márdero Arellano, 2022).

Os compromissos do Ibict e da rede Cariniana continuam sendo oferecer alternativas para que as instituições brasileiras possam colecionar, armazenar e promover o acesso ao conteúdo selecionado por meio de cópias autorizadas, com a aplicação de políticas voltadas para a preservação do patrimônio científico e cultural digital, condição fundamental para a democratização da informação no país.

## REFERÊNCIAS

ASOCIACIÓN IBEROAMERICANA DE PRESERVACIÓN DIGITAL (APREDIG). Disponível em: [www.apredig.org](http://www.apredig.org). Acesso em: 11 ago. 2025.

CLOCKSS. **Preserving Scholarship for Future Generations**. Disponível em: <https://clockss.org>. Acesso em: 11 ago. 2025.

FIGSHARE. Disponível em: <https://figshare.com>. Acesso em: 11 ago. 2025.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **[PDU]**: Plano Diretor [da Unidade ou] do Ibict: 2011-2015: [planejamento estratégico do Ibict]. Brasília: Ibict, [2010].

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **[PDU]**: Plano Diretor [da Unidade ou] do Ibict: 2018-2022: [planejamento estratégico do Ibict]. Brasília: Ibict, [2020].

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Grupo de Pesquisa Dríade**. Brasília: Ibict, 2025. Disponível em: [https://cariniana.ibict.br/?page\\_id=341](https://cariniana.ibict.br/?page_id=341). Acesso em: 11 ago. 2025.

ISSN INTERNATIONAL CENTRE. **Stats**. Disponível em: <https://keepers.issn.org/stats>. Acesso em: 11 ago. 2025.

MÁRDERO ARELLANO, Miguel Ángel. Cariniana: uma rede nacional de preservação digital. **Ciência da Informação**, Brasília, v. 41, n. 1, p. 83-91, jan./abr. 2012. DOI: 10.18225/ci.inf.v41i1.1354.

MÁRDERO ARELLANO, Miguel Ángel; LEITE, Fernando César Lima. Acesso aberto à informação científica e o problema da preservação digital. **Biblio**, Brasília, n. 35, mar./jun. 2009.

MÁRDERO ARELLANO, Miguel Ángel; SANTOS, Regina dos; FONSECA, Ramón da. SEER: Disseminação de um sistema eletrônico para editoração de revistas científicas no Brasil. **Arquivística.net**, Rio de Janeiro, v. 1, n. 2, p. 75-82, jul./dez. 2005.

SILVA, Fabiano Couto Corrêa da. **Gestión de datos de investigación oceanográfica**: propuesta de un modelo para Brasil. 2017. Tese (Doctorado en Información y Documentación en la Sociedad del Conocimiento) - Universidad de Barcelona, Departament de Biblioteconomia, Documentació i Comunicació Audiovisual, Barcelona, 2017.

SILVA, Mayane Paulino de Brito e et al. Contribuições da Rede Cariniana para a preservação digital nos repositórios digitais institucionais: o caso da UFRN. **Informação em Pauta**, Fortaleza, v. 4, n. especial, p. 99-116, 2019. DOI: 10.36517/ip.v4iespecial.42607.

TAVARES, Maria de Fátima Duarte; MÁRDERO ARELLANO, Miguel Ángel. O Ibict e a preservação da memória científica. In: BRAGA, Tiago Emmanuel Nunes; MÁRDERO ARELLANO, Miguel Ángel. (org.). **Hipátia**: modelo de preservação para repositórios arquivísticos digitais confiáveis. Brasília: Ibict, 2022. p. 26-39. DOI: 10.22477/9786589167501.cap2.

Como citar este capítulo:

MÁRDERO ARELLANO, Miguel Ángel. Rede Cariniana: um esforço colaborativo de preservação digital do Ibict. In: AMARO, Bianca; CAMPOS, Phillippe de Freitas; BARCELOS, Janinne. (org.). **Infraestruturas de Ciência e de Acesso Aberto no Brasil**: iniciativas do Instituto Brasileiro de Informação em Ciência e Tecnologia. Brasília, DF: Editora Ibict, 2025. Cap. 3, p. 77-94. DOI: 10.22477/9788570132543.cap3



# REPOSITÓRIO INSTITUCIONAL DO IBICT (RIDi)

Cássio Teixeira de Moraes



## APRESENTAÇÃO

O Repositório Institucional do Instituto Brasileiro de Informação em Ciência e Tecnologia (RIDI) é a base de dados responsável por registrar, preservar e disseminar a produção científica e técnica realizada no âmbito do Instituto. Alinhado ao contexto dos repositórios institucionais de Acesso Aberto, os repositórios digitais são bases de dados destinadas ao armazenamento e gerenciamento da produção científica, somando-se às estruturas que compõem a comunicação científica (Leite, 2009). O RIDI foi oficialmente instituído pela Portaria nº 043, de 10 de setembro de 2014, que estabelece sua política de funcionamento.

A produção armazenada no repositório comprehende resultados consolidados de pesquisa, publicados em canais com avaliação por pares ou aprovados por comitês científicos. Essa produção é originada da comunidade científica institucional, composta por servidores, colaboradores, bolsistas de pesquisa e discentes do Programa de Pós-Graduação em Ciência da Informação do Ibiict.

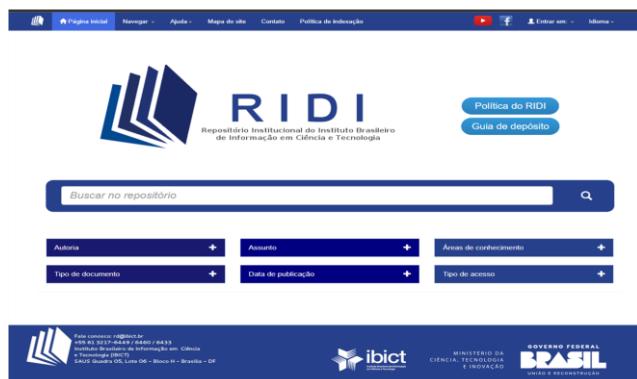
O acervo digital do RIDI é composto por artigos publicados em periódicos científicos, trabalhos apresentados em eventos científicos que tenham passado por revisão por pares, livros e capítulos de livros resultantes de projetos de pesquisa, além das dissertações, teses e relatórios de pós-doutorado dos cursos de pós-graduação da instituição.

O conteúdo armazenado no repositório, proveniente dessas publicações, abrange diversas áreas, com ênfase

em Ciência da Informação e Tecnologia da Informação. Dessa forma, os usuários potenciais do RIDI incluem, além de servidores e colaboradores do Ibitc, pesquisadores, acadêmicos e cidadãos que tenham interesse nas áreas de Ciência e Tecnologia da Informação. Essa iniciativa amplia o acesso ao conhecimento e fortalece as pesquisas desenvolvidas pelo instituto, promovendo maior integração entre a produção científica e a sociedade.

Desenvolvido com o software livre DSpace, o repositório oferece recursos avançados de navegação e busca, incluindo filtros por autor, assunto, tipo de documento, data de publicação, tipo de acesso e áreas do conhecimento. A Figura 1 apresenta a interface da página inicial do RIDI:

Figura 1 - Página inicial do RIDI



Fonte: Captura de tela do RIDI (Ibitc, 2024).

Conforme estabelecido em sua política, o povoamento do RIDI ocorre preferencialmente por meio do

autoarquivamento, modelo no qual cada autor é responsável pelo depósito de sua produção científica. Contudo, a adoção complementar do depósito mediado também é praticada, como estratégia para agilizar e ampliar a inclusão de documentos no repositório. Esse processo é conduzido pela Coordenação de Tratamento, Análise e Disseminação da Informação Científica (Codic), responsável pela curadoria dos dados - incluindo a revisão, edição, validação e disponibilização dos registros inseridos -, bem como pelo pleno funcionamento do repositório, pelo desenvolvimento contínuo e pela incorporação de aprimoramentos técnicos que assegurem sua atualização e eficiência.

Em consonância com sua política institucional, o RIDI tem como objetivo geral consolidar uma base de dados que reúna e preserve a produção científica do Ibict, assegurando sua autenticidade, integridade e acesso a longo prazo. Entre os objetivos específicos do repositório, destacam-se:

- a) Preservar as produções científicas do Ibict;
- b) Potencializar o intercâmbio com outras instituições;
- c) Acelerar e fortalecer desenvolvimento das pesquisas institucionais;
- d) Ampliar o acesso à sua produção científica;
- e) Facilitar o acesso à informação científica de uma forma geral;
- f) Otimizar a gestão de investimentos em pesquisa do Ibict.

O RIDI é organizado em quatro comunidades, subdivididas em coleções, de acordo com a origem institucional da produção científica preservada:

1. *IBBD - Instituto Brasileiro de Bibliografia e Documentação (1954-1975)*: contém a produção científica do IBBD, nomenclatura do Instituto entre os anos de 1954-1975. A preservação da memória institucional é um dos compromissos desta comunidade, composta pelas seguintes coleções:
  - 1.1 Artigos em Revistas publicadas no Brasil (IBBD);
  - 1.2 Artigos em Revistas publicadas no Exterior (IBBD);
  - 1.3 Dissertações de Mestrado (IBBD);
  - 1.4 Livros e Capítulos de Livros Nacionais (IBBD);
  - 1.5 Trabalhos apresentados em Congresso no Exterior (IBBD).
2. *Ibict - Instituto Brasileiro de Informação em Ciência e Tecnologia (1976 ---)*: contém a produção científica dos profissionais atuantes do Ibict a partir do ano de 1976. São sete coleções desta comunidade:
  - 2.1 Artigos em Revistas publicadas no Brasil;
  - 2.2 Artigos em Revistas publicadas no Exterior;
  - 2.3 Dissertações de Mestrado;
  - 2.4 Teses de Doutorado;
  - 2.5 Livros e Capítulos de Livros Nacionais;

- 2.6 Trabalhos apresentados em Congresso no Brasil;
- 2.7 Trabalhos apresentados em Congresso no Exterior.
3. *PPGCI IBICT-UFF*: comprehende a coleção de teses e dissertações do Programa de Pós-Graduação em Ciência da Informação desenvolvido em parceria entre o Ibict e a Universidade Federal Fluminense (UFF). Trata-se de uma coleção descontinuada e fechada, que não recebe novos trabalhos.
  4. *PPGCI IBICT-UFRJ*: Reúne as teses e dissertações do Programa de Pós-Graduação em Ciência da Informação mantido pelo Ibict em parceria com a Universidade Federal do Rio de Janeiro (UFRJ), com atualizações regulares da produção acadêmica do programa.

Atualmente o RIDI disponibiliza cerca de 1.170 publicações científicas. No entanto, esse número ainda não reflete a totalidade da produção científica realizada pela comunidade institucional do Instituto. Embora o depósito das publicações seja uma exigência estabelecida pela política do repositório, muitos trabalhos ainda não foram arquivados na base. A realização de atividades de formação e sensibilização sobre o autoarquivamento contribui para a transformação desse cenário, bem como esclarecimentos sobre a importância do depósito da produção científica institucional.

As publicações científicas depositadas no RIDI são licenciadas conforme os termos das licenças *Creative*

*Commons* (CC). A *Creative Commons* é uma organização sem fins lucrativos que visa ampliar o uso legal e compartilhado de obras protegidas por direitos autorais, permitindo que autores concedam permissões específicas para o uso de seus trabalhos. Existem seis tipos de licenças CC, e cabe ao autor selecionar a mais adequada à sua publicação, respeitando eventuais restrições impostas por periódicos ou editoras.

O RIDI é um repositório interoperável, permitindo a coleta automática de seus metadados por provedores de serviços e sistemas agregadores nacionais e internacionais, como a Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), o Portal Oasisbr, o Repositório Científico de Acesso Aberto de Portugal (RCAAP), a *Red Latinoameriana para la Ciencia Abierta* (LA Referencia), o OpenAIRE, entre outros.

A gestão do repositório prevê ainda a possibilidade de edição ou exclusão de publicações em casos específicos. Isso pode ocorrer por solicitação dos próprios autores, quando identificadas inconsistências ou restrições legais à divulgação, ou por demanda de terceiros, em situações que envolvam alegações de violação de direitos autorais ou de privacidade.

## HISTÓRICO

No contexto da Ciência Aberta e do Movimento de Acesso Aberto à Informação Científica (MAA), o Ibict tem se destacado, desde o início dos anos 2000, como protagonista na criação de políticas, serviços e produtos voltados à ampliação do acesso à produção científica. A

criação do RIDI é uma dessas iniciativas, concebida para assegurar a preservação, o acesso aberto e a disseminação da produção científica vinculada à instituição, em consonância com os princípios da Ciência Aberta.

O desenvolvimento do RIDI resultou de estudos e articulações anteriores a 2014, ano em que o repositório foi formalmente instituído e disponibilizado à comunidade científica do Ibit. Desde então, passou a integrar ecossistemas nacionais e internacionais de informação e pesquisa científica, fortalecendo a presença institucional nesses ambientes e promovendo maior visibilidade à produção do Instituto.

Café *et al.* (2003) destaca a implementação de repositórios institucionais como uma solução moderna diante dos mecanismos pagos e de acesso restrito que limitam a circulação do conhecimento científico. Duas décadas depois, os repositórios permanecem como instrumentos fundamentais para apoiar o trabalho dos pesquisadores e divulgar os resultados das pesquisas de maneira ampla, aberta e sustentável.

A Ciência Aberta, por sua vez, propõe mudanças estruturais na forma como o conhecimento científico é produzido, organizado, compartilhado e reutilizado. É um novo modo de fazer ciência, mais colaborativo, transparente e sustentável. Ao implementar o repositório com base na perspectiva dos arquivos abertos, onde os pesquisadores publicam seus trabalhos por meio do autoarquivamento inicia-se toda uma estrutura colaborativa.

O trabalho colaborativo é parte essencial para o fomento da Ciência Aberta, tanto para definição de diretrizes quanto para a disseminação dessas práticas. Entre os pilares constituintes da Ciência Aberta, têm-se os repositórios digitais, os quais são constituídos para o compartilhamento de publicações avaliadas por pares e também de dados provenientes das pesquisas (Sousa et al., 2024).

Nesta linha colaborativa, o RIDI foi implementado em DSpace, software livre desenvolvido em código aberto e mundialmente utilizado para construção e gerenciamento de repositórios digitais. Em 2025 iniciou-se o processo de atualização do RIDI, que migrará da versão 5.3 para a versão 7.6 do DSpace.

## **CONSIDERAÇÕES FINAIS**

O Ibict reafirma seu compromisso com a preservação e o acesso de longo prazo às publicações científicas depositadas no RIDI, alinhado às melhores práticas internacionais de preservação digital e certificação de repositórios confiáveis. Há mais de duas décadas o Instituto tem atuado em diferentes frentes de pesquisa, projetos e mediações de estudos que visam promover a Ciência Aberta brasileira por meio dos repositórios digitais e outras infraestruturas igualmente importantes.

A constituição e preservação do RIDI são fundamentais para a disponibilização de informações científicas da comunidade institucional, e como tal, torna-

se um exemplo de boas práticas de sustentabilidade para sociedade científica brasileira.

Durante o ano de 2025 o RIDI tem passado por uma significativa atualização tecnológica, que inclui desde a atualização da versão do software até alterações no padrão de metadados, melhorias de interface, padrões de acessibilidade digital, etc., garantindo a modernização e a eficiência do repositório.

## REFERÊNCIAS

CAFÉ, Lígia *et al.* Repositórios institucionais: nova estratégia para publicação científica na Rede. In: CONGRESSO BRASILEIRO DE CIÊNCIA DA COMUNICAÇÃO, 26., 2003, Belo Horizonte. **Anais** [...]. Belo Horizonte: Sociedade Brasileira de Estudos Interdisciplinares da Comunicação, 2003.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Repositório Institucional do Instituto Brasileiro de Informação em Ciência e Tecnologia**: RIDI, 2024. Disponível em: <https://ridi.ibict.br>. Acesso em: 01 maio 2024.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). Portaria nº 43, de 10 de setembro de 2014. Dispõe sobre a política para o repositório institucional da produção científica do Instituto Brasileiro de Informação em Ciência e Tecnologia. **Boletim de Serviço [do Ministério da Ciência, Tecnologia e Inovação]**, Brasília, n. 17, p. 84-88, 15 set. 2014.

LEITE, Fernando César Lima. **Como gerenciar e ampliar a visibilidade da informação científica brasileira: repositórios institucionais de Acesso Aberto**. Brasília: Ibict, 2009.

SOUZA, Juliana Araújo *et al.* Red Brasileña de Repositorios Digitales (RBRD): análisis de su constitución y representatividad a través del portal OASISBR. **Integración y Conocimiento**, Córdoba, v. 13, n. 1, p. 34-48, 1 feb. 2024. DOI: 10.61203/2347-0658.v13.n1.44207.

### Como citar este capítulo

MORAIS, Cássio Teixeira de. Repositório Institucional do Ibit (RIDI). In: AMARO, Bianca; CAMPOS, Phillipe de Freitas; BARCELOS, Janinne. (org.). **Infraestruturas de Ciência e de Acesso Aberto no Brasil**: iniciativas do Instituto Brasileiro de Informação em Ciência e Tecnologia. Brasília, DF: Editora Ibit, 2025. Cap. 4, p. 95-105. DOI: 10.22477/9788570132543.cap4



# REPOSITÓRIO COMUM DO BRASIL (DEPOSITA)

Blena Estevam dos Santos  
Millena Cordeiro Matos de Lima  
Cássio Teixeira de Moraes



## APRESENTAÇÃO

O Repositório Comum do Brasil (Deposita) foi desenvolvido pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) e teve sua primeira versão lançada oficialmente em 2017. O Deposita é uma base de dados dedicada a reunir e disponibilizar as produções científicas de pesquisadores brasileiros vinculados a instituições que ainda não possuem repositórios institucionais próprios. Além disso, também serve como um recurso para pesquisadores brasileiros que realizam pós-graduação em outros países, permitindo que suas publicações fiquem acessíveis à comunidade científica nacional. (Campos *et al.*, 2023; Ibict, 2025).

A criação do Deposita está diretamente relacionada à Portaria nº 13, de 15 de fevereiro de 2006, da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), que instituiu a obrigatoriedade de divulgação digital das teses e dissertações produzidas em programas de pós-graduação reconhecidos pela instituição (Brasil, 2006). Embora essa portaria tenha sido revogada posteriormente pela Portaria nº 182, de 14 de agosto de 2018 - que regulamenta a avaliação dos programas de pós-graduação stricto sensu - a exigência de visibilidade e acesso aberto à produção científica manteve-se como diretriz institucional (Brasil, 2018a). Complementarmente, a Portaria nº 289, de 28 de dezembro de 2018, também da Capes, passou a exigir que bolsistas no exterior publiquem os resultados de seus projetos em revistas de acesso aberto (Brasil, 2018b).

Seguindo os preceitos da Via Verde, uma das estratégias recomendadas pelo Movimento de Acesso Aberto para incentivar o depósito de pesquisas científicas em repositórios digitais, em complemento à Via Dourada, onde o acesso aberto é promovido pelas revistas científicas (Leite, 2009). O Deposita promove o autoarquivamento das produções acadêmicas e contribui para superar barreiras no acesso ao conhecimento científico brasileiro. Sua atuação é especialmente relevante para pesquisadores vinculados a instituições que ainda não implementaram repositórios próprios, garantindo-lhes uma alternativa segura e pública para disseminação de suas pesquisas.

O nome Deposita foi escolhido para incentivar os pesquisadores brasileiros a depositarem os resultados de seus trabalhos acadêmico-científicos. A escolha do termo buscou transmitir, de forma direta, a finalidade do serviço. Já a designação inicial de "Repositório Comum" foi substituída por não refletir adequadamente os objetivos do projeto, nem gerar identificação com a comunidade acadêmica.

A produção científica disponível no Deposita torna-se amplamente visível na internet, ampliando seu alcance e impacto. Os documentos depositados no repositório são automaticamente coletados pelo Portal Brasileiro de Publicações e Dados Científicos em Acesso Aberto (Oasisbr), iniciativa que reúne e integra conteúdos de repositórios de instituições brasileiras de ensino e pesquisa. No caso específico de teses e dissertações, esses registros também são coletados pela Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), contribuindo para a preservação, disseminação e acesso

público desses trabalhos. Dessa forma, o Deposita garante maior visibilidade, interoperabilidade e acesso aberto à produção acadêmico-científica, beneficiando tanto os pesquisadores quanto a sociedade em geral.

Assim, o principal objetivo do Deposita é ampliar o acesso e a visibilidade da produção científica brasileira, além de promover e incentivar a criação de repositórios institucionais em universidades e centros de pesquisa que ainda não dispõem desse tipo de infraestrutura. Os benefícios do Deposita são diversos, tanto para os pesquisadores quanto para o fortalecimento da Ciência Aberta no país. Dentre esses benefícios, destacam-se:

#### **Benefícios para o pesquisador:**

- a) Disponibiliza livremente sua produção científica na Internet;
- b) Torna a sua produção mais visível pelos motores de busca na Internet;
- c) Possibilita o aumento das citações e do impacto da produção;
- d) Garante a preservação da produção em um ambiente digital; e
- e) Oferece um endereço eletrônico confiável e único para a produção, garantindo o acesso permanente.

#### **Benefícios para a ciência brasileira:**

- a) Organiza e amplia a sua visibilidade;
- b) Permite que a produção científica realizada no Brasil seja acessada de qualquer lugar do

- mundo e a qualquer hora de forma livre e sem nenhum custo para o usuário;
- c) Preserva a produção científica realizada no Brasil;
  - d) Fornece dados para a gestão da ciência brasileira; e
  - e) Promove a criação de repositórios institucionais, que permitirão a gestão organizacional dos recursos de pesquisa.

Em 2025, o repositório Deposita foi atualizado para a versão 7.6.2 do software DSpace, trazendo melhorias significativas na interface e na experiência de navegação. A nova versão permite a exploração do acervo por meio de filtros como autor, área do conhecimento, assunto, instituição, data e idioma de publicação, facilitando o acesso às informações de forma mais eficiente. O Deposita reúne uma variedade de tipologias documentais, com destaque para artigo de revista, artigo de evento, trabalho de conclusão de curso, dissertação, tese, livro e capítulo de livro. O número de registros por tipo documental é dinâmico, refletindo as constantes contribuições dos pesquisadores. À medida que as instituições passam a contar com seus próprios repositórios institucionais, os documentos duplicados entre sistemas são identificados e removidos do Deposita, a fim de evitar redundância de registros e garantir a confiabilidade da base de dados.

A arquitetura informacional do Deposita é baseada no software DSpace, sendo que em sua atual versão o repositório apresenta a comunidade denominada “Deposita” e suas coleções baseadas nas tipologias

documentais aceitas no repositório. Essa estrutura facilita a organização, a busca e a recuperação das informações, promovendo uma navegação mais intuitiva e alinhada aos padrões internacionais de repositórios digitais.

Figura 1 - Arquitetura da informação

### Arquitetura da informação do Deposita publicações



Fonte: os autores (2025).

Ao acessar o portal ou realizar o cadastro, o usuário deve estar ciente da política do repositório e da licença de uso adotada. Os documentos disponibilizados no Deposita são distribuídos sob uma licença *Creative Commons* (CC BY-NC-SA) de distribuição não-exclusiva, que concede ao repositório o direito de reproduzir, traduzir e disseminar os conteúdos em diferentes formatos e meios. Ficam reservados aos autores todos os direitos morais, bem como os usos comerciais sobre as obras de sua autoria, salvo as exceções previstas em lei ou em instrumentos contratuais.

Entende-se que as Licenças *Creative Commons* constituem um conjunto de ferramentas jurídicas voltadas para o compartilhamento de conteúdos digitais, especialmente aqueles de cunho não comercial,

permitindo aos autores definirem com clareza os termos de uso e disseminação de suas obras. Segundo Shintaku e Meirelles (2010), trata-se de um recurso que promove a liberdade de disseminação, ao mesmo tempo em que garante os direitos autorais dos criadores.

As licenças *Creative Commons* oferecem diferentes modalidades de licenciamento, combinando tipos de permissões com condições específicas de uso. As permissões definem o que é permitido fazer com a obra, como compartilhar (copiar, distribuir e transmitir) ou remixar (criar obras derivadas). Já as condições estabelecem os critérios para esse uso, podendo incluir: atribuição (obrigatoriedade de crédito ao autor); uso não comercial (proibição de uso com fins lucrativos); vedação à criação de obras derivadas (proibição de alterações ou transformações); e compartilhamento pela mesma licença (exigência de que obras derivadas sejam distribuídas sob a mesma licença original ou equivalente) (Shintaku; Meirelles, 2010).

## HISTÓRICO

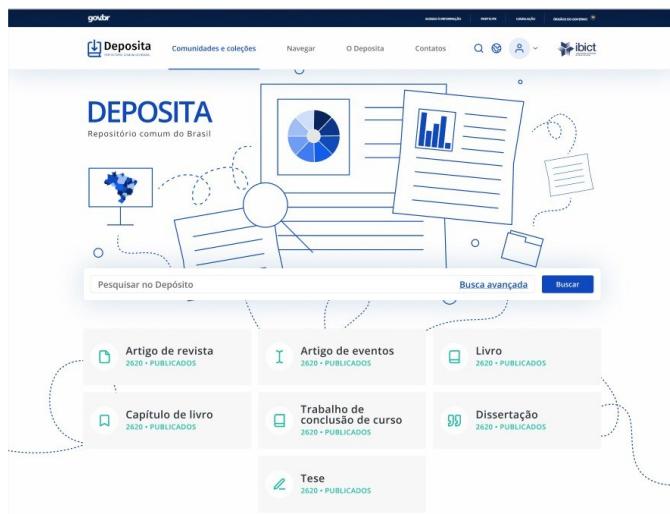
O Deposita começou a ser concebido em 2014, a partir de estudos e análises de iniciativas internacionais alinhadas ao Movimento de Acesso Aberto, como o Repositório Comum de Portugal, integrado ao portal RCAAP (Repositórios Científicos de Acesso Aberto de Portugal). Diante dessas referências e da ausência, no Brasil, de uma plataforma nacional voltada ao depósito de publicações científicas por pesquisadores sem repositórios institucionais disponíveis, identificou-se a necessidade de

criar um ambiente que incentivasse essa prática. Ainda em 2014, o projeto foi apresentado internacionalmente durante a 5<sup>a</sup> Conferência Luso-Brasileira sobre Acesso Aberto (ConfOA), realizada na Universidade de Coimbra, em Portugal.

O Deposita foi oficialmente lançado em 2017, ano em que os primeiros registros começaram a ser depositados no sistema. Um levantamento retrospectivo realizado com base na ferramenta *Internet Archive Wayback Machine* permite acompanhar a evolução do repositório desde 19 de junho de 2017, quando cinco documentos estavam disponíveis. Ainda em 2017, esse número subiu para oito documentos; em 2018, para 18; em 2019, para 30; em 2020, alcançou 116; em 2022, eram 240 documentos; em 2023, o número ultrapassava 332; em 2024, já somava 474; e até abril de 2025, o repositório contava com mais de 700 documentos depositados.

Esses dados evidenciam um crescimento gradual entre 2017 e 2019, seguido por um aumento mais expressivo a partir de 2020, possivelmente refletindo uma maior consolidação da plataforma e do compromisso institucional com a ciência aberta. Ainda conforme verificado por meio do Wayback Machine, a interface do Deposita manteve-se relativamente estável ao longo dos anos, com poucas alterações visuais — como a adição de logomarcas e links externos —, preservando sua estrutura principal. No entanto, em 2025, o site passou por uma grande atualização, que modernizou significativamente sua interface e aprimorou a experiência de navegação para os usuários.

Figura 2 -Tela inicial: Deposita



Fonte: Captura de tela do Deposita (Ibict, 2025).

Para acessar o Deposita e realizar o autoarquivamento de sua produção científica, o pesquisador deve entrar no repositório por meio do endereço eletrônico: [deposita.ibict.br](http://deposita.ibict.br). No ambiente do sistema, o usuário preencherá os metadados do documento conforme a tipologia documental selecionada, seguindo as orientações apresentadas em cada etapa do formulário de submissão.

## CONSIDERAÇÕES FINAIS

No contexto da Ciência Aberta, o Deposita constitui uma importante ferramenta para ampliar a visibilidade da produção científica de pesquisadores brasileiros, tanto oriundos de instituições nacionais quanto daqueles que realizaram sua formação no exterior. Por meio do autoarquivamento, o repositório contribui para o aumento da circulação e do acesso dessas produções, favorecendo sua preservação, o incremento de citações e a democratização do acesso à ciência.

Com sete anos de funcionamento, o repositório já reúne uma diversidade de documentos, como trabalhos de conclusão de curso, dissertações, teses e artigos, disponíveis em acesso aberto. Desde seu lançamento observa-se um crescimento contínuo no número de depósitos, o que reforça sua consolidação como espaço relevante para a disseminação do conhecimento científico.

O Deposita também exerce um papel estratégico e político, uma vez que reafirma a importância dos repositórios digitais e estimula instituições que ainda não implementaram seus próprios repositórios digitais a avançarem nessa direção. Dessa forma, além de preservar a memória da ciência nacional, o repositório contribui para fortalecer a cultura do acesso aberto no Brasil, ampliando o debate sobre a importância dos repositórios institucionais e promovendo a adoção de boas práticas em comunicação científica.

## REFERÊNCIAS

BRASIL. Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). **Portaria nº 13, de 15 de fevereiro de 2006.** Institui a divulgação digital das teses e dissertações produzidas pelos programas de doutorado e mestrado reconhecidos. Brasília: Capes, 2006.

BRASIL. Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). **Portaria nº 182, de 14 de agosto de 2018.** Dispõe sobre processos avaliativos das propostas de cursos novos e dos programas de pós-graduação stricto sensu em funcionamento. Brasília: Capes, 2006.

BRASIL. Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). **Portaria nº 289, de 28 de dezembro de 2018.** Institui o Regulamento que estabelece as normas para as modalidades de bolsas e auxílios no exterior. Brasília: Capes, 2006.

CAMPOS, Phillipe de Freitas et al. Contribución del Instituto Brasileño de Información en Ciencia y Tecnología a la promoción del acceso abierto y la ciencia abierta: análisis de sus infraestructuras. **Revista Científica**, Bogotá, v. 48, n. 3, sept./dic. 2023. DOI: 10.14483/23448350.20924.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Manual de depósito de obras no Deposita - Repositório Comum do Brasil.** Brasília: Ibitc, 2025. Disponível em: <https://deposita.ibict.br/info/deposit>. Acesso em: 11 jun. 2025.

LEITE, Fernando César Lima. **Como gerenciar e ampliar a visibilidade da informação científica brasileira: repositórios institucionais de Acesso Aberto.** Brasília: Ibitc, 2009.

SHINTAKU, Milton; MEIRELLES, Rodrigo. **Manual do DSpace:** administração de repositórios. Salvador: EDUFBA, 2010.

### Como citar este capítulo

SANTOS, Blena Estevam dos; LIMA, Millena Cordeiro Matos de; MORAIS, Cássio Teixeira de. Repositório Comum do Brasil (Deposita). In: AMARO, Bianca; CAMPOS, Phillippe de Freitas; BARCELOS, Janinne. (org.). **Infraestruturas de Ciência e de Acesso Aberto no Brasil:** iniciativas do Instituto Brasileiro de Informação em Ciência e Tecnologia. Brasília, DF: Editora Ibitc, 2025. Cap. 5, p. 107-118. DOI: 10.22477/9788570132543.cap5

# ALEIA: REPOSITÓRIO DE DADOS DE PESQUISA DO IBICT

Tatyane Guedes Martins da Silva  
Leticia Guarany Bonetti



O repositório de dados de pesquisa Aleia é uma das iniciativas do Instituto Brasileiro de Informação em Ciência e Tecnologia para promoção da informação científica e tecnológica nacional e internacional. Sua primeira versão foi oficialmente lançada em 2023 no XXII Seminário Nacional de Bibliotecas Universitárias (SNBU), em Florianópolis. Em maio de 2025, o Aleia conquistou a certificação internacional *CoreTrustSeal*, a mais adotada para a validação de repositórios digitais confiáveis. A certificação consolida o repositório como um repositório digital de alta confiabilidade para a comunidade científica do Instituto. O processo de certificação foi iniciado em 2023 e demandou o cumprimento de 16 requisitos que abrangem áreas como infraestrutura informacional, gestão de objetos digitais, tecnologia da informação e segurança, alinhando-se aos Princípios FAIR (*Findability, Accessibility, Interoperability e Reusability*). O Aleia tornou-se o segundo repositório de dados da América Latina a obter esse selo, atrás apenas do Arca Dados, da Fiocruz, destacando-se entre os 159 repositórios certificados no mundo (Ibict, 2025).

O Aleia surgiu da necessidade de apoiar e facilitar a adoção do acesso aberto aos dados de pesquisa, fomentar novas pesquisas e promover práticas de Ciência Aberta. Sua criação foi motivada pelo Compromisso 3, do 4º Plano de Ação Nacional que visava estabelecer mecanismos de governança de dados científicos para o avanço da Ciência Aberta no Brasil (*Open Government Partnership*, 2018).

Sanchez, Vidotti e Vechiato (2017, p. 3) descrevem os repositórios de dados como aqueles que "[...] buscam organizar, estruturar, permitir acesso, disseminar e

preservar todos os dados gerados por meio de pesquisas realizadas em sua maioria por Instituições de Ensino e Pesquisa". De acordo com Sayão e Sales (2015, p. 49), "Eles constituem o lugar mais apropriado para que seus dados sejam preservados e possam ser recuperados, acessados e citados por outros pesquisadores, ou seja, tenham visibilidade em escala mundial".

Sabe-se que os dados de pesquisa possuem um nível de abstração maior que outros objetos digitais como artigos, dissertações, teses e livros, o que pode prejudicar a interpretação, reutilização e citação adequada. Entretanto, os dados, quando reutilizados, podem trazer benefícios como: 1) transparência e prevenção de fraudes; 2) economia de recursos financeiros e humanos; 3) agilização do ciclo científico; 4) reproducibilidade, dentre outros. Por isso é fundamental o seu armazenamento em repositórios, uma vez que:

Sem uma descrição minuciosa do contexto tecnológico dos arquivos de dados, do contexto no qual os dados foram criados ou coletados, das medidas que foram feitas, dos detalhes espaciais e temporais, dos instrumentos usados, dos parâmetros e unidades e da qualidade dos dados e da sua proveniência, é improvável que os dados possam ser descobertos, interpretados, gerenciados e efetivamente usados e reusados (Sayão; Sales, 2015, p. 20).

Nota-se, portanto, a importância da descrição adequada dos dados, garantindo sua contextualização, o que é feito por meio da representação por metadados em

infraestruturas tecnológicas como os repositórios. A gestão adequada dos dados de pesquisa, de acordo com Sayão e Sales (2016, p. 92), tem como componente central os repositórios de dados que "[...] rapidamente se tornam parte essencial da infraestrutura mundial de pesquisa", o que pode ser ligado a duas demandas essenciais no contexto atual: "[...] as políticas mandatórias das agências financeiras de pesquisa e a incorporação pelas instituições de pesquisa e pela sociedade em geral dos valores e princípios preconizados pela Ciência Aberta".

O Aleia é um serviço voltado para o depósito de dados que foram coletados e produzidos pela comunidade científica do Ibict durante a pesquisa técnica ou científica, sendo comumente aceitos na comunidade para validar as descobertas e resultados de um estudo. Assim, seu principal objetivo é ser o repositório oficial do Ibict que custodia, publica, dissemina, preserva e compartilha conjuntos de dados da comunidade científica do Ibict constituída por seus servidores, pesquisadores visitantes, bolsistas e alunos do programa de pós-graduação (Ibict, 2024).

O repositório foi desenvolvido ao longo de oito etapas, sendo elas: 1) estudos para escolha do software em que seria desenvolvido (optou-se pelo Dataverse); 2) instalação do software escolhido em máquina de teste; 3) definição dos metadados; 4) definição das temáticas a serem definidas como comunidades; 5) criação e aprovação do design do repositório; 6) desenvolvimento dos conteúdos para página do repositório; 7) inclusão dos conteúdos no repositório; e 8) migração do repositório de ambiente de teste para ambiente de produção.

No que diz respeito à arquitetura informacional, o Aleia foi estruturado da seguinte forma: 1) Sobre; 2) Guia do usuário; 3) Política do repositório; 4) Termos; 5) Perguntas frequentes; 6) Contato; 7) Comunidades Dataverse, sendo elas: Tecnologias da Informação, Ensino e Pesquisa, Informação Científica, Informação para Sustentabilidade, Informação Tecnológica, Popularização da Ciência e Informação para a Sociedade; 8) Conjuntos de Dados; 9) Arquivos.

A arquitetura da informação foi definida pela equipe técnica do Aleia com o objetivo de agrupar assuntos gerais que mais são demandados pela comunidade científica do Ibict. Os formatos disponibilizados no Aleia são planilhas, imagens, vídeos, gráficos, arquivos tabulares, entre outros. Devem estar em formato natos digitais ou digitalizados, completos e prontos para disseminação.

Como mencionado anteriormente, o Aleia foi desenvolvido no Dataverse, software de código aberto direcionado para armazenamento e compartilhamento de dados de pesquisa. O Dataverse foi desenvolvido em 2006 pelo Instituto de Ciências Sociais Quantitativas da Universidade de Harvard em colaboração com profissionais de todo o mundo, por meio do *Dataverse Project* (Dataverse, 2023).

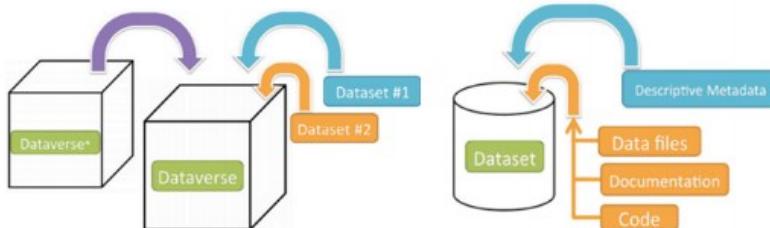
A essência do *Dataverse Project* é automatizar grande parte do trabalho do arquivista profissional, fornecer serviços e distribuir crédito ao criador de dados. Há a união da preservação dos dados e dos créditos ao autor, o que influenciou na escolha do Dataverse para ambos os repositórios de dados gerenciados pelo Ibict: Aleia e Deposita Dados. Além disso, já fala-se sobre como “[...]

uma plataforma de repositório como o Dataverse pode facilitar muito a criação de dados científicos FAIR" (Silva; Rodrigues, 2021, p. 127). Corroborando com os autores, Boyd (2021, p. 5, tradução nossa) afirma que "A plataforma de software Dataverse oferece suporte a dados FAIR por meio de uma combinação de funcionalidade de infraestrutura padrão e recursos necessários e opcionais voltados para o usuário". Isso inclui, por exemplo, identificadores persistentes, metadados específicos do domínio e suporte para formatos de arquivo abertos.

Os princípios FAIR são um acrônimo para "*Findable*", "*Accessible*", "*Interoperable*" e "*Reusable*", que em português significa "Encontrável", "Acessível", "Interoperável" e "Reutilizável". São diretrizes que buscam enfatizar o aprimoramento da capacidade das máquinas de encontrar e usar os dados automaticamente, além de apoiar sua reutilização por usuários humanos. Com o volume cada vez maior de dados sendo produzidos e processados nas pesquisas, o que demanda o uso de computação intensiva, é essencial se pensar em formas de tornar os dados científicos mais FAIR, justamente pela otimização do processamento automático. De acordo com Sales *et al.* (2020), os princípios FAIR já são reconhecidos mundialmente como elementos-chave para boas práticas em todos os processos de gestão de dados. Logo, a adoção do Dataverse está consoante com essa tendência internacional.

O Aleia hospeda vários arquivos virtuais chamados Comunidades Dataverses. Cada comunidade contém conjuntos de dados chamados datasets, e cada dataset contém objetos digitais (arquivos) e metadados.

Figura 1 - Estrutura do Dataverse



Fonte: Crosas (2015).

Cabe destacar, ainda, que o Dataverse investe em uma série de boas práticas que podem ser consultadas, na íntegra, no site oficial (Ibict, 2024). Alguns exemplos de suma importância para os repositórios geridos pelo Ibict são:

- 1) crédito acadêmico: ao depositar, os pesquisadores tornam seus conjuntos de dados mais localizáveis na web. Um dos motivos é a oferta do recurso *widgets* pelo Dataverse, que permite que o pesquisador incorpore os conjuntos em qualquer site para ajudar outros pesquisadores a encontrá-los com mais facilidade;
- 2) citação dos dados: o Projeto Dataverse padroniza a citação de conjuntos de dados para tornar mais fácil para os pesquisadores publicarem seus dados e obterem crédito, bem como reconhecimento por seu trabalho. Quando se cria um conjunto de dados em um repositório Dataverse, a citação é gerada e apresentada automaticamente;
- 3) gestão dos dados: muitos periódicos, editores e agências de financiamento exigem que os

pesquisadores depositem conjuntos de dados em um repositório público. O Projeto Dataverse ajuda os pesquisadores a cumprir esse requisito, apoiando o depósito de conjuntos de dados e tornando-os facilmente detectáveis para outros pesquisadores reutilizarem e verificarem se um estudo pode ser replicado sem ter que, necessariamente, entrar em contato com os autores do estudo.

Nota-se, portanto, que a adoção do Dataverse traz uma série de benefícios para os depositantes, trazendo maior segurança para os pesquisadores que desejam compartilhar seus dados. Com isso em mente, o repositório Aleia permite o armazenamento, a gestão e a preservação dos dados de pesquisa de diferentes comunidades acadêmicas, apoiando e facilitando também a adoção do acesso aberto aos dados de pesquisa, fomentando novas pesquisas e promovendo uma Ciência Aberta, indo de acordo com o Manifesto de Acesso Aberto a Dados de Pesquisa Brasileira para Ciência Cidadã lançado em 2016.

Para acessar e depositar no Aleia, basta acessar o repositório em seu endereço eletrônico: aleia.ibict.br. Ao clicar na aba “Iniciar sessão”, clique no botão “ORCID”.

O usuário já estará logado e pronto para iniciar o depósito dos conjuntos de dados que deseja compartilhar. O depositante deve escolher a comunidade de acordo com o assunto principal do conjunto de dados, lembrando que o povoamento/alimentação do Aleia deverá ser realizado por autoarquivamento, ou seja, caberá ao autor realizar o depósito do seu próprio conjunto de dados, que passará pela revisão da equipe gestora do Aleia.

Figura 2 - Tela inicial



Fonte: (Iibict, 2024).

Os arquivos devem estar em formato natos digitais ou digitalizados, completos e prontos para disseminação. E, sempre que possível, devem ser escolhidos formatos que ajudem a preservar e garantir o acesso a longo prazo aos arquivos depositados no Aleia, com preferência por arquivos em formatos abertos e não-proprietários, o que evita barreiras de software.

## CONSIDERAÇÕES FINAIS

Finalmente, almeja-se que o repositório possa atender à comunidade científica do Iibict, desempenhando o papel de uma fonte de informação confiável e segura. O Aleia tem como finalidade tornar os dados de pesquisa acessíveis, dar visibilidade aos dados de pesquisa

elaborados pela comunidade científica Ibjct; garantir a reprodutibilidade e reuso dos dados de pesquisa da comunidade científica institucional do Ibjct; funcionar como ferramenta para promover a transparência e por consequência a confiança e a solidez das pesquisas desenvolvidas, maximizando o retorno de aplicação de recurso público.

## REFERÊNCIAS

BOYD, Ceilynn. Use of Optional Data Curation Features by Users of Harvard Dataverse Repository. **Journal of eScience Librarianship**, Worcester, v. 10, n. 2, p. 1-20, 2021. DOI:

10.7191/jeslib.2021.1191.

CROSAS, Merce. **Dataverse on the MOC**. 2015. Disponível em: <https://pt.slideshare.net/mercecrosas/dataverse-on-the-moc>. Acesso em: 12 jun. 2024. (Apresentação slides).

DATAVERSE. **The Dataverse Project**. 2023. Disponível em: <https://dataverse.org>. Acesso em: 15 jan. 2024.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Aleia**. Disponível em: <https://aleia.ibict.br>. Acesso em: 12 jun. 2024.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Repositório Aleia conquista certificação internacional CoreTrustSeal, que reconhece repositórios digitais confiáveis**. 2025. Disponível em: [www.gov.br/ibict/pt-br/central-de-conteudos/noticias/2025/junho/repositorio-aleia-conquista-certificacao-internacional-coretrustseal-que-reconhece-repositorios-digitais-confiaveis](http://www.gov.br/ibict/pt-br/central-de-conteudos/noticias/2025/junho/repositorio-aleia-conquista-certificacao-internacional-coretrustseal-que-reconhece-repositorios-digitais-confiaveis). Acesso em: 11 jun. 2025.

**OPEN GOVERNMENT PARTNERSHIP. 4º Plano de Ação**

**Nacional em Governo Aberto.** Brasília: Ministério da Transparéncia e Controladoria-Geral da União, 2018.

SALES, Luana et al. GO FAIR Brazil: A Challenge for Brazilian Data Science. **Data Intelligence**, Cambridge, v. 2, n. 1-2, p. 238-245, 2020. DOI: 10.1162/dint\_a\_00046.

SANCHEZ, Fernanda Alves; VIDOTTI, Silvana Aparecida Borsetti Gregório; VECCHIATO, Fernando Luiz. A contribuição da curadoria digital em repositórios digitais. **Revista Informação na Sociedade Contemporânea**, Natal, v. 1, p. 1-17, 2017. DOI: 10.21680/2447-0198.2017v1n0ID12280.

SAYÃO, Luis Fernando; SALES, Luana Farias. Algumas considerações sobre os repositórios digitais de dados de pesquisa. **Informação & Informação**, Londrina, v. 21, n. 2, p. 90-115, maio/ago., 2016. DOI: 10.5433/1981-8920.2016v21n2p90.

SAYÃO, Luis Fernando; SALES, Luana Farias. **Guia de gestão de dados de pesquisa para bibliotecários e pesquisadores**. Rio de Janeiro: CNEN, 2015.

SILVA, Fabiano Couto Corrêa da; RODRIGUES, Marcello Mundim. Implementação dos princípios FAIR em repositórios de dados científicos: uma análise comparativa das infraestruturas de software do DSpace e Dataverse. In: SALES, Luana Farias; VEIGA, Viviane dos Santos de Oliveira; HENNING, P Patrícia; SAYÃO, Luis Fernando (org.). **Princípios FAIR aplicados à gestão de dados de pesquisa**. Rio de Janeiro: Ibiti, 2021, p. 123-136.

### Como citar este capítulo

SILVA, Tatyane Guedes Martins da; BONETTI, Leticia Guarany. Aleia: Repositório de dados de pesquisa do Ibict. *In: AMARO, Bianca; CAMPOS, Phillip de Freitas; BARCELOS, Janinne. (org.). Infraestruturas de Ciência e de Acesso Aberto no Brasil: iniciativas do Instituto Brasileiro de Informação em Ciência e Tecnologia.* Brasília, DF: Editora Ibict, 2025. Cap. 6, p. 119-130. DOI: 10.22477/9788570132543.cap6

# **REPOSITÓRIO COMUM DE DADOS DE PESQUISA: DEPOSITA DADOS**

Tatyane Guedes Martins da Silva  
Leticia Guarany Bonetti



O repositório de dados de pesquisa Deposita Dados é uma das iniciativas do Instituto Brasileiro de Informação em Ciência e Tecnologia para promoção da informação científica e tecnológica nacional e internacional. Sua primeira versão foi oficialmente lançada em 2023 no XXII Seminário Nacional de Bibliotecas Universitárias (SNBU), em Florianópolis.

O repositório de dados é uma plataforma online que contribui para o gerenciamento e o compartilhamento de dados de pesquisa, capazes de “[...] representar, processar, gerenciar, disseminar e dar acesso aos dados nele depositados, incluindo métodos de curadoria” (Monteiro; Sant’ana, 2016, p. 658).

O Deposita Dados é uma ferramenta que reúne e disponibiliza os dados de pesquisa de pesquisadores de instituições brasileiras que ainda não possuem repositórios de dados próprios. Além disso, é utilizado pelos pesquisadores brasileiros que executaram seus conjuntos de dados por meio de colaboração científica em instituições estrangeiras de ensino e pesquisa.

O Deposita Dados foi criado e desenvolvido por meio de estudos e análises de iniciativas similares mundialmente, com o objetivo de tornar mais acessível e visível os dados de pesquisa dos pesquisadores brasileiros, como também promover e incentivar a criação de repositórios em universidades e institutos de pesquisa que ainda não desenvolveram esse serviço em suas instituições. O Deposita Dados destina-se aos mais diversos atores envolvidos no processo de comunicação científica, estimulando a ampla circulação do

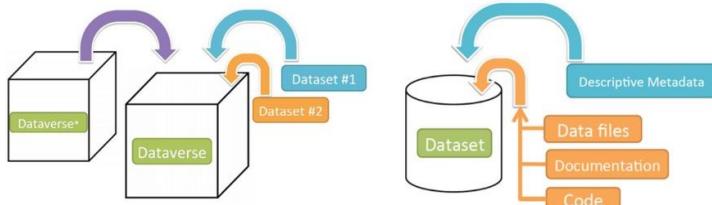
conhecimento científico. Sua utilização gera os seguintes benefícios:

- Ao pesquisador:
  - a) Disponibiliza livremente seus conjuntos de dados de pesquisa na internet.
  - b) Torna os seus conjuntos de dados mais visíveis pelos motores de busca na internet.
  - c) Possibilita o aumento das citações e o impacto da produção.
  - d) Garante a preservação dos seus conjuntos de dados em um ambiente digital.
  - e) Assegurar acesso permanente aos conjuntos de dados por meio do identificador persistente DOI.
  - f) Promove um meio para que o pesquisador apresente os conjuntos de dados coletados com recurso financeiro investido na pesquisa científica.
- Às instituições de ensino e pesquisa:
  - a) Amplia a visibilidade dos conjuntos de dados gerados pela produção científica no contexto institucional.
  - b) Promove a criação de repositórios de dados nas instituições que ainda não possuem este serviço, por meio da gestão dos depósitos dos conjuntos de dados.
- À ciência brasileira:
  - a) Permite que os conjuntos de dados gerados durante o processo de produção científica no Brasil seja acessada de qualquer lugar do mundo e a qualquer

- hora de forma livre e sem nenhum custo para o usuário.
- b) Fornece dados para a gestão da ciência e instituições brasileiras.

Atualmente, o repositório se encontra na versão do software Dataverse 5.11.1. O Deposita Dados hospeda vários arquivos virtuais chamados Comunidades Dataverses. Cada comunidade contém conjuntos de dados chamados *datasets*, e cada *dataset* contém objetos digitais (arquivos) e metadados.

Figura 1 - Estrutura do Dataverse



Fonte: Crosas (2015).

A interface utilizada permite a navegação por ano de publicação, autor, assunto e data de depósito, como também é possível a navegação específica por regiões do Brasil, com comunidades organizadas de acordo com as cinco regiões: Norte, Nordeste, Centro-Oeste, Sudeste e Sul, além de uma comunidade destinada para pesquisadores brasileiros que executaram seus conjuntos de dados por meio de colaboração científica em instituições estrangeiras de ensino e pesquisa: Pesquisas do exterior.

O Dataverse consiste em um software livre de código aberto, que vem sendo amplamente utilizado para construir e gerenciar repositórios de dados de pesquisa, pois disponibiliza vários recursos personalizáveis para atender à necessidade do usuário. O software é focado no armazenamento, acesso e preservação de dados de pesquisa e suporta uma variedade de tipos e formatos de arquivos, incluindo: imagens, gráficos, planilhas, dados audiovisuais, arquivos tabulares, entre outros (Dataverse, 2023).

Rocha et al. (2021, p. 8) explicam que o Dataverse "[...] possui recursos que possibilitam variadas configurações de ambientes de repositório de dados". É possível, portanto, a personalização do ambiente digital, se adequando melhor a diferentes comunidades e públicos alvo. Além disso, "O Dataverse também permite que metadados e arquivos que compõem *datasets* sejam extraídos do ambiente via interface de programação (API), facilitando a preservação a longo prazo" (Rocha et al., 2021, p. 10). Considerando a rápida obsolescência dos objetos digitais e o valor dos dados de pesquisa para os avanços científicos, a preservação torna-se essencial.

Considerado, portanto, como componente central da gestão dos dados de pesquisa, os repositórios de dados estão atrelados com vários benefícios quando se fala no compartilhamento de conjuntos de dados como: 1) preservação digital, 2) créditos ao autor, 3) memória científica e transparência, 4) segurança dos dados, 5) indicação da qualidade e produtividade da instituição, 6) curadoria digital, 7) reutilização dos dados, dentre outros

(Sayão; Sales, 2016). Logo, a criação do repositório Deposita Dados segue nessa linha, ciente de que:

[...] é imprescindível a conscientização entre os pesquisadores sobre a importância da disponibilização dos seus dados científicos, a criação de políticas que incentivem essa prática, o desenvolvimento tecnológico e a capacitação dos pesquisadores com ferramentas que facilitem o processo (Araújo; Dias; Autran, 2021, p. 102).

Para acessar e depositar no Deposita Dados, basta acessar o repositório em seu endereço eletrônico: [depositadados.ibict.br](http://depositadados.ibict.br), clicar na aba “Iniciar sessão” e logo em seguida clicar no botão “ORCID”. O usuário já estará logado e pronto para iniciar o depósito dos conjuntos de dados que deseja compartilhar. O depositante deve escolher a comunidade de acordo com a localização da sua instituição, lembrando que o povoamento do Deposita Dados deverá ser realizado por autoarquivamento, ou seja, caberá ao autor realizar o depósito do seu próprio conjuntos de dados.

Caso não tenha se cadastrado no ORCID, clique em “Registrar agora” e preencha os campos solicitados.

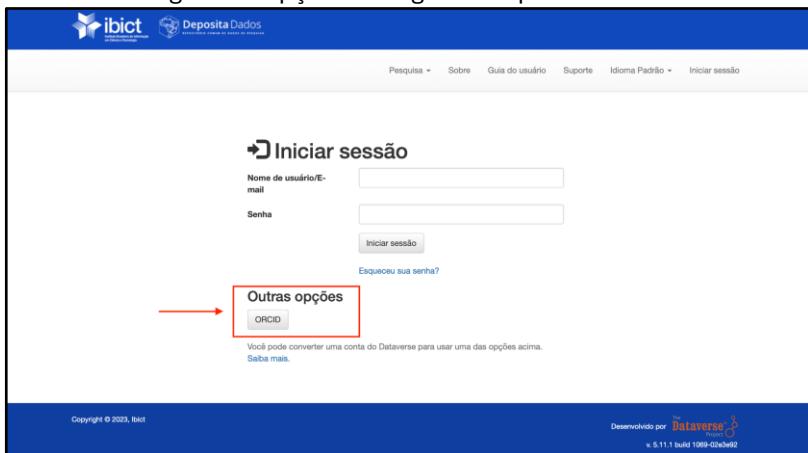
Os arquivos devem estar em formato natos digitais ou digitalizados, completos e prontos para disseminação. E, sempre que possível, devem ser escolhidos formatos que ajudem a preservar e garantir o acesso a longo prazo aos arquivos depositados no Deposita Dados, com preferência por arquivos em formatos abertos e não-proprietários, o que evita barreiras de software.

Figura 2 - Iniciar sessão no Deposita Dados



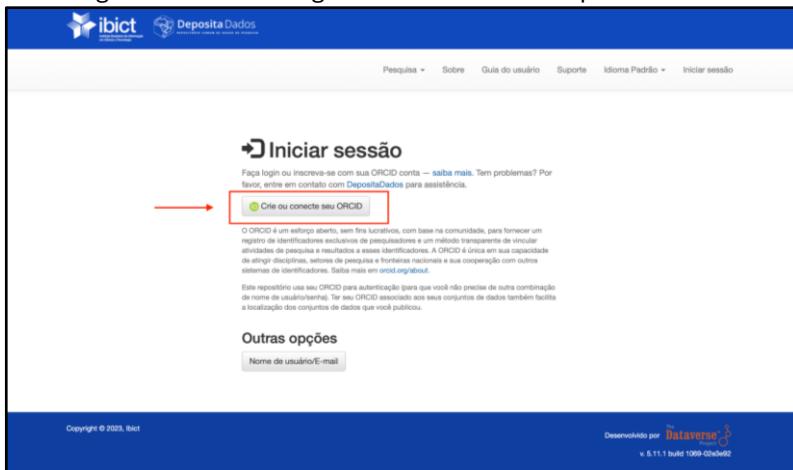
Fonte: Ibitc (2024).

Figura 3 - Opções de login no Deposita Dados



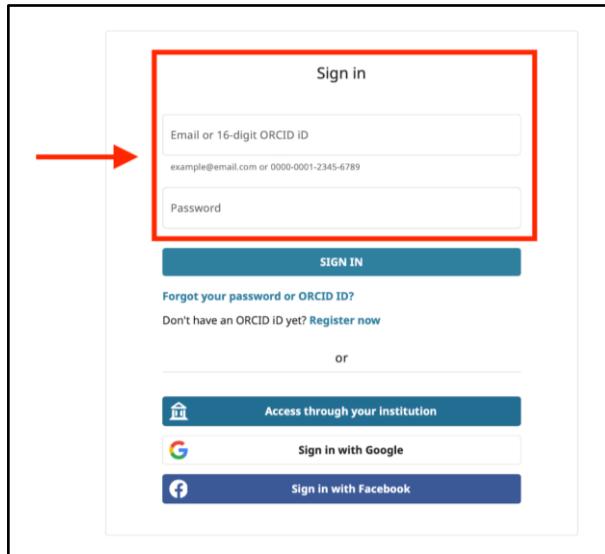
Fonte: Ibitc (2024).

Figura 4 - Fazendo login com o ORCID no Deposita Dados



Fonte: Ibitc (2024).

Figura 5 - Login com o ORCID no Deposita Dados



Fonte: Ibitc (2024).

Para conjuntos de dados que envolvem pesquisas com seres humanos, de qualquer natureza e área do conhecimento, os pesquisadores devem certificar-se e garantir que todos os preceitos e procedimentos éticos foram seguidos. Os dados pessoais devem ser anonimizados ou pseudonimizados antes do depósito, seguindo a Lei Geral de Proteção de Dados (Lei nº 13.709/18), que define anonimização como a "[...] utilização de meios técnicos razoáveis e disponíveis no momento do tratamento, por meio dos quais um dado perde a possibilidade de associação, direta ou indireta, a um indivíduo" (Brasil, 2018, Art. 5º, XI).

## **CONSIDERAÇÕES FINAIS**

Finalmente, almeja-se que o repositório possa atender pesquisadores brasileiros vinculados a instituições científicas que ainda não possuem seus repositórios de dados de pesquisa e/ou de pesquisadores brasileiros que executaram seus conjuntos de dados por meio de colaboração científica em instituições estrangeiras de ensino e pesquisa, desempenhando o papel de uma fonte de informação confiável e segura. O Deposita Dados tem papel fundamental para a promoção e consolidação da prática de Ciência Aberta ao prover o compartilhamento acessível dos dados de pesquisa, fortalecer a reprodutibilidade, transparência e incentivar o progresso científico, pois permite que pesquisadores ao redor do mundo acessem e utilizem os dados disponibilizados nesses ambientes digitais, facilitando a colaboração e o avanço coletivo do conhecimento científico.

## REFERÊNCIAS

ARAÚJO, Débora Gomes de; DIAS, Guilherme Ataíde; AUTRAN, Marynica de Medeiros Matos. Compartilhamento de dados no contexto da ciência brasileira: um estudo integrativo.

**Informação & Informação**, Londrina, v. 26, n. 3, p. 100-124, 2021. DOI: 10.5433/1981-8920.2021v26n3p100.

BRASIL. Lei n. 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). **Diário Oficial da União**, Brasília, 15 ago. 2018.

CROSAS, Merce. **Dataverse on the MOC**. 2015. Disponível em: <https://pt.slideshare.net/mercecrosas/dataverse-on-the-moc>. Acesso em: 12 jun. 2024. (Apresentação slides).

DATAVERSE. **The Dataverse Project**. 2023. Disponível em: <https://dataverse.org>. Acesso em: 15 jan. 2024.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Deposita Dados**. Disponível em: <https://depositadados.ibict.br>. Acesso em: 12 jun. 2024.

MONTEIRO, Elizabeth Cristina de Souza de Aguiar; SANT'ANA, Ricardo Cesar Gonçalves. Repositório de dados científicos nas universidades brasileiras e portuguesas. In: SEMINÁRIO EM CIÊNCIA DA INFORMAÇÃO, 6., 2016. **Anais** [...]. Londrina: UEL, 2016.

ROCHA, Rafael Port da et al. Análise dos sistemas DSpace e Dataverse para repositórios de dados de pesquisa com acesso aberto. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v. 17, p. 1-25. 2021.

SAYÃO, Luís Fernando; SALES, Luana Farias. Algumas considerações sobre os repositórios digitais de dados de pesquisa. **Informação & Informação**, Londrina, v. 21, n. 2, p.

90-115, maio/ago., 2016. DOI: 10.5433/1981-8920.2016v21n2p90.

### Como citar este capítulo

SILVA, Tatyane Guedes Martins da; BONETTI, Leticia Guarany. Repositório Comum de Dados de Pesquisa: Deposita Dados *In:* AMARO, Bianca; CAMPOS, Fhillipe de Freitas; BARCELOS, Janinne. (org.). **Infraestruturas de Ciência e de Acesso Aberto no Brasil:** iniciativas do Instituto Brasileiro de Informação em Ciência e Tecnologia. Brasília, DF: Editora Ibict, 2025. Cap. 7, p. 131-141. DOI: 10.22477/9788570132543.cap7



# DIRETÓRIO DE POLÍTICAS EDITORIAIS DAS REVISTAS CIENTÍFICAS BRASILEIRAS (DIADORIM)

Phillipe de Freitas Campos

Bianca Amaro

Denise Aparecida Freitas de Andrade



## APRESENTAÇÃO

O Diretório de políticas editoriais das revistas científicas brasileiras (Diadorim) é uma das iniciativas do Ibict que remonta à chegada do Movimento de Acesso Aberto à Informação Científica (MAA) no Brasil. Tendo sua primeira versão oficialmente lançada em 2012, a construção do Diretório tornou-se possível por meio de recursos recebidos da Financiadora de Estudos e Projetos (Finep) (Costa; Alves, 2019).

O Diadorim é um serviço de informações relativas às autorizações concedidas para o armazenamento e o acesso dos artigos das revistas brasileiras em repositórios digitais de acesso aberto. Assim, seu principal objetivo é identificar, sistematizar e disponibilizar as informações das políticas editoriais estabelecidas pelas revistas científicas brasileiras referentes ao armazenamento e acesso dos artigos em repositórios institucionais de acesso aberto (Campos *et al.*, 2023).

Assim, evidencia-se a interlocução promovida pelo Diadorim entre a Via Dourada e a Via Verde do MAA, que pode ser explicada da seguinte forma: os pesquisadores utilizam preferencialmente as revistas científicas (Via Dourada) para fazer a publicação inicial de seus resultados de pesquisa. Com a ascensão do MAA, surgem os repositórios institucionais (Via Verde), que tem como uma de suas funcionalidades resguardar a produção intelectual da instituição na qual foram criados. Todavia, há de se considerar que a partir do momento que um artigo é submetido e/ou publicado em uma revista científica passam a incidir sobre ele uma série de questões

relacionadas a direitos autorais, acesso e uso. Assim, para que um repositório institucional possa fazer o depósito deste artigo e disponibilizá-lo livremente à sua comunidade de usuários, é importante que seja verificado previamente quais são as políticas editoriais definidas pela revista, de modo a não infringir qualquer questão relativa aos direitos autorais e correlatos.

No que diz respeito à política editorial, os dados providos pelo Diadorim fazem uma descrição minuciosa das políticas estabelecidas pela revista. Todavia, essa descrição pode ser resumida no campo “Selo de armazenamento e acesso”. Utilizando uma categorização internacional por cores, este campo informa qual é a política da revista para armazenamento e acesso dos artigos submetidos e/ou publicados em repositórios institucionais de acesso aberto. No Quadro 1 é possível visualizar o resumo desta categorização.

Evidencia-se, portanto, que o Diadorim desempenha um papel crucial na promoção da transparência e conformidade com as políticas editoriais das revistas científicas brasileiras, contribuindo para a integridade e acessibilidade da produção científica nacional no contexto do movimento de acesso aberto.

Quadro 1 - Classificação dos selos do Diadorim

<b>Cor</b>	<b>Definição</b>	<b>Selo novo</b>
Verde	Permite o armazenamento e acesso das versões <i>preprint</i> e <i>pós-print</i> dos documentos em repositórios institucionais/digitais	

Cor	Definição	Selo novo
Azul	Permite o armazenamento e acesso das versões <i>pós-print</i> dos documentos em repositórios institucionais/digitais	
Amarelo	Permite o armazenamento e acesso das versões <i>preprint</i> dos documentos em repositórios institucionais/digitais	
Branco	Apresenta restrições para o armazenamento e acesso das versões <i>preprint</i> e <i>pós-print</i> dos documentos em repositórios institucionais/digitais	

Fonte: os autores (2025).

## HISTÓRICO

A proposição deste serviço se inicia entre os anos de 2010 e 2011, mas os primeiros resultados somente são apresentados em 2012, quando são publicados dois trabalhos por pesquisadoras do Instituto apontando para a necessidade de um serviço brasileiro que auxiliasse na interlocução entre a Via Verde e a Via Dourada do MAA. O primeiro deles foi apresentado no 3º Simpósio Brasileiro de Comunicação Científica (SBCC) e tem como título “Serviços de difusão de políticas para o acesso aberto: Sherpa/Juliet, Melibea e ROARMAP”. Ao passo que apresentam serviços internacionais com função similar ao que se pretendia ter com o Diadorim, as autoras enfatizam

que “nos três serviços analisados foram identificadas apenas três políticas brasileiras para a promoção do acesso aberto”. Desde modo, enfatizam que tal realidade “[...] aponta para a necessidade de criação de iniciativas similares à desses serviços [...] que poderia contribuir para a criação de políticas nacionais para o acesso aberto que sejam coerentes à realidade nacional [...]” (Costa; Amaro; Assis, 2012, p. 52).

Na sequência, em ocasião da 3<sup>a</sup> Conferência Luso-Brasileira sobre Acesso Aberto (ConfOA), é apresentado o segundo trabalho, cujo título é “Diretório de políticas editoriais brasileiras sobre o acesso aberto por meio dos repositórios: o serviço Diadorim”. Nesta publicação, já são apresentados os primeiros resultados do serviço em operação, que naquele então contava com 215 políticas editoriais das revistas científicas brasileiras. As autoras apontam ainda que um dos principais benefícios do serviço é a possibilidade dele “tornar-se uma mola propulsora para o povoamento dos repositórios institucionais e instrumentalizar as discussões sobre acesso aberto” (Costa; Assis; Silva; Amaro, 2012, n. p.).

Desse momento em diante, o Diretório entra em operação, de modo que a força de trabalho deixa de estar concentrada em sua construção para se concentrar no povoamento, trabalho esse que ocorreu majoritariamente por contato direto da equipe gestora com os editores científicos brasileiros, tanto por telefone quanto por e-mail. Durante este período o Diretório se fortaleceu junto à comunidade editorial brasileira, chegando no final do ano de 2014 com o registro da política editorial de 736 revistas. Já em 2015 é quando o Diretório passa por suas primeiras

alterações, neste caso especificamente em relação ao *layout* da página inicial. Tais alterações tiveram por principal objetivo propiciar aos usuários uma interface mais agradável e com algumas informações sobre as políticas editoriais disponíveis na tela inicial.

No ano seguinte o Diretório passa por alterações mais substanciais, especialmente em relação ao seu nome, que passa ser “Diretório de políticas editoriais das revistas científicas brasileiras”, nome esse que mantém até o presente momento. Essa mudança foi acarretada especialmente por conta de uma questão prática acerca das revistas científicas brasileiras: ainda que o Diretório tenha sido criado tendo como um de seus objetivos fomentar a aplicação do Movimento de Acesso Aberto no Brasil, nem todas as revistas nele cadastradas são de acesso aberto. Assim, a retirada do termo torna o serviço mais abrangente e fidedigno. A Figura 1 apresenta a interface do Diadorim já com seu novo nome.

Figura 1 - Interface do Diadorim com novo nome

The screenshot shows the Diadorim website interface. At the top is a logo with a yellow sun-like icon and the word "Diadorim" in orange. Below it is the text "Diretório de Políticas Editoriais das Revistas Científicas Brasileiras". The main search interface consists of a search bar labeled "Buscar no Diadorim" and a magnifying glass icon. Below the search bar are three filter sections: "Revista" (with a plus sign), "Editora" (with a plus sign), and "Área do conhecimento" (with a plus sign). Each section has a dropdown menu with options and counts: "Disponibilidade para depósito" (Azul: 163, Verde: 197, Amarelo: 36, Branca: 11), "Tipo de acesso" (Acesso aberto: 105, Acesso aberto após período de emb.: 34, Acesso Restrito: 19), and "Licença de conteúdo".

Fonte: os autores (2025).

De lá até o final de 2023 as alterações de interface no Diretório foram pontuais e sem grandes impactos. Todavia, sinaliza-se que foram feitas mudanças “invisíveis” para o usuário externo, a saber: a) alterações no padrão de metadados adotado; e b) atualização da versão do software para a versão 6.3 do DSpace. Destaca-se que durante todo este período o foco principal da equipe técnica foi concentrado em sua popularização e disseminação à comunidade científica e editorial, de modo que no final do ano de 2023 o Diretório contava com 3192 revistas.

Em janeiro de 2024 inicia-se um projeto maior para modernização do Diadorim, em virtude do lançamento de um novo serviço do Ibict, o Diretório das revistas científicas eletrônicas brasileiras (Miguilim), que será apresentado nos capítulos seguintes. Ocorre que o lançamento do novo serviço possibilitou às revistas espaço para maior detalhamento de seus dados, que vão além dos metadados abarcados pelo Diadorim. Frente a este cenário, a equipe técnica - que responde por ambos os serviços - optou por fazer modificações mais incisivas no Diadorim, objetivando otimizar tanto o trabalho dos editores quanto da própria equipe. Deste modo, iniciou-se um processo de interoperabilidade entre os dois serviços, de modo que os dados das revistas constantes no Diadorim serão importados diretamente do Miguilim. Pontua-se que para viabilizar essa atualização foi necessário modificar o software que hospeda o Diadorim, de modo que ele deixou de utilizar o DSpace para utilizar o Vufind. Na Figura 2 é possível visualizar a tela inicial da nova versão do Diadorim

Figura 2 - Interface atual do Diadorim



Fonte: os autores (2025)

## CONSIDERAÇÕES FINAIS

O MAA é uma realidade no contexto da comunicação científica mundial, que se estruturou por meio da interlocução entre suas duas vias originárias: Via Dourada e Via Verde. Nesse contexto, o Ibict nota a necessidade de criar um serviço capaz de mediar e facilitar a relação entre essas duas Vias, criando assim o Diadorim.

Com mais de 10 anos em pleno desenvolvimento, o Diadorim tem se fortalecido e procurado pela comunidade científica e editorial brasileira, que o utiliza prioritariamente para registrar e ter acesso à política editorial das revistas no que diz respeito ao armazenamento e acesso dos artigos em repositórios institucionais de acesso aberto.

Desde sua criação até meados do primeiro semestre de 2024 o serviço funcionou de maneira autônoma. Todavia, a partir da criação de um novo serviço pelo Ibict evidencia-se a necessidade de completa reformulação do Diadorim, a qual entrou em fase de execução a partir do segundo semestre de 2024, com interface mais moderna e novas formas de registrar as políticas editoriais.

## REFERÊNCIAS

CAMPOS, Phillip de Freitas et al. Contribución del Instituto Brasileño de Información en Ciencia y Tecnología a la promoción del acceso abierto y la ciencia abierta: análisis de sus infraestructuras. **Revista Científica**, Bogotá, v. 48, n. 3, p. 56-66, 2023. DOI: 10.14483/23448350.20924.

COSTA, Michelli Pereira da; ALVES, Larissa de Araújo. Diretórios de políticas editoriais sobre o acesso aberto. In: SHINTAKU, Milton; SALES, Luana Farias. **Ciência Aberta para editores científicos**. São Paulo: Abec Brasil, 2019. DOI: 10.21452/978-85-93910-02-9. p. 41-49.

COSTA, Michelli Pereira da; AMARO, Bianca; ASSIS, Tainá Batista de. Serviços de difusão de políticas para o acesso aberto: Sherpa/Juliet, Melibea e ROARMAP. **Encontros Bibli**, Florianópolis, p. 36-54, 2012. DOI: 10.5007/1518-2924.2012v17nesp2p36.

COSTA, Michelli Pereira da; ASSIS, Tainá Batista de; SILVA, Rayanne; AMARO, Bianca. Diretório de políticas editoriais brasileiras sobre o acesso aberto por meio dos repositórios: o serviço Diadorim. In: CONFERÊNCIA LUSO-BRASILEIRA SOBRE ACESSO ABERTO, 3., 2012, Lisboa. **Anais** [...]. Lisboa: Universidade de Lisboa, 2012.

## Como citar este capítulo

CAMPOS, Phillip de Freitas; AMARO, Bianca; ANDRADE, Denise Aparecida Freitas de. Diretório de políticas editoriais das revistas científicas brasileiras (Diadorim). *In:* AMARO, Bianca; CAMPOS, Phillip de Freitas; BARCELOS, Janinne. (org.). **Infraestruturas de Ciência e de Acesso Aberto no Brasil:** iniciativas do Instituto Brasileiro de Informação em Ciência e Tecnologia. Brasília, DF: Editora Ibict, 2025. Cap. 8, p. 143-152. DOI: 10.22477/9788570132543.cap8

# **PORTAL BRASILEIRO PARA AS REVISTAS CIENTÍFICAS (MANUELZÃO)**

Denise Aparecida Freitas de Andrade  
Phillipe de Freitas Campos  
Raphael Faria Vilas Boas



## APRESENTAÇÃO

O Portal brasileiro para revistas científicas (*Manuelzão*) surgiu da necessidade de centralizar informações, produtos e serviços relacionados à editoração de revistas científicas. A iniciativa foi motivada pela equipe do Ibict em 2019, que após anos de pesquisa e desenvolvimento sobre revistas científicas identificou que havia uma lacuna no processo de gerenciamento destas publicações. Eles apontaram que, em geral, as equipes editoriais lideradas pelos editores-chefes das revistas raramente possuíam formação adequada para este tipo de atividade, ainda que sejam pesquisadores de excelência em suas áreas de atuação (Amaro, Campos, Vilas Boas, 2022).

Sobre a carência deste tipo de formação, vale trazer a visão de Gomes (2010, p. 157) ao dizer que “[...] são praticamente inexistentes as oportunidades e espaços de formação de editores, tanto no âmbito de graduação como de pós-graduação”. Não obstante, o autor aponta que “Não há uma formação específica profissional para o editor de revistas científicas, função normalmente ocupada por pesquisadores da área sem a necessária formação técnica para promover ou coordenar processos editoriais como um todo”. Ainda que se tenha passado mais de 10 anos do citado estudo, a realidade foi pouco modificada.

Estudos mais recentes publicados por editores de periódicos, como o de Melo, Manduca e Santos (2018) e Formiga (2024), enfatizam o anseio por instrumentos de apoio, orientação e investimento voltados à qualificação editorial das revistas. Para os autores, a falta de

treinamento editorial adequado e contínuo para as equipes editoriais é um dos maiores desafios enfrentados pelos periódicos científicos. Nesse mesmo sentido, Esteves, Sartori e Teixeira (2023, p. 4) apontam que

a qualificação dos periódicos científicos tem sido uma tarefa árdua, sobretudo ao considerar os diferentes parâmetros que incidem sobre a avaliação final do Qualis do periódico.

De fato, há iniciativas profícuas que tentam dirimir esta lacuna deixada pela não formação profissional dos editores científicos, como aquelas desempenhadas pela Associação Brasileira de Editores Científicos (Abec), que promove cursos, palestras, seminários, workshops, dentre outros eventos visando à capacitação dos editores. Todavia, nem sempre essas ações atingem grandes públicos, ainda mais se comparada à quantidade de revistas científicas existentes no Brasil.

Frente ao contexto, o Manuelzão surge como uma iniciativa que tem por missão reunir e disseminar informações editoriais de fácil compreensão que auxiliem de modo efetivo o trabalho dos editores e equipes editoriais em sua atuação em uma revista científica. Trata-se, portanto, de um serviço dirigido a editores científicos e equipes editoriais.

A escolha do nome Manuelzão foi inspirada na tendência internacional de nomear diretórios e serviços de informação científica homenageando obras literárias do país onde o sistema foi desenvolvido. Neste caso, o nome é uma referência à obra de João Guimarães Rosa, intitulada

"Manuelzão e Miguilim", publicada pela primeira vez em 1964.

## HISTÓRICO

O Manuelzão foi desenvolvido ao longo de sete etapas, elencadas pela equipe gestora do Portal na primeira publicação científica sobre ele:

1) estudos para escolha do software em que seria desenvolvido (optou-se ao final pelo wordpress); 2) instalação do software escolhido em máquina de teste; 3) definição das temáticas a serem abordadas; 4) criação, aprovação e implementação do design do Portal; 5) desenvolvimento dos conteúdos; 6) inclusão dos conteúdos desenvolvidos no ambiente do Portal e; 7) migração do Portal de ambiente de teste para ambiente de produção (Amaro, Campos, Vilas Boas, 2022).

No que tange à arquitetura informacional, o Manuelzão foi inicialmente estruturado em onze categorias que abrangem o universo editorial científico, sendo estas:

1) Produtos e serviços do Ibict; 2) Portais brasileiros de revistas científicas; 3) Como criar uma revista científica; 4) Acesso Aberto; 5) Ciência Aberta; 6) Avaliação por pares; 7) Dados de pesquisa; 8) Auto-diagnóstico de qualidade editorial; 9) Visibilidade e indexação; 10) Estratégias para internacionalização e; 11) Taxas de

publicação (Amaro, Campos, Vilas Boas, 2022).

A arquitetura informacional foi definida pela equipe desenvolvedora do Manuelzão por entender que ela é capaz de agrupar os conteúdos que mais são demandados pelos editores e equipes editoriais que procuram as equipes do Ibict. Dentro de cada uma destas classes são disponibilizados uma série de conteúdos que podem auxiliar os consultentes na editoração científica. Os conteúdos disponibilizados no Manuelzão são tanto autorais (elaborados pela própria equipe com base no trabalho já realizado junto às revistas científicas brasileiras) quanto provenientes de fontes externas (artigos científicos, livros, palestras, cursos online, etc.). Ressalta-se, por fim, que os conteúdos ora apresentados não são exaustivos, o que significa dizer que podem haver diversas outras fontes informacionais desconhecidas pela equipe desenvolvedora. Deste modo, dá-se destaque ao caráter colaborativo do Manuelzão, onde o público tem total abertura e liberdade para propor a inclusão de materiais ainda não disponíveis no Portal.

Em 2022, diante da interrupção do projeto devido a fatores externos, como a pandemia da Covid-19, a equipe do Manuelzão optou por reativá-lo. Após análises, os gestores optaram por reestruturar e atualizar o conteúdo do site, o qual foram redistribuídos em 7 classes que seguem descritas abaixo:

1. Produtos e serviços do Ibict para revistas científicas: o conteúdo desta categoria foi preservado conforme

o original, destacando os principais produtos e serviços desenvolvidos e mantidos pelo Ibict destinados às revistas científicas.

2. Como criar uma revista científica: essa seção delinea o processo de criação de uma revista. Anteriormente, estava organizada em cinco categorias - efetiva necessidade, políticas editoriais, software, ISSN, editoração. No entanto, após análises, foram incluídas novas categorias, sendo elas: efetiva necessidade; financiamento de uma revista; definição e estrutura da revista; equipe editorial; políticas editoriais; diretrizes para autores; plataforma de publicação; registro ISSN; preservação digital e indexação e divulgação.
3. Revistas científicas e Ciência Aberta: esta seção abrange conteúdos relacionados a acesso aberto, ciência aberta, dados de pesquisa e taxas de submissão.
4. Indexação e qualificação de revistas científicas: nesta seção, são abordados temas como auto diagnóstico de qualidade editorial, Qualis Periódicos, revistas predatórias, estratégias para internacionalização e visibilidade, e indexação.
5. Boas práticas editoriais: esta seção reúne recomendações voltadas à qualificação das revistas científicas brasileiras, com foco na adoção de práticas editoriais responsáveis, éticas e alinhadas aos princípios da Ciência Aberta.
6. Portais brasileiros de revistas científicas: apresenta breves descrições e exemplos dos portais brasileiros

de revistas científicas, bem como discute a importância dessas infraestruturas.

7. Conteúdos de interesse para o editor científico: esta seção reúne as principais notícias relacionadas à editoração científica e está organizada em quatro classes principais, a saber: sugestões de leitura, sugestões de produtos e serviços, sugestões de vídeos e eventos, sugestões de documentos. Além disso, os conteúdos são classificados de acordo com seus temas, que incluem categorias como revistas predatórias, editoração, inteligência artificial, entre outros. Essas categorias podem se expandir à medida que novos conteúdos são incluídos.

Adicionalmente foram efetuadas alterações na estrutura do código da página para cumprir as diretrizes de acessibilidade digital. As principais modificações incluíram a incorporação do *plugin* VLibras (Brasil, 2025), a modificação nas cores e no contraste do site, ajustes na cor e tamanho da fonte, a inclusão de texto alternativo em imagens e links, além da hierarquização dos cabeçalhos e textos. Após seu lançamento foram incorporadas novas funcionalidades com o objetivo de ampliar a disseminação das informações disponíveis no site, bem como de conteúdos relacionados à editoração científica.

Entre as melhorias implementadas, destaca-se a divulgação semanal de notícias oriundas de diversas plataformas, portais e iniciativas que contribuem para a qualificação e o fortalecimento da editoria de revistas científicas. Com o intuito de estreitar a comunicação entre

o Manuelzão e os editores, foi criado um boletim informativo, distribuído por meio do canal Google Grupos, facilitando o acesso contínuo às atualizações e conteúdos relevantes para a comunidade editorial. A Figura 1 apresenta a tela inicial do Manuelzão.

Figura 1 - Tela inicial do Manuelzão



Instituto Brasileiro de Informação  
em Ciência e Tecnologia

PÁGINA INICIAL NOTÍCIAS SOBRE O PORTAL CONTATO

# MANUELZÃO

Portal brasileiro para as revistas científicas

Informações, produtos e serviços para a editoração de revistas  
científicas em um só lugar

Pesquisar ...

PESQUISAR

## Últimas notícias



Toolkit para periódicos de  
acesso aberto: um guia  
para editores científicos

9 de junho de 2025

Acesso Aberto, Edição de  
revistas



Cartilha – Publicação  
Contínua em Periódicos  
Científicos: convite à  
reflexão editorial

4 de junho de 2025

Edição de revistas



Webinar – Melhoria  
contínua da qualidade dos  
periódicos na Redalyc: um  
processo aberto e  
iterativo

30 de maio de 2025

Bons práticas editoriais, Indexação e  
indexadores



Novo recurso do Diretório  
Migullum: gráfico de  
percentual de  
preenchimento dos  
metadados

28 de maio de 2025

Diretório Migullum, Indexação e  
indexadores

Inscrir-se na newsletter

Fonte: os autores (2025).

## CONSIDERAÇÕES FINAIS

Espera-se que o Portal atenda de forma eficaz às necessidades da comunidade acadêmica, científica e editorial, consolidando-se como uma fonte confiável de informação e contribuindo, assim, para o aprimoramento da qualidade editorial das revistas científicas brasileiras.

Ressalta-se que o Portal Manuelzão mantém o compromisso com a atualização contínua, incorporando novos conteúdos, recursos e funcionalidades em consonância com as transformações no cenário da comunicação e da editoração científica. A expectativa é de que o site continue sendo regularmente atualizado, mantendo-se relevante e alinhado às demandas do meio científico.

## REFERÊNCIAS

AMARO, Bianca; CAMPOS, Phillipe de Freitas; BOAS, Raphael Faria Vilas. Manuelzão e Miguilim: iniciativas do Ibict para os editores e revistas científicas brasileiras. *In: ABEC MEETING, 2022, São Paulo. Anais [...]*. São Paulo: Abec, 2022. DOI: 10.21452/abecmeeting2022.148.

BRASIL. Ministério da Gestão e Inovação em Serviços Públicos. **VLibras**. Brasília: MGI, 2025. Disponível em: [www.gov.br/governodigital/pt-br/vlibras](http://www.gov.br/governodigital/pt-br/vlibras). Acesso em: 11 ago. 2025.

ESTEVES, Pâmela; SARTORI, Leandro; TEIXEIRA, Marcelle. Revista Periferia: quatorze anos dedicados à Educação, Comunicação e Cultura. *In: ABEC MEETING 2023, Foz do Iguaçu. Anais [...]*. Foz do Iguaçu: Abec, 2023. DOI: 10.21452/abecmeeting2023.198.

FORMIGA, Magno. De repente, editor-chefe: desafios de um jovem doutor na revitalização de um periódico científico. *In: ABEC MEETING, 2024, São Paulo. Anais [...]*. São Paulo: Abec, 2024. DOI: 10.21452/abecmeeting2024.222.

GOMES, Valdir Pereira. O editor de revista científica: desafios da prática e da formação. **Informação & Informação**, Londrina, v.

15, n. 1 p. 147-172, 2010. DOI: 10.5433/1981-8920.2010v15n1p147.

MELO, Paula Lorena Silva; MANDUCA, Marcelo; SANTOS, Elisama Nunes dos. Ferramentas e desafios na gestão do processo editorial de revistas via Open Journal Systems: relato de caso. *In: ABEC MEETING, 2., 2018, São Paulo. Anais [...]*. São Paulo: Abec, 2018. p. 1-5. DOI: 10.21452/abecmeeting.2018.186.

### Como citar este capítulo

ANDRADE, Denise Aparecida Freitas de; CAMPOS, Fhillipe de Freitas; VILAS BOAS, Raphael Faria. Portal brasileiro para as revistas científicas (Manuelzão). *In: AMARO, Bianca; CAMPOS, Fhillipe de Freitas; BARCELOS, Janinne. (org.). Infraestruturas de Ciência e de Acesso Aberto no Brasil: iniciativas do Instituto Brasileiro de Informação em Ciência e Tecnologia*. Brasília, DF: Editora Ibict, 2025. Cap. 9, p. 153-162. DOI: 10.22477/9788570132543.cap9

# DIRETÓRIO DAS REVISTAS CIENTÍFICAS ELETRÔNICAS BRASILEIRAS (MIGUILIM)

Denise Aparecida Freitas de Andrade  
Phillipe de Freitas Campos  
Bianca Amaro



## APRESENTAÇÃO

O Diretório das revistas científicas eletrônicas brasileiras (Miguilim) é um serviço cujo objetivo central é “agregar, em um único local, informações sobre as revistas científicas editadas e publicadas no Brasil que se encontravam dispersas em diferentes plataformas” (Ibict, 2024). Além disso, busca-se por meio do Diretório atingir outros objetivos, tais como: facilitar o acesso às revistas científicas brasileiras, aumentar sua visibilidade internacional, promover transparência editorial, disseminar boas práticas, instruir editores sobre critérios de avaliação, incentivar pesquisas em comunicação científica, servir como porta de entrada para outros produtos do Ibict, evitar retrabalho de editores, promover padronização de dados e fomentar ações relacionadas à Ciência Aberta e Acesso Aberto. Estes esforços são fundamentais para fortalecer a pesquisa e a comunicação científica no país, bem como para promover a inserção internacional da produção científica nacional.

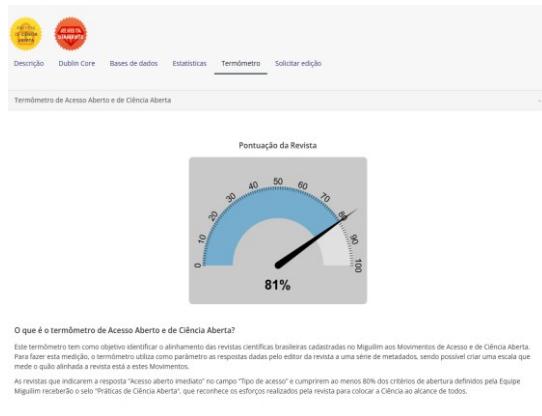
O Diretório é composto por duas coleções principais: a coleção de revistas científicas, que possui 77 campos descritivos abrangendo dados cadastrais da revista e elementos da política editorial; e a coleção de Portais de revistas, com 14 campos descritivos sobre os portais, incluindo dados cadastrais sobre a instituição e a lista de revistas que dele fazem parte (Ibict, 2025).

O registro das informações pode ser realizado de duas maneiras: por meio do pré-cadastro feito pela equipe do Miguilim, que coleta em lote os dados das revistas de outras bases de dados e posteriormente permite a

atualização do registro pelos editores científicos; ou por meio do registro direto no sistema. Em ambos os casos, cabe à equipe gestora apenas a validação e eventuais correções nos dados informados. É importante ressaltar que o formulário é auto declaratório, sendo responsabilidade do editor preencher corretamente as informações sobre as políticas editoriais das revistas.

Dentre as finalidades do Miguilim destaca-se a criação do “Termômetro de Acesso Aberto e de Ciência Aberta”, que avalia o alinhamento das revistas aos Movimentos de Acesso e de Ciência Aberta. Esse recurso utiliza as respostas fornecidas pelos editores das revistas a uma série de metadados e atribui pontuações para cada uma delas. A Figura 1 apresenta a aplicação do termômetro no Miguilim.

Figura 1 - Registro no Miguilim com o “Termômetro de Acesso Aberto e de Ciência Aberta”



Fonte: os autores (2025).

Por consequência, todas as revistas que indicam ser de “Acesso aberto imediato” e atingem a partir de 80% da

pontuação total aferida pelo termômetro recebem um selo denominado “Práticas de Ciência Aberta”. Esse selo tem como objetivo destacar revistas comprometidas com a adoção efetiva dos princípios da Ciência Aberta.

Outra funcionalidade do Miguilim é o selo “Revista Diamante”, atribuído às revistas que adotam o modelo de Acesso Aberto conforme os princípios estabelecidos pela BOAI. Recebem este selo todas as revistas que afirmam oferecer acesso aberto imediato e que não cobram quaisquer taxas de processamento editorial, seja de submissão ou de publicação. Adicionalmente, o Diretório conta com o selo “Associação à Abec Brasil”, criado a partir da parceria entre o Ibict e a Associação Brasileira de Editores Científicos (Abec Brasil). Esse selo foi concebido com o intuito de fortalecer a colaboração entre as instituições e permitir à Abec Brasil o acesso facilitado aos dados das revistas associadas, ampliando a visibilidade e o reconhecimento dessas publicações no cenário nacional.

A Figura 2 apresenta um registro com todos os selos atribuídos:

Figura 2 - Registro no Miguilim com os três selos atribuídos

Revista Brasileira de Cancerologia (online)					
Use este identificador para citar ou linkar para este item: <a href="https://miguilim.ibict.br/handle/miguilim/3483">https://miguilim.ibict.br/handle/miguilim/3483</a>					
Descrição	Dublin Core	Bases de dados	Estatísticas	Termômetro	Solicitar edição
Descrição:					
Título:	Revista Brasileira de Cancerologia (online)				
Título abreviado:	Rev. Bras. Cancerol. (online)				
Título próprio:	Revista Brasileira de Cancerologia				
ISSN:	2176-9745				

Fonte: os autores (2025).

Além disso, o Miguilim disponibiliza um painel de indicadores que reúne informações sobre diversos aspectos das revistas científicas cadastradas na plataforma. Seu principal objetivo é oferecer aos pesquisadores, editores e demais interessados subsídios para compreender o panorama das publicações científicas no país, além de incentivar ações voltadas à melhoria da qualidade editorial, visibilidade e internacionalização dessas revistas. O painel também desempenha um papel estratégico do ponto de vista gerencial, ao fornecer dados que orientam e qualificam o trabalho de curadoria da base.

O Miguilim conta também com uma funcionalidade que calcula o percentual de preenchimento das informações de cada revista cadastrada. Essa ferramenta tem como objetivo indicar o grau de completude dos registros, sem, contudo, exigir o preenchimento integral de todos os campos, uma vez que nem todas as informações se aplicam a todas as revistas. Espera-se, no mínimo, que sejam preenchidos os 48 campos obrigatórios. A partir de 62% de preenchimento, a revista já é considerada como tendo um bom nível de completude em seu registro, o que contribui para a qualidade e a confiabilidade das informações disponíveis no Diretório.

## HISTÓRICO

As proposições para criação do Miguilim remontam aos anos de 2017-2018, momento em que se percebe a maturidade do Diretório de políticas editoriais das revistas científicas brasileiras (Diadorim), que teve um aumento considerável no número de revistas registradas e passou a

ser amplamente utilizado pela comunidade editorial brasileira. Apesar disso, não era incomum os gestores do Diadorim receberem demandas relativas às revistas e precisarem consultar outros serviços de informação para obtê-las. Neste cenário é que surge a ideia de expandir o serviço por meio da criação de um padrão de metadados mais robusto e com novas funcionalidades, capazes de melhor retratar as revistas científicas brasileiras.

Todavia, a expansão vislumbrada não poderia ser feita no Diadorim tendo em vista a especificidade de seu objetivo. Concomitantemente, nota-se uma ascensão notável do Movimento de Ciência Aberta no Brasil, indicando uma transformação significativa na cultura científica mundial e, por conseguinte, demandando que as revistas se ajustem ao novo paradigma em vigor. Diante deste cenário é que se inicia a criação do Miguilim.

O processo de desenvolvimento do Diretório foi dividido em diversas etapas e sub-etapas, descritas por Amaro, Campos e Vilas Boas (2022) em uma das primeiras publicações científicas sobre o Miguilim:

[...] no que diz respeito ao Miguilim, a metodologia de desenvolvimento seguiu as seguintes etapas: 1) estudos para escolha do software em que seria desenvolvido (optou-se ao final pelo DSpace em sua versão 6.3); 2) instalação do DSpace em máquina de teste; 3) definição do padrão de metadados do Diretório; 4) criação, aprovação e implementação do design do Diretório; 5) implementação do padrão de metadados em ambiente de teste; 6) coleta de informações de revistas científicas eletrônicas brasileiras

de diferentes serviços de informação de acesso aberto; 7) busca ativa por portais brasileiros de revistas científicas; 8) limpeza e correção dos dados coletados; 8) upload em lote dos registros das revistas científicas e dos portais de revistas coletados nas etapas anteriores; 9) migração do Diretório de ambiente de teste para ambiente de produção (Amaro, Campos e Vilas Boas, 2022, p. 4).

A metodologia de desenvolvimento apresentada pelos autores supracitados é, definitivamente, exemplificativa, isso porque dois fatores precisam ser considerados ao falar sobre a metodologia de desenvolvimento do Miguilim: a) o universo de dados e informações relevantes sobre as revistas científicas; e b) a quantidade de revistas científicas existentes no Brasil. Esses dois fatores, atrelados a questões de cunho tecnológico, tornaram necessário um longo período para desenvolvimento e o trabalho cooperativo de profissionais de diferentes áreas do conhecimento para tornar factível o estabelecimento do Diretório.

Para efeitos exemplificativos, cita-se as fases 5, 6 e 7 de desenvolvimento, responsáveis pela implementação do padrão de metadados para ambas as coleções e coleta de dados sobre as revistas em outros sistemas de informação. Somente essas três fases demandaram boa parte do período total de desenvolvimento, tendo em vista a complexidade das tarefas nelas executadas. Especificamente sobre a fase 6 (coleta de informações de revistas científicas eletrônicas brasileiras de diferentes serviços de informação de acesso aberto), há de se

destacar que foram feitas coletas de dados em diversos outros serviços, como Diadorim, Latindex, DOAJ, Portal ISSN, *Wikidata*, *Google Scholar Metrics*, etc. Tais coletas retornaram mais de 5.000 registros, que posteriormente precisaram ser deduplicados para que na Fase 8 pudessem ser feitas limpezas e correções nos dados.

Superadas as fases de desenvolvimento, o Miguilim foi lançado em fase de testes do mês de janeiro de 2022, momento em que os editores começaram a ser contatados para que eles próprios pudessem atualizar os dados de suas revistas e registrar revistas que eventualmente ainda não estivessem registradas. Desse período em diante foi iniciada uma fase de interlocução com a comunidade científica e editorial, que não se furtou a fazer considerações de possíveis melhorias no Diretório, as quais foram e continuam sendo implementadas pela equipe gestora.

Oficialmente, o Miguilim foi lançado em novembro de 2023 durante o XXII Seminário Nacional de Bibliotecas Universitárias (SNBU), especificamente durante o evento complementar intitulado “Ciência Aberta Brasil: ações, produtos e serviços do Ibict para o desenvolvimento da ciência nacional”. Durante o evento, a equipe técnica e gestora teve a oportunidade de apresentar o Diretório como um novo serviço do Ibict para as revistas científicas brasileiras, bem como demonstrar suas potencialidades para a realização de pesquisas que as tenham por base.

## CONSIDERAÇÕES FINAIS

A relevância das revistas no cenário da comunicação científica mundial é de fundamental importância, razão pela qual diversos são realizados diversos estudos e implementação de políticas públicas que consideram sua influência neste contexto. Tal relevância fez com que o Ibict notasse a necessidade de prover a comunidade científica e editorial brasileiro de um sistema capaz de, em alguma medida, traçar um retrato das revistas científicas brasileiras, ainda que esse objetivo seja um tanto quanto desafiador.

Neste sentido, e ancorado em iniciativas já nacionalmente reconhecidas e consolidadas, a partir de 2017 o Instituto inicia o processo de construção do Miguilim, um Diretório nacional que sistematiza e difunde dados sobre mais de 5.000 revistas científicas brasileiras e quase 300 portais de revistas. Contando com diferentes funcionalidades, almeja-se que o Diretório possa ser usado para diversas finalidades no contexto da pesquisa científica nacional, inclusive na formação de políticas públicas pelo Estado brasileiro.

Sabe-se que um serviço dessa proporção precisa passar por constantes correções e atualizações, seja pela ocorrência de erros de programação, seja pela necessidade de aperfeiçoamentos de maneira geral. Ciente disso, a equipe técnica e gestora do Miguilim se compromete a estar sempre atenta às mais diversas transformações no cenário mundial que possam afetar a comunicação científica e, especificamente, as revistas científicas, de modo a dotar o Miguilim das funcionalidades necessárias para que ele seja

uma fonte de informação confiável sobre as revistas científicas brasileiras.

## REFERÊNCIAS

AMARO, Bianca; CAMPOS, Phillip de Freitas; BOAS, Raphael Faria Vilas. Manuelzão e Miguilim: iniciativas do Ibict para os editores e revistas científicas brasileiras. In: ABEC MEETING, 2022, São Paulo. **Anais** [...]. São Paulo: Abec, 2022. DOI: 10.21452/abecmeeting2022.148.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Documento de apoio do Miguilim**. Brasília: Ibict. Disponível em:  
<https://miguilim.ibict.br/static/pages/Documento-de-apoio.pdf>. Acesso em: 30 abr. 2024.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Conheça os padrões de metadados do Miguilim**. Brasília: Ibict. Disponível em:  
<https://miguilim.ibict.br/static/pages/padroes-de-metadados.jsp>. Acesso em: 11 ago. 2025.

Como citar este capítulo

ANDRADE, Denise Aparecida Freitas de; CAMPOS, Phillip de Freitas; AMARO, Bianca. Diretório das revistas científicas eletrônicas brasileiras (Miguilim). In: AMARO, Bianca; CAMPOS, Phillip de Freitas; BARCELOS, Janinne. (org.). **Infraestruturas de Ciência e de Acesso Aberto no Brasil**: iniciativas do Instituto Brasileiro de Informação em Ciência e Tecnologia. Brasília, DF: Editora Ibict, 2025. Cap. 10, p. 163-172. DOI: 10.22477/9788570132543.cap10

# **REDE MOARA: PLATAFORMA PARA COMPARTILHAMENTO DE CÓDIGOS-FONTE PARA PESQUISA**

Milton Shintaku

Bernardo Dionízio Vechi

Rebeca dos Santos de Moura

Lucas Rodrigues Costa

Diego José Macedo



## INTRODUÇÃO

A criação da Web nos anos 1990 revolucionou a democratização da informação, comparável à invenção da prensa de tipos móveis de Gutenberg. Segundo Berners-Lee, Cailliau e Groff (1992), a Web permite acessar um universo de informações online em qualquer lugar. Inicialmente, popularizou o uso da internet antes restrita a redes empresariais e institucionais. Com a criação da Web, surgiram inúmeras iniciativas em todas as áreas. Nas ciências, iniciaram-se com páginas estáticas das universidades, evoluindo para blogs de pesquisadores e outras formas de disseminação da informação. Um marco dessa época foi o lançamento do arXiv em 1991, voltado à publicação de pré-prints.

Van de Sompel e Lagoze (2000) vislumbraram a possibilidade de os autores publicarem livremente seus trabalhos de todas as áreas, como a literatura cinzenta. Assim, em 1999, foi criado o movimento dos arquivos abertos (*Open Archives*) pela Convenção de Santa Fé, Novo México. Triska e Café (2001) destacaram os principais pontos do movimento: definição de um conjunto mínimo de metadados, uso do formato XML, definição de um protocolo comum e políticas de gestão dos documentos. No Brasil, o maior marco desse movimento é a Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), que fomentou a disponibilização da literatura cinzenta das universidades.

Pouco tempo após o surgimento dos arquivos abertos, surgiu o movimento de acesso aberto, em grande parte devido à crise dos periódicos. Harnad (2005) descreve dois caminhos para o acesso aberto à informação

científica: as revistas de acesso aberto (Via Dourada) e os repositórios que mantêm artigos publicados nas revistas (Via Verde).

Em paralelo, na ciência da computação, o movimento dos softwares livres ganhava força. Inicialmente, os softwares eram desenvolvidos internamente nas organizações e aos poucos viraram itens de comercialização com licenças de uso, os chamados softwares proprietários. Como reação, surgiram os softwares livres, que permitem o uso sem cobrar licença e a distribuição dos seus códigos-fonte.

Costa (2006) advoga pela junção dessas iniciativas em uma filosofia aberta: software aberto, arquivos abertos e acesso aberto. Esse ponto foi retomado no movimento de Ciência Aberta, que fomenta a transparência nas atividades científicas com reuso de dados, colaboração entre pesquisadores e participação social (Silva; Silveira, 2019).

Silva e Silveira (2019) relatam que a Ciência Aberta é um movimento que fomenta a transparência nas atividades científicas, com reuso de dados, colaboração entre pesquisadores, participação social e outros. Já Albagli (2015), a Ciência Aberta ainda está em construção e não há uma delimitação de sua abrangência, na medida em que possibilita uma variedade de práticas e abordagens, possibilitando diferentes interpretações. Assim, demonstra que a Ciência Aberta é a mais abrangente dos movimentos, pois não se restringe aos resultados, mas sim em todo o processo científico, incluindo todos os seus elementos.

Nesse contexto, Silveira *et al.* (2023), ao revisitar a taxonomia da Ciência Aberta inicialmente proposta por Pontika *et al.* (2015), propõe, entre outros pontos, as

infraestruturas e ferramentas científicas abertas, a qual engloba várias iniciativas voltadas à abertura das tecnologias que apoiam as atividades científicas. Entre as iniciativas deste tópico, tem-se os repositórios de códigos abertos, já famosos na ciência da computação por hospedar softwares livres de códigos abertos.

Assim, desde o início da abertura das ciências, o Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) tem aderido a todos os movimentos, com várias iniciativas, incluindo as questões relacionadas aos softwares livres. Entre tantas iniciativas relacionadas à Ciência Aberta, neste capítulo, destaca-se os repositórios de código aberto, que no Ibict se materializa na Rede Moara.

## **REDE MOARA: HISTÓRICO E APRESENTAÇÃO**

A Rede Moara (Ibict, 2025) é uma iniciativa do Ibict alinhada aos princípios da Ciência Aberta, projetada para promover um ecossistema colaborativo no desenvolvimento e compartilhamento de software livre para pesquisa no Brasil. O nome "Moara" origina-se do Tupi e significa "aquele que ajuda no nascimento", simbolizando a missão da Rede de facilitar a criação e disseminação de novas tecnologias livres. De certa forma, a Rede Moara é aquela que ajuda o software livre, seu código, a ser disseminado para a comunidade de usuários e colaboradores, ou seja, nascer para o público.

Ao apresentar uma revisão da taxonomia da Ciência Aberta, Silveira *et al.* (2023) apresenta o tema da Infraestrutura aberta. Assim, a Rede Moara se posiciona no contexto da Ciência Aberta, buscando atender à

necessidade de plataformas abertas para o compartilhamento de código-fonte, alinhados aos preceitos apresentados na missão do Ibict. Assim como a Biblioteca Digital Brasileira de Teses e Dissertações (BDTD) promoveu o acesso aberto à literatura cinzenta no Brasil, a Rede Moara facilita o acesso e a reutilização de códigos-fonte desenvolvidos por pesquisadores brasileiros, promovendo a inovação e a eficiência na pesquisa científica.

Assim, a Rede Moara se alinha a outras iniciativas do Instituto na abertura das ciências, iniciadas no movimento dos arquivos abertos. Para tanto, iniciou seus estudos enfocando o problema da não existência de um repositório aberto de códigos-fonte governamental. Com os resultados dos primeiros estudos, verificou-se a necessidade de criação de um ecossistema de informação, que atendesse as necessidades da comunidade de desenvolvedores de códigos-fonte voltados para a pesquisa.

Como relatam de Moura *et al.* (2024), na criação da Rede Moara, os estudos voltaram-se para formulação das suas políticas, dentre as quais, estavam as relacionadas aos direitos autorais dos códigos, na medida em que essa produção não é coberta por legislação específica. Na maioria dos casos, códigos-fonte estão baseados em licenças de uso e distribuição, mas não levam em conta a questão autoral. Assim, a Rede atende a questões básicas de direitos autorais a uma produção intelectual que nem sempre se preocupa com a autoria, mas sim com a distribuição e uso.

Nesse quesito, a Rede Moara é um sistema de informação, composto por métodos, padrões, processos e

outros voltados a gerir e dar visibilidade a softwares livres de pesquisas, por meio dos seus códigos-fonte. Um sistema de informação que está implementado em um ecossistema de informação, com várias ferramentas, em que cada um atua de forma a complementar o objetivo geral do sistema.

## HISTÓRICO

O Ibict tem atuado com apoio ao uso de tecnologias no decorrer de sua história, tendo sido o distribuidor de software de gestão de bibliotecas, além de desenvolver algumas ferramentas. Nesse caminho, tem apoiado ferramentas que se tornaram quase uma hegemonia em algumas áreas, como no caso do DSpace para repositórios e bibliotecas digitais e o *Open Journal Systems* (OJS) para revistas científicas de acesso aberto.

Assim, em agosto de 2021, a Coordenação de Tecnologias para Informação (Cotec) iniciou um projeto de pesquisa, a fim de estudar um repositório de códigos fontes, embrião da Rede Moara. Esse primeiro projeto de pesquisa, tinha por objetivo “Organizar e fornecer infraestrutura para organização de códigos livres (*Open Source*) de sistemas voltados para a pesquisa em CT&I”, com duração de um ano. Com isso, deu-se início ao projeto que culminaria com a Rede Moara.

Os primeiros estudos iniciados em 2021 resultaram na determinação de todas as tecnologias a serem utilizadas na rede, sendo a infraestrutura básica, assim como a criação inicial das políticas a serem adotadas pelo sistema. Possivelmente o maior dos resultados se deu em relação

aos direitos autorais dos softwares livres, no qual gerou o “Guia Prático de Proteção para Softwares Livres” (Sousa, 2022).

De forma a dar continuidade aos estudos, em outubro de 2023, um novo projeto de pesquisa iniciou-se com o intuito de disseminar a rede, dando o viés da ciência aberta e ao compromisso do governo aberto. Assim, a partir de 2024, passou-se a disseminar a rede, sendo apresentado, inclusive em conferência livre da Semana Nacional de Ciência e Tecnologia (SNCT).

Nesse caminho, apesar do pouco tempo de existência, na medida em que a Rede Moara ter sido disponibilizada ao final de 2022, apresenta potencial por estar alinhado a vários pontos, sendo a ciência aberta e a questão de soberania tecnológica as que mais se destacam. Assim, a Rede Moara torna-se mais um sistema de informação do Ibict, que contribui na promoção da infraestrutura informacional para a democratização da informação.

## **ESTRUTURA E FUNCIONAMENTO**

A estrutura da Rede Moara foi desenvolvida para promover a transparência e a colaboração entre desenvolvedores e pesquisadores. Para tanto, conta com uma estrutura diversificada, composta por vários softwares livres, que atendem a uma determinada função no ecossistema. O *front-end* é formado por um portal criado com o *Content Management System* (CMS) *WordPress*, sendo um agregador para as outras ferramentas. Para cada código-fonte gerenciado na Rede Moara, cria-se uma

página no *WordPress*, que atua como um repositório centralizado para o cadastro de todos os códigos-fonte depositados. Esses códigos-fonte são organizados conforme a área de atuação, seguindo a taxonomia estabelecida pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). A estrutura da Rede Moara é composta pelos seguintes componentes:

- GitLab: *locus* principal para o depósito dos códigos-fonte, possibilitando a colaboração eficiente entre desenvolvedores e pesquisadores. A plataforma facilita a colaboração, acessibilidade e usabilidade da ciência e das ferramentas desenvolvidas no contexto científico brasileiro. O *GitLab* centraliza o gerenciamento de códigos-fonte, oferecendo ferramentas robustas para controle de versão, documentação e colaboração, promovendo um ambiente integrado e colaborativo.
- Fórum de Discussão: espaço dedicado à troca de ideias e à discussão sobre utilização e desenvolvimento dos códigos e softwares. O objetivo do fórum é ser utilizado pela comunidade para interação entre os pares, proporcionando um ambiente de suporte mútuo, fundamental para resolução colaborativa de problemas associados aos produtos tecnológicos desenvolvidos pelas pesquisas brasileiras.
- Wiki do Ibict: serve como um recurso para documentação online dos softwares apoiados

pelo Instituto. Devido à facilidade de produção colaborativa de conteúdos, a Wiki permite que desenvolvedores e pesquisadores acessem e contribuam com informações atualizadas e detalhadas sobre os sistemas e tecnologias utilizadas, o que corrobora para a promoção da transparência e eficiência na disseminação de conhecimento técnico.

- Portal de Documentação Técnica: reúne publicações técnicas e científicas úteis tanto para desenvolvedores quanto para pesquisadores. Este portal oferece acesso a vários documentos técnicos, metodologias e estudos científicos dos projetos desenvolvidos pela Coordenação de Tecnologias para Informação (Cotec), fortalecendo a base de conhecimento disponível para a comunidade científica e tecnológica.

Utilizando o *GitLab* como plataforma principal, a Rede Moara centraliza o gerenciamento de códigos-fonte, oferecendo ferramentas para controle de versão, documentação e colaboração. Além do *GitLab*, a Rede dispõe de um fórum de discussão, uma wiki para esclarecimento de dúvidas e um portal de documentação técnica, criando um ambiente integrado e colaborativo.

Essa infraestrutura não apenas simplifica o compartilhamento de códigos-fonte entre pesquisadores, mas também promove a visibilidade da produção tecnológica nacional, contribuindo para o avanço da Ciência Aberta no Brasil. A Rede Moara reforça o compromisso com a transparência, colaboração e

compartilhamento de conhecimento, princípios fundamentais da Ciência Aberta. Como relatam, Moura *et al.* (2024), A Rede Moara foi desenvolvida com o uso de softwares livres para compartilhamento de códigos fontes de softwares livres de pesquisa.

## BENEFÍCIOS E IMPACTOS

Toda iniciativa que atua em uma lacuna impacta o cenário em que está podendo trazer benefícios às pessoas, instituições e outros. No caso da Rede Moara, tem-se uma lacuna na oferta de uma iniciativa governamental aberta para depósito de códigos fontes. Dessa forma, os benefícios da Rede Moara são diversos e alinhados aos princípios discutidos anteriormente e são:

- Fomento à Ciência Aberta: assim como a BD TD promoveu o acesso aberto à literatura cinzenta, a Rede Moara facilita o acesso e a reutilização de códigos desenvolvidos para pesquisa, promovendo a transparência e a colaboração, por meio da reutilização.
- Colaboração e Inovação: a plataforma estimula o desenvolvimento colaborativo e a inovação tecnológica, refletindo os princípios da Ciência Aberta e do movimento de software livre.
- Democratização do Acesso ao Conhecimento: torna as tecnologias desenvolvidas acessíveis a um público mais amplo, alinhando-se ao objetivo da Ciência Aberta de democratizar o acesso à informação.

- Interoperabilidade e Reutilização de Códigos: facilita a integração de diferentes sistemas e a reutilização de soluções existentes, promovendo a eficiência e a economia de recursos.
- Suporte e Desenvolvimento Comunitário: cria uma comunidade de suporte mútuo entre desenvolvedores e pesquisadores, fomentando a colaboração interdisciplinar.
- Interação entre Pesquisadores, Desenvolvedores e o PÚblico em Geral: promove a troca de conhecimentos e experiências, enriquecendo o processo de desenvolvimento científico.

A Rede Moara representa um avanço significativo para a Ciência Aberta no Brasil, promovendo a transparéncia, a colaboração e a inovação tecnológica. Ao facilitar o compartilhamento de códigos-fonte, a Rede contribui para a construção de um ecossistema mais aberto e inclusivo. A plataforma simplifica o processo de desenvolvimento colaborativo e fortalece a comunidade científica, criando um ambiente de suporte mútuo e intercâmbio de conhecimentos.

Como modelo para outras iniciativas, a Rede Moara destaca a importância de uma infraestrutura aberta e colaborativa para o avanço científico e tecnológico. Ela inspira uma mudança cultural na forma como a ciência é feita, promovendo a inclusão, o acesso aberto e a colaboração global. O compromisso com a Ciência Aberta e a inovação contínua garantirá que a Rede Moara seja um

recurso essencial para pesquisadores e desenvolvedores, contribuindo significativamente para o progresso científico no Brasil.

## **CONSIDERAÇÕES FINAIS**

Desde o movimento dos arquivos abertos até a abertura ampla das ciências, o Ibict tem atuado nesses movimentos, alinhado à sua missão. Historicamente, o Instituto tem estado na vanguarda dos estudos envolvendo a informação científica e tecnológica, criado por meio de uma provocação da Organização das Nações Unidas para a Educação, a Ciência e a Cultura (Unesco), liderada por Lydia Sambaquy.

Nesse sentido, muitas outras iniciativas do Ibict têm promovido ações relacionadas à Ciência Aberta. Ao longo das atividades, o Ibict tem inúmeros exemplos de esforços para tornar acessível de forma ampla e sem restrições, adaptando-se à tecnologia da época e aos preceitos teóricos. Inicialmente, o Ibict atuava com os resultados de pesquisa publicados em canais tradicionais, mas depois expandiu para novas formas de disseminação da informação.

Com a Rede Moara, o Instituto passa a atuar na infraestrutura de pesquisa, abrindo os resultados de desenvolvimento de softwares de projetos de pesquisa. Dessa forma, oferece à comunidade científica e tecnológica um sistema de informação para hospedagem de softwares livre, em iniciativa governamental, mantida por instituição tradicional na disseminação da informação. Embora faltem dados gerais, é provável que muitos

códigos-fonte criados em projetos de pesquisa não sejam compartilhados, resultando em desconhecimento e re-desenvolvimento desnecessários, onerando as pesquisas. Além disso, o uso de softwares estrangeiros mantém a dependência do norte global, afetando a soberania tecnológica.

A Rede Moara visa dar maior visibilidade à produção nacional de softwares livres de pesquisa, fomentando a comunidade brasileira na evolução de códigos promissores. A comunidade de desenvolvedores brasileiros contribui com projetos estrangeiros, mas há pouca interação interna. Com a Rede, há uma oportunidade de colaboração interna, mostrando a importância da Rede Moara no cenário tecnológico nacional. Ela fomenta a colaboração entre desenvolvedores e requer apoio das agências de fomento para orientação de depósito dos produtos tecnológicos na Rede, a posicionando como uma estratégia nacional para soberania tecnológica no desenvolvimento de softwares livres para pesquisa.

## REFERÊNCIAS

ALBAGLI, Sarita. Ciência aberta em questão. In: ALBAGLI, Sarita; MACIEL, Maria Lucia; ABDO, Alexandre Hannud (org.). **Ciência aberta, questões abertas**. Brasília: Ibitc; Rio de Janeiro: Unirio, 2015. cap. 1, p. 9-26.

BERNERS-LEE, Tim; CAILLIAU, Robert; GROFF, Jean-François. The world-wide web. **Computer Networks and ISDN Systems**, North-Holland, v. 25, p. 454-459, 1992.

COSTA, Sely Maria de Souza. Filosofia aberta, modelos de negócios e agências de fomento: elementos essenciais a uma

discussão sobre o acesso aberto à informação científica.

**Ciência da Informação**, Brasília, v. 35, v. 2, 2006. DOI:  
10.18225/ci.inf.v35i2.1139.

HARNAD, Stevan. The green and gold roads to maximizing journal article access, usage and impact. **Haworth Press**, Binghamton, 2005.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Rede Moara**. Brasília: Ibitc. Disponível em: <https://redemoara.ibict.br>. Acesso em: 11 ago. 2025.

MOURA, Rebeca dos Santos de et al. Rede Moara: sistema de disponibilização de códigos-fonte abertos. In: WORKSHOP DE INFORMAÇÃO, DADOS E TECNOLOGIA, 7., 2024. **Anais** [...]. Porto Velho: Ibitc, 2024. DOI: 10.22477/vii.widat.163.

PONTIKA, Nancy et al. Fostering open science to research using taxonomy and an elearning portal. In: INTERNATIONAL CONFERENCE ON KNOWLEDGE TECHNOLOGIES AND DATA-DRIVEN BUSINESS, 15., 2015. **Proceedings** [...]. Graz, Áustria: Association for Computing Machinery, 2015. DOI: 10.1145/2809563.2809571.

SILVA, Fabiano Couto Corrêa da; SILVEIRA, Lúcia. O ecossistema da ciência aberta. **Transinformação**, Campinas, v. 31, e190001, 2019. DOI: 10.1590/2318-0889201931e190001.

SILVEIRA, Lúcia da et al. Taxonomia da Ciência Aberta: revisada e ampliada. **Encontros Bibl**: revista eletrônica de biblioteconomia e ciência da informação, Florianópolis, v. 28, p. 1-22, 2023. DOI: 10.5007/1518-2924.2023.e91712.

SOUZA, Rosilene Paiva Marinho de et al. **Guia Prático de Proteção para Software Livre**. Brasília: Ibitc, 2022. DOI: 10.22477/9786589167471.

TRISKA, Ricardo; CAFÉ, Lígia. Arquivos abertos: subprojeto da Biblioteca Digital Brasileira. **Ciência da Informação**, Brasília, v. 30, n. 3, 2001. DOI: 10.18225/ci.inf.v30i3.917.

VAN DE SOMPEL, Herbert. The Santa Fe convention of the open archives initiative. **D-Lib Magazine**, Charlottesville, v. 6, n. 2, fev. 2000. DOI: 10.1045/february2000-vandesompel-oai.

### Como citar este capítulo

SHINTAKU, Milton; VECHI, Bernardo Dionízio; MOURA, Rebeca dos Santos de; COSTA, Lucas Rodrigues; MACEDO, Diego José. Rede Moara: plataforma para compartilhamento de códigos-fonte para pesquisa. In: AMARO, Bianca; CAMPOS, Fhillipe de Freitas; BARCELOS, Janinne. (org.). **Infraestruturas de Ciência e de Acesso Aberto no Brasil**: iniciativas do Instituto Brasileiro de Informação em Ciência e Tecnologia. Brasília, DF: Editora Ibitc, 2025. Cap. 11, p. 173-188. DOI: 10.22477/9788570132543.cap11



# O PGD BR COMO FERRAMENTA PARA A PROMOÇÃO DA CIÊNCIA ABERTA

Fabiano Couto Corrêa da Silva

Sandra de Albuquerque Siebra

Alexandre Faria de Oliveira

Laura Vilela Rodrigues Rezende



## APRESENTAÇÃO

Os PGDs têm emergido como elementos fundamentais na pesquisa científica contemporânea, especialmente devido ao crescente volume e complexidade dos dados gerados. A implementação eficaz de PGDs garante a integridade, transparência e reproduzibilidade da pesquisa, aspectos essenciais para o avanço do conhecimento científico. Além disso, agências de fomento como a Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp), o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), a National Science Foundation (NSF) nos Estados Unidos, a European Research Council (ERC) na Europa e o Natural Sciences and Engineering Research Council (NSERC) no Canadá têm exigido a inclusão de PGDs nas propostas de financiamento, reforçando a necessidade de acompanhamento do processo de gestão de dados estruturada e responsável por parte dos pesquisadores. Adicionalmente, PGDs têm sido paulatinamente recomendados em universidades e institutos de pesquisa, como parte de orientações e políticas relacionadas à gestão dos dados das pesquisas realizadas dentro destas instituições (Monteiro *et al.*, 2023).

Um PGD é um documento que descreve como os dados serão gerenciados ao longo do ciclo de vida de um projeto de pesquisa. Isso inclui a coleta, organização, armazenamento, compartilhamento e preservação dos dados. Os PGDs visam assegurar que os dados sejam manejados de forma ética e eficiente, promovendo sua reutilização e interoperabilidade, facilitando a colaboração

entre pesquisadores e instituições. Com isso, aumentam a confiança nos resultados da pesquisa e otimizam o impacto científico e social dos estudos realizados (Williams; Bagwell; Zozus, 2017).

O PGD BR é uma iniciativa brasileira desenvolvida pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) com o objetivo de promover, de maneira otimizada, o planejamento da gestão de dados científicos no Brasil. Esta ferramenta de elaboração guiada de PGDs é voltada para pesquisadores e instituições de pesquisa que buscam atender às exigências das agências de fomento e aprimorar a qualidade e a reproduzibilidade de suas pesquisas.

## A IMPORTÂNCIA DOS PLANOS DE GESTÃO DE DADOS NA CIÊNCIA ABERTA

A Ciência Aberta tem se consolidado como um modelo essencial para a pesquisa científica contemporânea, promovendo a transparência, a colaboração e o compartilhamento de conhecimento. Nesse contexto, os PGDs desempenham um papel essencial ao fornecer diretrizes e estratégias para a gestão eficaz dos dados de pesquisa.

De acordo com Balbinotti *et al.* (2022), os PGDs são documentos que descrevem como os dados de pesquisa serão coletados, organizados, armazenados, preservados e compartilhados ao longo do ciclo de vida de um projeto. Essas diretrizes são essenciais no contexto da Ciência Aberta, pois facilitam a transparência, a reproduzibilidade e o compartilhamento de dados científicos. Andrade *et al.*

(2022) destacam a importância da padronização e aprimoramento das bases de dados para ampliar o acesso, a transparência e a integração entre os campos do Direito e da Saúde Coletiva.

Os benefícios dos PGDs se estendem para pesquisadores, instituições e a sociedade em geral. Conforme apontado por Veiga *et al.* (2019), os PGDs representam uma ferramenta para organizar e documentar dados, facilitando a colaboração e a reutilização de informações. Para as instituições, a implementação de PGDs promove a integridade e a transparência, fortalecendo a reputação e contribuindo para avanços científicos significativos, como discutido por Rocha (2020).

Os PGDs descrevem as ações que serão realizadas conforme o fluxo: Planejamento > Coleta de Dados > Organização > Armazenamento > Documentação > Compartilhamento > Preservação. Esse fluxo simplificado demonstra como os PGDs estruturam o gerenciamento de dados em todas as etapas do ciclo de vida da pesquisa, garantindo que os dados sejam bem geridos desde o planejamento e preparação da pesquisa, passando por sua coleta até a sua preservação a longo prazo. Um PGD possui uma estrutura textual composta por três elementos principais: dados, gestão e plano. Esses elementos são interdependentes e juntos garantem que os dados sejam bem geridos e utilizados de forma eficiente.

Desse modo, os PGDs desempenham um papel essencial na interseção entre a gestão de dados e os princípios da Ciência Aberta. De acordo com Silva e Santos (2021) e Vitale e Sandy (2019), a importância dos PGDs na Ciência Aberta reside na formalização das intenções dos

pesquisadores em relação à gestão de dados, garantindo transparência e reutilização dos mesmos. Esses planos são essenciais para que os pesquisadores possam definir claramente como os dados serão coletados, armazenados, compartilhados e preservados, promovendo uma prática mais ética e eficiente na gestão de dados científicos.

Veiga *et al.* (2025) e Miksa, Oblasser e Rauber (2021) destacam os desafios e oportunidades na implementação de PGDs em projetos de pesquisa, ressaltando a necessidade de uma abordagem estruturada para assegurar a eficácia na gestão dos dados. Os desafios incluem a necessidade de treinamento adequado para pesquisadores, a criação de infraestrutura tecnológica robusta e a superação de barreiras culturais dentro das instituições de pesquisa. Contudo, existem também oportunidades significativas proporcionadas pelos PGDs, como a melhoria na qualidade dos dados e o aumento da colaboração científica. Um exemplo de oportunidade é a potencial facilidade de auditoria e validação dos dados, que pode fortalecer a credibilidade e a confiança nos resultados de pesquisa.

A evolução dos PGDs e sua relação com a Ciência Aberta foram discutidas por Ibict (2024) e Buys e Shaw (2015), evidenciando a importância crescente desses documentos na promoção da transparência e reutilização dos dados científicos. Ao longo dos anos, os PGDs passaram de simples ferramentas de planejamento para elementos centrais na estratégia de gestão de dados de muitas instituições. Além disso, Stueber *et al.* (2023), Costa e Sousa (2020) e Holler e Kedron (2022) ressaltam a importância das boas práticas em PGDs para impulsionar a

Ciência Aberta, destacando a necessidade de tornar os dados acessíveis e reutilizáveis. A integração dos princípios FAIR (*Findable, Accessible, Interoperable, Reusable*) nos PGDs é fundamental para garantir a qualidade e usabilidade dos dados, conforme observado por Carvalho e Lima (2022). Estes princípios fornecem um quadro claro e prático para a gestão de dados, permitindo que os dados sejam planejados para serem facilmente localizados, acessados e reutilizados por outros pesquisadores, promovendo assim uma cultura de transparência e colaboração.

A Ciência Aberta, ao garantir a disponibilidade e usabilidade sem barreiras de acesso às publicações acadêmicas e aos dados resultantes de pesquisas, promove a transparência e o acesso aberto a dados e publicações científicas, como enfatizado por Silva e Silveira (2019). A transparência e a disponibilidade dos dados científicos são essenciais para a confiança nos resultados de pesquisa e para a ampliação da visibilidade das investigações, conforme apontado por Ribeiro (2022) e Ribeiro *et al.* (2022). Quando os dados estão acessíveis e reutilizáveis, eles podem ser verificados e replicados por outros pesquisadores, o que é vital para a robustez e a credibilidade da pesquisa científica, como também para o progresso da ciência.

Neste cenário, o PGD BR, desenvolvido pelo Iibict, exemplifica a aplicação prática dos princípios da Ciência Aberta no Brasil. Ele facilita a criação de PGDs alinhados aos princípios FAIR e promove boas práticas de gestão de dados entre os pesquisadores brasileiros. Com um suporte robusto para metadados, garante que os dados científicos sejam bem documentados, facilitando que estejam

acessíveis e reutilizáveis, amplificando os benefícios da Ciência Aberta. Desta forma, o PGD BR se posiciona como uma ferramenta potencialmente relevante na promoção da transparência, reproduzibilidade e colaboração na pesquisa científica. Ao estabelecer diretrizes claras e padronizadas, contribui para a integridade e a confiabilidade da pesquisa, beneficiando pesquisadores, instituições e a sociedade como um todo.

## HISTÓRICO DE CRIAÇÃO E DESENVOLVIMENTO

A concepção do PGD BR se iniciou com um processo de modelagem conceitual voltado para mapear e detalhar os principais atores e ações interligadas em um sistema de elaboração de PGD, considerando a realidade brasileira. A partir daí, após um estudo realizado de análise detalhada de diversas ferramentas internacionais de gestão de dados, optou-se por customizar um software livre já existente, ao invés de se desenvolver desde o início. O software escolhido para servir como base para o PGD BR foi o DMPTool, tendo sido considerados fatores como flexibilidade, capacidade de integração e suporte aos Princípios FAIR (*Findable, Accessible, Interoperable, and Reusable*), que o destacaram como a melhor opção, mostrando-se como uma solução robusta e escalável para a gestão de dados científicos (Stueber *et al.*, 2023; Márdero Arellano *et al.*, 2024).

O DMPTool, desenvolvido pela *University of California Curation Center (UC3)* do *California Digital Library*, é um software amplamente utilizado por instituições de pesquisa nos Estados Unidos para facilitar a

criação e a gestão de PGDs. Por exemplo, a Universidade de Stanford adota o DMPTool para ajudar seus pesquisadores a desenvolver PGDs que estejam em conformidade com as exigências de agências de fomento, como a *National Institutes of Health (NIH)* e a *National Science Foundation (NSF)* (*Stanford University Libraries*, 2023). A Universidade de Michigan também integra o DMPTool em seu processo de gestão de dados, proporcionando aos seus pesquisadores modelos personalizados e recursos de apoio para a criação de PGDs (*University of Michigan Library*, 2023). Além disso, o *Massachusetts Institute of Technology (MIT)* utiliza o DMPTool para promover práticas de gestão de dados entre seus pesquisadores, assegurando que os dados de pesquisa sejam bem organizados, acessíveis e reutilizáveis (*MIT Libraries*, 2023). Essas implementações exemplificam como o DMPTool tem sido fundamental para melhorar a gestão de dados científicos em diversas instituições de renome, promovendo a transparência, a reprodutibilidade e a conformidade com as diretrizes de financiamento e endossando a decisão do Ibict em utilizá-lo como base para o PGD BR.

O modelo conceitual desenvolvido pela equipe de pesquisadores à frente do projeto de implementação do PGD BR posiciona a ferramenta brasileira de elaboração de PGDs como um centralizador de informações referentes às pesquisas. Na parte superior do fluxo, estão as informações sobre os pesquisadores e custos envolvidos no projeto, incluindo o acesso à plataforma PGD com credenciais definidas. A modelagem conceitual trouxe agentes com informações provenientes de várias fontes. A Comunidade Acadêmica Federada (Cafe), a primeira federação

acadêmica do país, sendo uma das cinco maiores do mundo em número de clientes. O ORCID (*Open Researcher and Contributor Identifier*), que fornece um identificador único ao pesquisador e vincula-o com suas informações profissionais, como afiliações, bolsas, publicações e revisões por pares. O Currículo Lattes, a principal plataforma de gestão de currículos brasileira, e o GovBr, serviço governamental de identificação em meios digitais, que permite identificação segura ao acessar serviços digitais governamentais.

As agências de fomento são fornecedoras de informações sobre chamadas de financiamento, projetos financiados e seus respectivos vínculos com a equipe de pesquisadores. Existe a possibilidade de registro de um identificador único das agências de fomento via recurso *Crossref Founder Registry*, favorecendo a parametrização dessas informações. A relação entre o PGD e os sistemas das agências de fomento é bidirecional, pois além de fornecerem informações, essas agências também podem utilizar as informações provenientes dos PGDs para gerar indicadores e estatísticas de financiamento.

O PGD BR passou por um processo evolutivo significativo desde a sua concepção inicial até a implementação atual. Concebido pelo Grupo de Pesquisa DataLab, seu desenvolvimento foi posteriormente assumido pelo Ibiict, que ampliou e aperfeiçoou a plataforma.

A versão inicial do PGD BR foi projetada para fornecer uma solução básica para a gestão de PGDs. A plataforma permite que os pesquisadores registrem e gerenciem seus planos de maneira centralizada. Esta

abordagem inicial ofereceu uma estrutura fundamental para a organização dos dados de pesquisa, promovendo uma maior transparência e reproduzibilidade científica. No entanto, à medida que a demanda por Ciência Aberta crescia, tornou-se evidente que funcionalidades mais avançadas eram necessárias para atender às complexidades da gestão de dados em diferentes disciplinas científicas no contexto brasileiro.

Reconhecendo a necessidade de uma infraestrutura mais robusta, o IbiCT assumiu o desenvolvimento do PGD BR, introduzindo uma série de melhorias e novas funcionalidades. Esta transição foi motivada por vários fatores, incluindo:

- Gestão de Metadados: A plataforma passou a permitir a inclusão e a gestão detalhada de metadados, que são informações descritivas essenciais sobre os dados de pesquisa. Isso inclui detalhes sobre como os dados foram coletados, as metodologias utilizadas e os contextos de uso, facilitando a reproduzibilidade e a interpretação dos dados. Por exemplo, em um estudo epidemiológico, os metadados podem incluir informações sobre a população estudada, os instrumentos de coleta de dados e os métodos de análise estatística. Isso permite que outros pesquisadores possam entender e replicar o estudo com precisão.
- Integração com Plataformas e Repositórios: A nova versão do PGD BR foi projetada para se integrar com outras plataformas de dados e repositórios, tanto nacionais quanto internacionais. Isso facilita o

compartilhamento e a disseminação de dados, promovendo uma maior colaboração entre pesquisadores e instituições. Por exemplo, a integração com o repositório de dados da Open Science Framework (OSF) permite que os dados de pesquisa sejam facilmente compartilhados com uma comunidade global de pesquisadores, aumentando a visibilidade e o impacto dos projetos de pesquisa brasileiros.

Ressalta-se que além das funcionalidades, também buscou-se aprimorar a usabilidade da plataforma, a fim de contribuir para que pesquisadores de diferentes áreas do conhecimento, independentemente de sua familiaridade com tecnologias digitais, possam utilizar o PGD BR com facilidade. As melhorias incluíram:

- Design Intuitivo: A interface foi redesenhada com um foco em simplicidade e clareza, utilizando princípios de design centrado no usuário para facilitar a navegação e o uso da plataforma.
- Tutoriais e Guias: Estão sendo desenvolvidos tutoriais interativos e guias detalhados que acompanham os usuários em cada etapa do processo de gestão de dados, desde o registro até o compartilhamento.
- Feedback do Usuário: A implementação de mecanismos de feedback permite que os usuários contribuam com sugestões e relatem problemas, garantindo que a plataforma possa ser

continuamente aprimorada com base nas necessidades reais dos pesquisadores.

Para garantir que os usuários pudessem maximizar o uso do PGD BR, o Ibict investiu em uma série de programas de treinamento e suporte técnico. Estes esforços incluíram:

- Workshops Presenciais e Online: Frequentemente são organizados workshops, tanto presenciais, quanto online, para capacitar os pesquisadores no uso da plataforma e em melhores práticas de gestão de dados. Os workshops abrangeram diferentes públicos e áreas de pesquisa, como ciências da saúde ou ciências sociais, de forma a abordar as necessidades únicas de cada disciplina.
- Materiais Didáticos: Foram produzidos materiais didáticos, como manuais, vídeos tutoriais e FAQs, que estão disponíveis na plataforma e podem ser acessados a qualquer momento pelos usuários. Esses recursos educacionais fornecem instruções passo a passo e dicas práticas sobre como utilizar a plataforma de forma eficaz, facilitando o aprendizado e a adoção da ferramenta.
- Equipe de Suporte Técnico: Uma equipe dedicada de suporte técnico foi estabelecida para oferecer assistência personalizada aos usuários; respondendo a dúvidas, enviadas, por exemplo, via email ou formulário; resolvendo problemas técnicos e fornecendo orientações específicas conforme necessário.

## IMPLEMENTAÇÕES FUTURAS DE RECURSOS AVANÇADOS

Como evolução da ferramenta, a equipe do PGD BR planeja implementar uma série de recursos avançados para promover a Ciência Aberta:

- Integração com o diretório Re3data (2025): Facilitando a identificação de repositórios por área de conhecimento, garantindo que os pesquisadores possam localizar e utilizar repositórios adequados para o depósito e recuperação de seus dados.
- Suporte aos Princípios FAIR: Adoção dos Princípios FAIR como diretriz central para a gestão de dados, garantindo que os dados sejam encontráveis, acessíveis, interoperáveis e reutilizáveis, promovendo uma maior transparência e colaboração na pesquisa científica.
- *Upload* do logo da instituição, grupo ou projeto de pesquisa: Permitindo a personalização dos PGDs com a identidade visual das instituições, grupos ou projetos de pesquisa, fortalecendo a marca e a visibilidade das iniciativas de pesquisa envolvidas.
- Acionável por máquina (mais campos de preenchimento automático) lattes, orcid, etc.

A elaboração de um Plano de Gestão de Dados (PGD) no PGD BR é facilitada por meio de *templates* que atendem a diferentes necessidades de projetos de pesquisa. O PGD BR oferece dois principais *templates*: o *Template FAIR* e o *Template Resumido*. O *Template FAIR* é detalhado e visa

garantir que os dados de pesquisa sejam Encontráveis, Acessíveis, Interoperáveis e Reutilizáveis. Este *template* orienta os pesquisadores a descreverem de forma abrangente todas as etapas do ciclo de vida dos dados, desde a coleta até a preservação, promovendo a transparência e a conformidade com os princípios FAIR. É ideal para projetos que requerem um alto nível de detalhamento. Em contraste, o *Template Resumido* é mais conciso e aborda os aspectos essenciais da gestão de dados de maneira simplificada. Ele é apropriado para pesquisadores que necessitam de um PGD mais básico, mas ainda eficaz, permitindo uma gestão eficiente sem sobrecarregar com detalhes excessivos.

No processo de elaboração do PGD, a seção "Detalhes do Projeto" é fundamental, pois fornece uma visão abrangente dos elementos principais do projeto. Os pesquisadores devem incluir informações como o título e o resumo do projeto, área de conhecimento, datas de início e conclusão, e quaisquer questões éticas associadas. Essas informações são cruciais para contextualizar a pesquisa e garantir que todos os aspectos do projeto sejam bem compreendidos e documentados. Por exemplo, o título deve capturar a essência do projeto, enquanto o resumo deve fornecer uma descrição breve dos objetivos, métodos e importância do estudo. A área de conhecimento especifica o campo científico ao qual o projeto pertence, facilitando a categorização e o acesso por outros pesquisadores interessados no mesmo domínio. As datas de início e conclusão delineiam o período de execução do projeto, permitindo um planejamento adequado das atividades. As questões éticas, por sua vez, são essenciais

para assegurar que a pesquisa esteja em conformidade com normas e regulamentos éticos, garantindo a proteção dos participantes e a integridade do estudo.

A seção de coleta de dados descreve como os dados serão gerados ou reutilizados durante a pesquisa, promovendo a transparência e a integridade do processo de coleta de dados. Os pesquisadores podem optar por coletar novos dados, reutilizar dados de coletas anteriores ou utilizar dados de outras fontes externas. É necessário especificar os tipos e formatos de dados que serão utilizados, como dados numéricos, software, visualizações, textos, artefatos ou processos. Cada tipo de dado deve ser claramente descrito para facilitar sua organização e reutilização, assegurando que os dados possam ser facilmente interpretados e reutilizados por outros pesquisadores. A descrição detalhada dos tipos e formatos de dados inclui informações sobre a natureza dos dados, como medidas, resultados de levantamentos, simulações, tabelas, gráficos, entrevistas, espécimes e procedimentos operacionais padronizados. Os formatos dos dados, como CSV, DOC, JSON, JPEG, MP4, PDF, entre outros, também devem ser especificados para garantir a interoperabilidade e a reutilização eficaz dos dados.

A documentação e os metadados são elementos essenciais na gestão de dados, proporcionando informações descritivas que acompanham os dados e facilitam sua compreensão e reutilização. Os pesquisadores devem detalhar como os dados serão documentados e quais padrões de metadados serão utilizados, assegurando que os dados possam ser facilmente interpretados e reutilizados por outros

pesquisadores. Isso inclui a criação de *readme files*, cadernos de laboratório e outros documentos descritivos que forneçam contexto e informações detalhadas sobre os dados. A escolha de padrões de metadados adequados é essencial para garantir a qualidade e a usabilidade dos dados, permitindo que outros pesquisadores entendam e reutilizem os dados de maneira eficaz. Os metadados devem incluir detalhes sobre a coleta, organização, metodologia, contexto de uso e qualquer outra informação relevante que facilite a interpretação e reutilização dos dados.

O armazenamento, *backup* e a preservação dos dados são aspectos críticos abordados no PGD, essenciais para garantir que os dados permaneçam disponíveis e utilizáveis no futuro (Oliveira *et al.*, 2023). É importante descrever onde e como os dados serão armazenados e preservados, incluindo os métodos de *backup* que serão utilizados para garantir a segurança dos dados. O planejamento adequado dessas etapas garante a integridade e a disponibilidade dos dados a longo prazo. Os pesquisadores devem especificar onde os dados serão armazenados durante a pesquisa, como em nuvem, dispositivos externos ou repositórios institucionais, e onde serão armazenados após a conclusão do projeto. O *backup* dos dados deve ser feito de maneira regular e sistemática para evitar perda de dados, utilizando métodos como backup institucional, serviços de terceiros ou processos manuais. A preservação a longo prazo dos dados em repositórios seguros e acessíveis é essencial para garantir que os dados permaneçam disponíveis e utilizáveis no futuro.

O acesso e o compartilhamento de dados são componentes fundamentais do PGD, promovendo a transparência e a colaboração na pesquisa científica. Os pesquisadores devem definir como os dados serão acessados e compartilhados durante e após a pesquisa, utilizando controles de acesso adequados e publicando os dados em repositórios institucionais ou temáticos. Isso inclui a gestão do acesso por meio de *logins* e senhas, garantindo que apenas indivíduos autorizados possam acessar os dados. O compartilhamento dos dados em repositórios institucionais ou temáticos facilita a disseminação e a reutilização dos dados, promovendo a colaboração entre pesquisadores e instituições. É importante que os dados sejam compartilhados de maneira que preservem a confidencialidade e os direitos dos participantes da pesquisa, garantindo que os dados sejam utilizados de maneira ética e responsável.

Por fim, as questões éticas e legais relacionadas aos dados devem ser cuidadosamente consideradas e documentadas. Isso inclui a proteção de direitos autorais, a aplicação de licenças apropriadas para o uso e reutilização dos dados, e medidas para garantir a proteção da identidade e privacidade dos participantes da pesquisa. Os pesquisadores devem garantir que todos os aspectos éticos e legais sejam abordados de maneira adequada, promovendo a integridade e a responsabilidade na gestão dos dados. A consideração das questões éticas envolve a obtenção de consentimento informado dos participantes, a proteção da privacidade e a confidencialidade dos dados pessoais, e a conformidade com regulamentos e diretrizes éticas. As questões legais incluem a propriedade

intelectual dos dados, os direitos de uso e a aplicação de licenças que permitam a reutilização dos dados por outros pesquisadores, garantindo que os dados sejam compartilhados de maneira ética e legal.

Figura 1 - Campos de preenchimento do Plano de Gestão de Dados

The screenshot shows the PGD BR interface with the following details:

- Header:** Meu Painel de Controle, Criar plano, Requisitos da Agência de Fomento, PGDs Públicos, Ajuda, Fabiano Couto corrêa da silva, Idioma.
- Project Title:** Captura do som das baleias jubarte no litoral de Pernambuco
- Project Status:** Detalhes do Projeto, Colaboradores, Escrever PGD, Resultados da pesquisa científica, Finalizar plano, Baixar plano.
- Project Description:** Este plano é baseado no modelo "Plano de Gestão de Dados Ibict" fornecido por Instituto Brasileiro de Informação em Ciência e Tecnologia (ibict.br) - (ver: 25, pub: 2024-01-24). Visualizar todos | Fechar todos.
- Content Structure:**
  - + Iniciar o plano (1 / 1)
  - + Coleta de Dados (5 / 7)
  - + Documentação e Metadados (2 / 2)
  - + Armazenamento, Backup e Preservação (3 / 6)
  - + Requisitos Éticos e Legais (5 / 5)
  - + Compartilhamento de Dados (4 / 4)
  - + Responsabilidades e Recursos (1 / 1)

Fonte: PGD BR (2024).

## CONSIDERAÇÕES FINAIS

O Plano de Gestão de Dados (PGD) BR representa uma ferramenta essencial na promoção da Ciência Aberta no Brasil, contribuindo significativamente para a gestão eficaz e ética dos dados científicos. Desde sua concepção inicial pelo Grupo de Pesquisa DataLab até seu desenvolvimento e aprimoramento pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), o PGD BR tem evoluído para atender às crescentes demandas de transparência, reproduzibilidade e acessibilidade na pesquisa científica.

As funcionalidades robustas do PGD BR, como a gestão detalhada de metadados, o controle de versões e a

integração com plataformas e repositórios nacionais e internacionais, são fundamentais para assegurar que os dados de pesquisa sejam bem organizados, acessíveis e reutilizáveis. Essas características, alinhadas aos princípios FAIR (*Findability, Accessibility, Interoperability, Reusability*), garantem que os dados científicos geridos pelo PGD BR possam ser facilmente encontrados, acessados, integrados e reutilizados por outros pesquisadores, promovendo assim uma maior colaboração e inovação científica.

A escolha do DMPTool como base para o PGD BR foi estratégica, considerando sua flexibilidade, capacidade de integração e suporte aos princípios FAIR. A adoção de uma ferramenta amplamente utilizada por instituições de renome internacional, como Stanford, Michigan e MIT, reforça a robustez e a eficácia do PGD BR na gestão de dados científicos. A implementação bem-sucedida dessas funcionalidades nas versões Simplificada e Extendida do PGD BR oferece uma solução adaptável tanto para pesquisadores que buscam uma abordagem rápida e fácil quanto para aqueles que necessitam de critérios avançados e detalhados na gestão de dados.

Além das funcionalidades técnicas, a modernização da infraestrutura tecnológica, incluindo servidores de alta capacidade, sistemas de backup automatizados e protocolos de segurança aprimorados, assegura a confiabilidade e a segurança dos dados geridos pelo PGD BR. A ênfase na usabilidade, com um design intuitivo, tutoriais e guias detalhados, e uma equipe de suporte técnico dedicada, garante que pesquisadores de diferentes

áreas do conhecimento possam utilizar a plataforma de maneira eficiente e eficaz.

Os recursos avançados, como a integração com o diretório Re3data e as ferramentas de análise de metadados, fortalecem ainda mais a capacidade do PGD BR de promover a Ciência Aberta. Ao facilitar a identificação de repositórios adequados e a criação de metadados ricos, o PGD BR assegura que os dados científicos sejam devidamente documentados e descritos, aumentando sua descoberta e compartilhamento.

Em suma, o PGD BR desempenha um papel fundamental na promoção da Ciência Aberta no Brasil, estabelecendo diretrizes padronizadas para a gestão de dados de pesquisa. Sua evolução contínua e alinhamento com as melhores práticas internacionais posicionam o PGD BR como uma ferramenta indispensável para a comunidade científica brasileira. Ao facilitar a transparência, a reproduzibilidade e o compartilhamento de dados, o PGD BR não apenas atende às exigências das agências de fomento e instituições de pesquisa, mas também contribui para um ecossistema de pesquisa mais aberto, colaborativo e eficiente, beneficiando pesquisadores, instituições e a sociedade em geral.

## REFERÊNCIAS

ANDRADE, Nayla Rochele Nogueira de et al. Limites e possibilidades para desenvolvimento de pesquisas em saúde pública no judiciário. **Revista de Saúde Pública**, São Paulo, v. 56, p. 76, 2022. DOI: 10.11606/s1518-8787.2022056004203.

BALBINOTTI, Stheve et al. Competência em dados: uma necessidade contemporânea para pesquisadores e para a sociedade. **AtoZ**: novas práticas em informação e conhecimento, Curitiba, v. 11, p. 1-11, ago. 2022. DOI: 10.5380/atoz.v11i0.83569.

BUYS, Cunera M.; SHAW, Pamela L. Data Management Practices Across an Institution: Survey and Report. **Journal of Librarianship and Scholarly Communication**, Iowa, v. 3, n. 1, 2015. DOI: 10.7710/2162-3309.1225.

CARVALHO, Luiz Carlos Michilis de; LIMA, Renato Abreu. Em tempos de pandemia: articulando possibilidades teórico-metodológicas na pesquisa com professores de ciências biológicas do Instituto Federal do Amazonas. **Revista Ibero-Americana de Humanidades, Ciências e Educação**, São Paulo, v. 8, n. 7, p. 1-19, 2022. DOI: 10.51891/rease.v8i7.6214.

COSTA, Levi Cadmiel Amaral da; SOUSA, Marckson Roberto Ferreira de. Análise da literatura sobre a temática “Gestão de dados” no âmbito da Ciência da Informação. **InCID: Revista de Ciência da Informação e Documentação**, Ribeirão Preto, v. 11, n. 2, p. 60-81, 2020. DOI: 10.11606/issn.2178-2075.v11i2p60-81.

HOLLER, J.; KEDRON, P. Mainstreaming metadata into research workflows to advance reproducibility and open geographic information science. **The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. XLVIII-4/W1-2022, p. 201-208, 2022. DOI: 10.5194/isprs-archives-xlviii-4-w1-2022-201-2022.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **PGD BR**: impulsionando a Ciência Aberta no Brasil. Brasília: Canal Ciência, 2024. Disponível em: [https://canalcienca.ibict.br/ciencia-em-sintese/textos-de-divulgacao/artigo/?item\\_id=28737](https://canalcienca.ibict.br/ciencia-em-sintese/textos-de-divulgacao/artigo/?item_id=28737). Acesso em: 04 ago. 2025.

MÁRDERO ARELLANO, Miguel Ángel *et al.* Vista de Planes de gestión de datos brasileños en DMPTool: caracterización y diversidad de datos científicos. **Revista Hipertext**, Barcelona, 2024.

MIKSA, Tomasz; OBLASSER, Simon; RAUBER, Andreas. Automating Research Data Management Using Machine-Actionable Data Management Plans. **ACM Transactions on Management Information Systems**, New York, 2021. DOI: 10.1145/3490396.

MIT Libraries. **Using DMPTool to manage your research data**. 2023. Disponível em: <https://libraries.mit.edu/datamanagement/dmptool>. Acesso em: 18 jul. 2024.

MONTEIRO, Elizabete Cristina de Souza de Aguiar *et al.* Plano de Gestão de Dados no Apoio à Gestão de Dados Científicos FAIR. **Revista Científica da Universidade Eduardo Mondiana**, Série: Letras e Ciências Sociais, v. 4, n. 1, pp 117-119, 2023.

OLIVEIRA, Alexandre Faria de *et al.* Preservação de Planos de Gestão de Dados: modelo Hipátia. **Revista científica UEM**: Série ciências sociais, v. 4, n. 1, p. 129-131, 2023.

REGISTRY OF RESEARCH DATA REPOSITORIES (RE3DATA.ORG). Disponível em: [www.re3data.org](http://www.re3data.org). Acesso em: 11 ago. 2025.

RIBEIRO, Nivaldo Calixto *et al.* Importância das práticas de Ciência Aberta e de comunicação científica na perspectiva de atores envolvidos. **RDBCi**: Revista Digital de Biblioteconomia e Ciência da Informação, Campinas, v. 20, p. e022019, 2022. DOI: 10.20396/rdbc.v20i00.8670366.

RIBEIRO, Nivaldo Calixto. **Ciência aberta em universidades públicas federais brasileiras**: políticas, ações e iniciativas. 2022. Tese (doutorado) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2022.

ROCHA, Lucas. As potencialidades do data paper na ciência atual. In: SHINTAKU, Milton; SALES, Luana Farias; COSTA, Michelli (org). **Tópicos sobre dados abertos para editores científicos**. Botucatu, SP: Abec, 2020. p. 159-163.

SILVA, Fabiano Couto Corrêa da; SILVEIRA, Lúcia da. O ecossistema da Ciência Aberta. **Transinformação**, Campinas, v. 31, 2019. DOI: 10.1590/2318-0889201931e190001.

STANFORD UNIVERSITY LIBRARIES. **DMPTool**: A data management plan creation tool. 2023. Disponível em: <https://library.stanford.edu/research/data-management-services/dmptool>. Acesso em: 18 jul. 2024.

STUEBER, Ketlen *et al.* Planes de gestión de datos brasileños en DMPTool: caracterización y diversidad de datos científicos.

**Hipertext.net**, Barcelona, n. 27, p. 47-56, 2023. DOI: 10.31009/hipertext.net.2023.i27.05.

UNIVERSITY OF MICHIGAN LIBRARY. **Data management planning using DMPTool**. 2023. Disponível em: [www.lib.umich.edu/research-data-management/data-management-planning](http://www.lib.umich.edu/research-data-management/data-management-planning). Acesso em: 18 jul. 2024.

VEIGA, Viviane Santos de Oliveira *et al.* Plano de gestão de dados fair: uma proposta para a Fiocruz. **Liinc em Revista**, Brasília, v. 15, n. 2, 2019. DOI: 10.18617/liinc.v15i2.5030.

VEIGA, Viviane Santos de Oliveira *et al.* Qualidade do Plano de Gestão de Dados. **Ciência da Informação**, Brasília, v. 52, n. 1, p. 1-17, 2023. DOI: 10.18225/ci.inf.v54i2.7237.

VITALE, Cynthia Hudson; SANDY, Heather Moulaision. Data Management Plans: A Review. **Desidoc Journal of Library & Information Technology**, Delhi, v. 39, n. 6, p. 15086, 2019. DOI: 10.14429/djlit.39.06.15086.

WILLIAMS, Mary; BAGWELL, Jacqueline; ZOZUS, Meredith Nahm. Data management plans: the missing perspective. **Journal of Biomedical Informatics**, San Diego, v. 71, p. 130-142, 2017. DOI: <https://doi.org/10.1016/j.jbi.2017.05.004>.

### Como citar este capítulo

SILVA, Fabiano Couto Corrêa da; SIEBRA, Sandra de Albuquerque; OLIVEIRA, Alexandre Faria de; REZENDE, Laura Vilela Rodrigues. O PGD BR como ferramenta para a promoção da ciência aberta. In: AMARO, Bianca; CAMPOS, Fhillipe de Freitas; BARCELOS, Janinne. (org.). **Infraestruturas de Ciência e de Acesso Aberto no Brasil**: iniciativas do Instituto Brasileiro de Informação em Ciência e Tecnologia. Brasília, DF: Editora Ibitc, 2025. Cap. 12, p. 189-212. DOI: 10.22477/9788570132543.cap12

# INFRAESTRUTURAS ABERTAS PARA CIÊNCIA CIDADÃ: A EXPERIÊNCIA DA CÍVIS

Sarita Albagli  
Luana Rocha  
Josir Cardoso Gomes  
Matheus Dantas



## INTRODUÇÃO

Este capítulo apresenta a Cívis, uma plataforma de ciência cidadã desenvolvida pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), com o objetivo de fornecer infraestrutura, disseminar conteúdos e se constituir em um espaço de aprendizado e inovação em ciência cidadã.

A ciência cidadã é uma das vertentes menos conhecidas dentre os movimentos pela ciência aberta, tendo se disseminado, no Brasil, desde a última década. O termo ciência cidadã tem sido usado com diferentes significados e abordagens (Albagli; Rocha, 2021; 2024). A Cívis adota uma definição ampla de ciência cidadã, que comporta a contribuição e o engajamento social na pesquisa, incluindo atividades de produção e análise de conhecimentos, dados e informações relevantes à ciência e a questões de importância social, ambiental e territorial.

Os principais propósitos da Cívis são dar visibilidade a iniciativas e métodos de ciência cidadã, com foco no Brasil e em demais países da América Latina e Caribe, bem como facilitar o acesso a ferramentas de apoio e formação para iniciativas e projetos com essa abordagem.

O capítulo apresenta um histórico e principais desenvolvimentos da Cívis, bem como desafios atuais e futuros.

## HISTÓRICO E ESTRUTURAÇÃO INICIAL

A Cívis (Ibict, 2025) foi lançada oficialmente em 27 de abril de 2022, de início em versão beta, acompanhando

um movimento internacional de constituição de infraestruturas desse tipo (Albagli; Rocha; Dantas, 2024).

O planejamento inicial do desenvolvimento e implementação da Cívís envolveu a realização de oficinas com membros da equipe<sup>4</sup> para a definição de objetivos, características, funcionalidades e públicos-alvo da plataforma; e o levantamento de plataformas digitais de ciência cidadã existentes em diferentes países, que pudesse servir de inspiração para o projeto. Esse levantamento foi feito por meio de pesquisa bibliográfica e documental, além de buscas na internet. Entende-se aqui por plataformas digitais infraestruturas online que facilitam interações entre diversos usuários, permitindo múltiplas funcionalidades (Trinca; Papi; Albagli, 2022). A pesquisa considerou plataformas digitais que atuam tanto como diretórios de informações, como as que possibilitam coleta, registro, análise e visualização de dados e informações, em formatos variados, com a contribuição de públicos variados e não acadêmicos.

Como resultado, identificaram-se infraestruturas desse tipo principalmente nos Estados Unidos e em países europeus, tais como: Anecdata (2025); Australian Citizen Science Association (ACSA, 2025); *Observatorio de la Ciencia Ciudadana En España* (2025); CitizenScience.gov (2025); CitSci.org (2025); EU-Citizen.Science (2025);

---

<sup>4</sup> A equipe inicial foi composta por Sarita Albagli e Tiago Braga (coordenadores do projeto), Luana Rocha, Leandro Chaves, Larissa Alves, Nathaly Leite Rocha, Marcos Novais, Jacqueline Coelho e Juliana Colem. Posteriormente passaram também a colaborar Josir Gomes, Fábio Ferreira da Silva, Matheus Dantas e Alisson Eugênio Costa.

iNaturalist (2025); Public Lab (2025); Scistarter (2025); SiBBr Ciência Cidadã (2025); Zooniverse (2025).

O levantamento reforçou a importância de se desenvolver uma plataforma com foco no Brasil e na América Latina. A equipe definiu então que a plataforma europeia EU-Citizen.Science serviria de base para a plataforma que seria desenvolvida pelo Ibict. As seguintes características da plataforma europeia motivaram essa decisão:

- É desenvolvida em código aberto, disponível na plataforma *GitHub* (2025), com possibilidade de adaptação do design e aprimoramento e desenvolvimento de funcionalidades, contando também com um banco de dados aberto;
- Contém diretórios diversos para registro e disponibilização de informações sobre iniciativas e plataformas de ciência cidadã, bem como materiais de apoio, com a possibilidade de tradução para diferentes idiomas;
- Possibilita a contribuição do público na plataforma, seja com o registro de conteúdo, seja com a discussão de tópicos de interesse, no blog e no fórum.

A partir da escolha da EU-Citizen.Science, definiram-se os aspectos que orientaram o desenvolvimento da plataforma brasileira, quais sejam:

- Nome e slogan: a plataforma foi batizada de “Cívís”, tendo como mote o slogan “A ciência construída por meio da colaboração entre todos/todas/todes”. A

adoção de pronomes masculino, feminino e neutro reflete a proposta de uma plataforma que se pretende inclusiva. A solução gráfica adotada para a exibição dos três pronomes foi a criação de uma imagem em movimento tipo GIF.

- Design e webdesign responsivo: foi elaborada a proposta de identidade visual para a plataforma, buscando transmitir leveza e diversidade, garantindo também a adaptação do tamanho das páginas ao tamanho da tela de diferentes dispositivos (computador, celular, *tablet*).
- Formulários de cadastro: foi feita a revisão dos formulários utilizados pela plataforma europeia para cadastro dos conteúdos, ajustando-os às características e objetivos da plataforma brasileira.
- Versão e localização do conteúdo: todo o conteúdo textual da plataforma foi revisado e adaptado para o português do Brasil e para o espanhol latino-americano. Além disso, os termos de uso e de privacidade foram adaptados para atender à legislação brasileira.
- Mapa: foi incluído um mapa para geolocalização das iniciativas cadastradas, integrado à infraestrutura do Sistema Aberto de Observatórios para Visualização de Informações (Visão) (Ibict, 2025d) (Braga; Alves; Leite, 2021).

## CONTEÚDOS DA PLATAFORMA

A Cívis possui atualmente seis grandes grupos de conteúdos:

- Iniciativas e projetos de ciência cidadã, com foco no Brasil e outros países da América Latina e Caribe, entendendo-se por iniciativas as que têm um caráter contínuo, enquanto que projetos têm objetivos e período de execução delimitados.
- Recursos, incluindo ferramentas, aplicativos e tutoriais pertinentes e úteis para iniciativas e projetos de ciência cidadã.
- Formações, abrangendo cursos e webinários orientados para introdução de conceitos e capacitação em metodologias e métodos em ciência cidadã.
- Eventos, contendo uma agenda de eventos de ciência cidadã atuais, passados e futuros, em diferentes países.
- Plataformas, entendidas como infraestruturas que atuam como diretórios de projetos e iniciativas de ciência cidadã, geralmente com abrangência nacional ou internacional, e que, em alguns casos, permitem também a alimentação de bases de dados das iniciativas ali cadastradas ou mesmo o envio de dados cidadãos aos respectivos projetos cadastrados.
- Organizações, como universidades, institutos de pesquisa e outras, em que são realizados os projetos e iniciativas de ciência cidadã cadastrados na plataforma, bem como redes de pessoas e iniciativas com atuação em ciência cidadã.
- Usuários cadastrados na plataforma e que optaram por tornar seu perfil público.

- Blog destinado à publicação de artigos curtos, entrevistas, textos de discussão e opinião, sobre temas diversos relacionados à ciência cidadã.

É possível navegar e ter acesso aos principais conteúdos da Cívís sem necessidade de cadastro prévio. Entretanto, caso o usuário queira contribuir com novos conteúdos, é necessário que faça um cadastro, informando seu nome, e-mail, idioma de preferência, biografia breve, áreas de interesse e uma foto (sendo os três últimos de preenchimento opcional).

Importante ressaltar que a plataforma tem grande preocupação com a privacidade dos usuários, tendo uma área denominada “Centro de Privacidade”, na qual quem se cadastrava pode verificar a qualquer momento quais dos seus dados estão armazenados e definir quais podem ser públicos. Além disso, foram removidas todas as ferramentas proprietárias que estavam integradas à plataforma europeia, a citar: o *Google Analytics* e o *Google Captcha*. Com as denúncias de invasão de privacidade por parte do *Google Analytics* (Dentons, 2022), decidiu-se por utilizar o software Matomo, também construído sob uma licença livre e com boa reputação documentada (Quintel; Wilson, 2020).

Na página inicial da Cívís, é possível acessar os formulários específicos para registro e submissão de cada tipo de conteúdo. O conteúdo submetido passa então pela moderação pela equipe da Cívís, de modo a garantir o uso exclusivo da plataforma para conteúdos relacionados à ciência cidadã, sem outras finalidades ou usos indevidos. Na estruturação da Cívís, houve a preocupação de prover

informações que permitissem orientar e dar transparência a seus propósitos e formas de uso, tal qual na plataforma europeia, incluindo:

- Sobre: página de apresentação da plataforma e da definição de ciência cidadã adotada.
- Perguntas frequentes: página com as principais perguntas e respostas a respeito da plataforma.
- Contato: formulário que permite ao público entrar em contato com a equipe da plataforma.
- Termos de uso: documento que define as regras para uso da plataforma.
- Licença: A licença CC BY 4.0 utilizada pela plataforma europeia foi mantida pois permite que o código fonte possa ser reutilizado e adaptado em outros projetos, contanto que os créditos sejam dados aos autores originais.
- Privacidade: Política de Privacidade da plataforma, com informações a respeito das salvaguardas no uso das informações registradas na plataforma.
- Guia de uso: guia para apoio ao usuário nos três idiomas, com orientações de funcionamento e uso das principais áreas da plataforma.

## **DESENVOLVIMENTO**

A equipe da Cívis vem trabalhando continuamente para aperfeiçoamentos na plataforma, com a atuação conjunta das equipes de conteúdo, desenvolvimento de tecnologia da informação (TI), arquitetura da informação e design. Tem contado ainda com metodologia colaborativa, aplicada no Ibict, para levantamento e identificação de

necessidades informacionais de seus projetos (Rocha; Alves; Braga, 2023). Como parte desse esforço, foi realizado, em 2023, um exercício de codesenho, em que usuários da Cívis e outros interessados fizeram sugestões de melhorias na plataforma, por meio de um formulário online.

Seguindo os passos da equipe europeia, o desenvolvimento da Cívis manteve a adoção da linguagem Python em conjunto com a biblioteca Django e o banco de dados relacional PostgreSQL. Tal escolha se deu pela segurança, flexibilidade e facilidade de uso do Python/Django para projetos dessa natureza, bem como pelo sucesso de implantações recentes no Ibict utilizando a mesma infraestrutura.

Ao longo do projeto, percebeu-se que apenas parte do sistema europeu aceitava mais de um idioma, ou seja, vários conteúdos não permitiam a tradução para o português e os *e-mails* eram enviados apenas em inglês. Dessa forma, foi necessário revisar todos os módulos, além de construir novas rotinas para que o sistema enviasse os *e-mails* no idioma selecionado pelo usuário.

Também houve um esforço para integrar a Cívis ao sistema Visão, por meio da API do sistema. Dessa forma, assim que as iniciativas são registradas, elas podem ser consultadas no mapa do Visão que se encontra na página inicial da Cívis. Da mesma forma, os dados de projetos e iniciativas registrados na Cívis ficam disponíveis no Visão, de forma que usuários do sistema possam usá-los para construir suas visualizações.

No cômputo geral, as melhorias sugeridas pelos usuários e pela equipe interna da Cívis implicaram a adição

de milhares de contribuições ao código-fonte original da EU-Citizen.Science. Tais contribuições foram comunicadas aos desenvolvedores da plataforma europeia para que avaliassem a incorporação do código da Cívis em seu sistema. Essa conduta corrobora um dos preceitos centrais do movimento de software livre que preconiza que, ao disponibilizar o software como livre, a comunidade que o utiliza passa a contribuir para a sua melhoria (Aquino, 2015).

#### a) Funcionalidades

- Ferramenta de busca - O sistema de buscas foi aperfeiçoado, de modo a facilitar a localização e ampliar o acesso ao conteúdo, permitindo assim atender a diferentes interesses e objetivos. A ferramenta de busca, antes baseada apenas nos títulos e palavras-chave, passou a contar com uma pesquisa booleana em todos os conteúdos, e respectivos campos de descrição, registrados na plataforma: projetos e iniciativas; recursos; formações; organizações, plataformas, eventos e usuários. Foi incluído também um botão para busca avançada, em caixa *pop-up* (exibida ao se clicar no botão), com explicação para uso dos operadores. Foi feita ainda a inclusão de busca na página de eventos. Essas modificações agregam valor à plataforma, permitindo cruzar dados dispersos, bem como atender a interesses por diferentes tipos de conteúdo.
- Mapa e geolocalização - Foi implementado no Mapa de Iniciativas, integrado ao Visão, o filtro por temas,

o que possibilita aos usuários da plataforma refinar as buscas dos projetos ali exibidos.

- Estatísticas de Uso - Para monitoramento das visitações e comportamento dos visitantes na plataforma foi adotada a ferramenta de código aberto Matomo, no lugar de ferramentas como o *Google Analytics*. Além disso, foi incorporada à Cívis a ferramenta de estatísticas de projetos desenvolvida pela plataforma europeia, que permite visualizar quantas vezes uma iniciativa foi vista, curtida e adicionada aos favoritos de usuários da plataforma.

### **b) Engajamento e melhoria da interface**

- Ferramenta de coedição - Os conteúdos iniciais da Cívis foram registrados por sua equipe, com base em pesquisas prévias e atualizações contínuas. A partir daí os responsáveis pelos projetos e iniciativas foram convidados a editar e atualizar as informações cadastradas. Para isso, a equipe da Cívis desenvolveu uma ferramenta específica para convite e inclusão de novos editores em iniciativas e projetos registrados na plataforma, permitindo a mais de um usuário cadastrado gerir e editar as informações de um projeto - originalmente, apenas o responsável pelo cadastro inicial tinha permissão para tal. Essa funcionalidade também colabora para aumentar o engajamento, na Cívis, de gestores e responsáveis por projetos de ciência cidadã, bem como para a divulgação da plataforma e de seus recursos junto ao público interessado e envolvido no

tema. Caso essas pessoas queiram atuar como editores das informações sobre as iniciativas que lideram, é necessário que se cadastrem na plataforma, o que eleva também o número de usuários da Cívis.

- *Newsletter* - Em 2023, foi lançada a *newsletter* da Cívis, destacando as principais novidades da plataforma. Para tanto, foi desenvolvida uma ferramenta integrada à área de gestão da plataforma, com um formulário próprio para edição e seleção dos conteúdos, possibilitando, ainda, o envio do boletim de forma automática. Foi criado um projeto gráfico do *newsletter*, além de um cadastro de e-mails independente do banco de usuários registrados na plataforma, permitindo o envio do boletim a um público mais amplo. Seguindo as boas práticas para envio de boletins eletrônicos, as pessoas incluídas no cadastro de e-mails do *newsletter* podem optar por continuar a receber ou não o informativo. A série completa de *newsletters* publicadas na Cívis está acessível no menu localizado no rodapé da plataforma. As edições dos *newsletters* incluem, entre outros conteúdos: a postagem mais recente no blog da Cívis; um recurso de ciência cidadã registrado na plataforma; os próximos eventos; e os projetos e iniciativas cadastrados na Cívis mais recentemente. Em suas últimas edições, o *newsletter* teve uma taxa de abertura de 64%, bem acima da taxa média de 20% reportada pela literatura para campanhas sem personalização ou segmentação (Jenkins, 2009).

- Fórum - Ainda em fase de teste, o Fórum da Cívis pretende oferecer um espaço para o debate, a troca de experiências e a constituição de redes temáticas e comunidades de prática em ciência cidadã no Brasil e demais países da América Latina e do Caribe. Após pesquisa sobre diferentes tipos de fórum como estratégia para estabelecer comunidades online, optou-se por manter o modelo adotado pela plataforma europeia, o django-machina. Na Cívis, desenvolveu-se uma funcionalidade que permite a qualquer usuário cadastrado na plataforma propor a criação de um fórum de discussão. As propostas são moderadas pela equipe da Cívis e, em caso de aprovação, o proponente se torna o administrador daquele fórum, sendo o responsável por monitorar e sinalizar, à equipe da plataforma, caso haja a publicação de conteúdos impróprios e/ou destoantes dos objetivos do Fórum da Cívis e da própria plataforma. Em desenvolvimento estão o estabelecimento de uma política de uso e administração do Fórum e ajustes no código para permitir um motor de busca próprio. Um desafio é a adoção de estratégia de incentivo ao uso do fórum pelos atuais e novos usuários da Cívis.
- Compartilhamento e pré-visualização nas redes sociais: Para dar mais visibilidade à Cívis nas redes sociais, incluiu-se um botão de compartilhamento, nos diversos conteúdos registrados na plataforma, que permite o compartilhamento de suas páginas web por link ou diretamente. Além disso, foram

criadas pré-visualizações de parte do conteúdo das páginas compartilhadas.

### **c) Cooperação nacional e internacional**

A busca de parcerias nacionais e internacionais tem sido uma estratégia da Cívis desde seu início. O desenvolvimento das versões em inglês e espanhol (Ibict, 2025b; 2025c) da plataforma, com tradução de todos seus conteúdos fixos, é parte dessa estratégia.

Destacam-se as parcerias com a Rede Brasileira de Ciência Cidadã (RBCC, 2025); a Rede Iberoamericana de Ciência Participativa (Ricap); a Eu-Citizen.Science; a Parceria do Governo Aberto; e o Instituto Nacional de Ciência Cidadã (INCC), um dos Institutos Nacionais de Ciência e Tecnologia (INCT) apoiados pelo MCTI e CNPq, em parceria com a Capes e as Fundações de Amparo à Pesquisa (FAPs).

## **CONSIDERAÇÕES FINAIS**

Desde seu lançamento, a Cívis logrou alguns incrementos importantes, particularmente nos seguintes eixos: novas funcionalidades; melhor interação e engajamento com os usuários; cooperação nacional e internacional; uso de códigos abertos, ao mesmo tempo em que garante segurança e privacidade.

Figura 1 - Evolução do número de iniciativas e usuários cadastrados na Cívis



Fonte: Cívis.

O esforço de divulgação e engajamento na plataforma, desde seu lançamento, resultou num expressivo crescimento dos conteúdos cadastrados. Em abril de 2022, quando foi lançada, a Cívis contabilizava 19 projetos e iniciativas e 23 usuários cadastrados. Em julho de 2024, esses números passaram para 222 projetos e iniciativas e 292 usuários.

Algumas novas frentes estão em estudo, em planejamento e em desenvolvimento, contando com a colaboração de parcerias de outras organizações e de diversas áreas do Ibict, tais como:

- O estabelecimento de um repositório de dados para projetos e iniciativas de ciência cidadã;
- O incremento de ferramentas, guias e materiais de informação e capacitação em ciência cidadã, incluindo protocolos e regimes de governança dos dados produzidos, bem como políticas de gestão, compartilhamento e abertura de dados, levando em

conta os princípios FAIR de abertura de dados (dados encontráveis, acessíveis, interoperáveis, reutilizáveis), e CARE (tratamento ético, responsável e equitativo dos dados);

- A ativação do Fórum como ambiente de troca e aprendizado;
- A organização de dossiês sobre temas de destaque que envolvem abordagens de ciência cidadã.

## REFERÊNCIAS

ALBAGLI, Sarita; ROCHA, Luana. Ciência cidadã no Brasil: um estudo exploratório. In: BORGES, Maria Manuel; CASADO, Elias Sanz. (org.). **Sob a lente da Ciência Aberta: olhares de Portugal, Espanha e Brasil**. Coimbra: Universidade de Coimbra, 2021, p. 489-511.

ALBAGLI, Sarita; ROCHA, Luana. Ciência cidadã: um conceito polissêmico. **Boletim do Museu de Biologia prof. Mello Leitão**, v. 40, n. 1, 2024

ALBAGLI, Sarita; ROCHA, Luana; DANTAS, Matheus. Cívis: plataforma de ciência cidadã. In: CORRÊA, Fabiano Couto Corrêa da; CARVALHO-SEGUNDO, Washington Luís Ribeiro de; STUEBER, Ketlen. **Ciência Aberta no Brasil: conquistas e desafios**. Porto Alegre: Letra 1, 2024.

ANECDATA. Disponível em: [www.anecdata.org](http://www.anecdata.org). Acesso em: 11 ago. 2025.

AQUINO, Jakson Alves de. Software livre e desenvolvimento de trabalhos científicos: o R como exemplo a ser seguido. **Revista Política Hoje**, Recife, v. 24, n. 2, p. 75-86, 2015.

AUSTRALIAN CITIZEN SCIENCE ASSOCIATION (ACSA).

Disponível em: <https://biocollect.ala.org.au/acsa>. Acesso em: 11 ago. 2025.

BRAGA, Tiago Emmanuel Nunes; ALVES, Larissa de Araújo; LEITE, Nathaly Cristine. Ferramentas de visualização de dados e informações e suas características. **Tendências da Pesquisa Brasileira em Ciência da Informação**, São Paulo, v. 14, 2022.

CITIZENSCIENCE. Disponível em: [www.citizenscience.gov](http://www.citizenscience.gov). Acesso em: 11 ago. 2025.

CITSCI. Disponível em: [www.citsci.org](http://www.citsci.org). Acesso em: 11 ago. 2025.

DENTONS Law Firm. **Google Analytics and privacy concerns surrounding it**. [s. l.: s. n.], 2022. Disponível em: [www.dentons.com/en/insights/articles/2022/march/3/google-analytics-and-privacy-concerns-surrounding-it](http://www.dentons.com/en/insights/articles/2022/march/3/google-analytics-and-privacy-concerns-surrounding-it). Acesso em: 19 ago. 2024.

EU-CITIZEN.SCIENCE. Disponível em: <http://eu-citizen.science>. Acesso em: 11 ago. 2025.

GITHUB. Disponível em: [https://github.com/Ibercivis/EU-CS\\_platform](https://github.com/Ibercivis/EU-CS_platform). Acesso em: 11 ago. 2025.

INATURALIST. Disponível em: [www.inaturalist.org](http://www.inaturalist.org). Acesso em: 11 ago. 2025.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Cívis**: plataforma de Ciência. Brasília: Ibict, 2025. Disponível em: <https://civis.ibict.br>. Acesso em: 11 ago. 2025.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Cívis** citizen science platform. Brasília: Ibict, 2025b. Disponível em: <https://civis.ibict.br/en>. Acesso em: 11 ago. 2025.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Cívis**: plataforma de ciencia ciudadana. Brasília: Ibict, 2025c. Disponível em: <https://civis.ibict.br/es>. Acesso em: 11 ago. 2025.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Conheça o Visão**. Brasília: Ibict, 2025d. Disponível em: <https://visao.ibict.br>. Acesso em: 11 ago. 2025.

JENKINS, Simms. **The truth about email marketing**. Upper Saddle River, N.J: FT Press, 2009.

OBSERVATORIO DE LA CIENCIA CIUDADANA EN ESPAÑA. Zaragoza. Disponível em: <https://ciencia-ciudadana.es>. Acesso em: 11 ago. 2025.

PUBLIC LABORATORY FOR OPEN TECHNOLOGY AND SCIENCE (PUBLICLAB). Disponível em: <https://publiclab.org>. Acesso em: 11 ago. 2025.

QUINTEL, Denise FitzGerald; WILSON, Robert. Analytics and Privacy: Using Matomo in EBSCO's Discovery Service. **Information Technology and Libraries**, v. 39, n. 3, 21 set. 2020. DOI: 10.6017/ital.v39i3.12219.

REDE BRASILEIRA DE CIÊNCIA CIDADÃ (RBCC). **Home**. Disponível em: [www.rbcc.org.br](http://www.rbcc.org.br). Acesso em: 11 ago. 2025.

ROCHA, Nathaly; ALVES, Larissa de Araújo; BRAGA, Tiago Emmanuel Nunes. Metodologia colaborativa para levantamento de necessidades informacionais em projetos de pesquisa. In: WORKSHOP DE INFORMAÇÃO, DADOS E TECNOLOGIA, 6., 2023. Brasília. **Anais** [...], Brasília: Ibict, 2023. DOI: 10.22477/vi.widat.76.

SCISTARTER. Disponível em: <https://scistarter.org>. Acesso em: 11 ago. 2025.

SISTEMA DE INFORMAÇÃO SOBRE A BIODIVERSIDADE BRASILEIRA (SiBBr). Disponível em:  
<https://sibbr.gov.br/cienciacidada/projetos.html>. Acesso em: 11 ago. 2025.

TRINCA, Tatiane Pacanaro; PAPI, Miguel Enrique Silveira; ALBAGLI, Sarita. Capitalismo de plataforma e plataformização na ciência. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO. 22., 2022. Porto Alegre. **Anais** [...], Porto Alegre: UFRGS, 2022.

ZOONIVERSE: People-Powered Research. Disponível em: [www.zooniverse.org](http://www.zooniverse.org). Acesso em: 11 ago. 2025.

### Como citar este capítulo

ALBAGLI, Sarita; ROCHA, Luana; GOMES, Josir Cardoso; DANTAS, Matheus. Infraestruturas abertas para ciência cidadã: a experiência da Cívis. In: AMARO, Bianca; CAMPOS, Fhillipe de Freitas; BARCELOS, Janinne. (org.). **Infraestruturas de Ciência e de Acesso Aberto no Brasil**: iniciativas do Instituto Brasileiro de Informação em Ciência e Tecnologia. Brasília, DF: Editora Ibiti, 2025. Cap. 13, p. 213-231. DOI: 10.22477/9788570132543.cap13



# **ECOSSISTEMA DE INFORMAÇÃO DA PESQUISA CIENTÍFICA BRASILEIRA (BrCris)**

Thiago Magela Rodrigues Dias  
Washington Luís R. de Carvalho Segundo  
Marcel Garcia de Souza



## APRESENTAÇÃO

O Brasil tem uma parcela relevante na produção científica internacional principalmente em nichos específicos ao atuar em sua economia. O país é líder na produção do conhecimento no contexto de América Latina (Collazo-Reyes, 2014) e um atrator de talentos no contexto regional (Saraiva; Miranda, 2004). Destaca-se pela sua implementação de plataformas digitais de registro nacional do atuar de seus pesquisadores.

O destaque internacional é a Plataforma Lattes, em que se observa a importância estratégica de ter as informações científicas curriculares disponíveis de forma ampla (Lane, 2010). Similar à Plataforma Lattes existem outras plataformas nacionais que registram parte do atuar acadêmico e tecnológicos, como o Sucupira da Capes, o Banco de teses e dissertações do Ibict, o Banco de propriedade industrial do INPI e Portal Transparência do Governo Federal, por exemplo. Embora todas as informações dessas fontes de dados sejam abertas e livres para consulta, elas estão restritas a suas bases de dados, não existindo uma integração de forma a possibilitar a exploração dessas informações de forma ampla a ciência e a sociedade do Brasil.

Neste contexto, é importante destacar que a informação proveniente de pesquisas em Ciência, Tecnologia e Inovação (CT&I) representa um dos principais pilares para o desenvolvimento econômico e social de um país. Essa informação constitui-se como um elemento orientador fundamental para a geração sistemática de conhecimento, tanto teórico quanto aplicado.

A informação gerada no âmbito de Ciência, Tecnologia e Inovação (CT&I) é altamente especializada, distinguindo-se de outros tipos de informação. Sua produção é fundamentada em um método específico - o método científico - e a divulgação de seus resultados seguem procedimentos distintos, abrangendo avaliação, validação, publicação e acesso por meio de fontes especializadas (Meadows, 1999).

A diversidade de fontes de informação, juntamente com a variedade de seus modelos de dados (metadados), resultante das pesquisas em Ciência, Tecnologia e de Inovação (CT&I) são frequentemente armazenados em bases de dados e repositórios que tem estruturas de dados distintas, visando assegurar a preservação, visibilidade e recuperação eficaz das informações contidas, e em poucos casos objetivam a integração.

Complementando as fontes de informações brasileiras, existem fontes com abrangência internacional, que possibilitam a troca de dados tanto para seres humanos como para computadores. Dentre essas bases de dados pode-se citar o Wikidata, o Crossref, OpenCitations, OpenAIRE Research Graph, o Latindex e o DOAJ entre outras fontes abertas de informação em CT&I.

Reunir e integrar os dados dessas e de outras fontes é um desafio que exige um grande esforço intelectual e um grande poder computacional. A capacidade de armazenamento necessária ultrapassa a que um sistema gerenciador de banco de dados tradicional consegue suportar, exigindo soluções tecnológicas avançadas, provenientes da chamada Web 4.0, bem como, de análises que envolvam técnicas de inteligência computacional.

A partir desse cenário, surgiram iniciativas voltadas para a criação de sistemas destinados a gerenciar a produção acadêmica, sejam eles institucionais, nacionais ou temáticos. Esses sistemas são conhecidos pela sigla CRIS. CRIS (*Current Research Information System*) define um sistema de informação abrangente sobre todo o ecossistema do processo científico.

O modelo CRIS organiza em um só lugar todas as informações do ciclo da pesquisa Científica, desde o Fomento, passando pelos projetos, pesquisadores, instituições de pesquisa e laboratórios, até os outputs de uma pesquisa científica, tais como artigos científicos, teses, dissertações, livros, capítulos de livro, patentes e conjuntos de dados científicos (Sivertsen, 2019).

Iniciativas de mapeamento do ecossistema ocorrem na Europa com o *Directory of Research Information Systems* (DRIS) executado pela euroCRIS (Eurocris, 2023), no Brasil, o BrCris é o mapeamento do ecossistema de informação da pesquisa científica brasileira. Sua concretude se faz com uma plataforma agregadora de diversas fontes de informações, o que permite recuperar, certificar e visualizar dados e informações relativas aos diversos atores que atuam na pesquisa científica do contexto brasileiro. Dentre as principais fontes estão os dados curriculares de indivíduos, organizações, programas de pós-graduação, publicações, orientações acadêmicas, revistas científicas, patentes, grupos de pesquisa, softwares e outras fontes que ainda serão agregadas (Dias; et al., 2022).

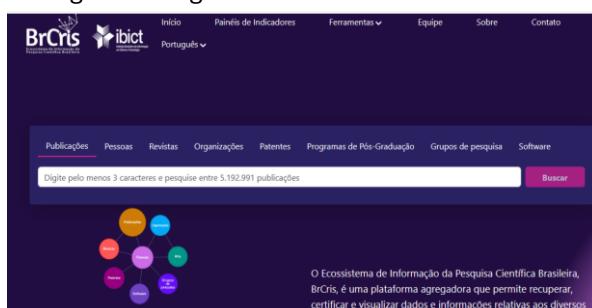
Logo, o BrCris tem por objetivo estabelecer um modelo único de organização da informação científica de todo o ecossistema de pesquisa brasileira. Entre os agentes

deste ecossistema estão os pesquisadores, os projetos, as infraestruturas, os laboratórios e as instituições de pesquisa, os financiadores, além dos resultados da pesquisa expressos principalmente por publicações científicas, teses, dissertações, conjuntos de dados científicos, software e patentes.

## HISTÓRICO E DESENVOLVIMENTO

O BrCris oferece uma interface unificada de busca de informações, a visualização de redes de colaboração e painéis de indicadores em ciência, tecnologia e inovação (Figura 1). E estabelece um modelo único de organização da informação científica de todo o ecossistema da pesquisa brasileira (Dias; et al., 2022). Entre os agentes deste ecossistema estão os pesquisadores, os projetos, infraestruturas, laboratórios e instituições de pesquisa, os financiadores, além dos resultados da pesquisa expressos principalmente por publicações científicas, teses, dissertações, conjuntos de dados científicos e patentes (Kong, et al., 2019).

Figura 1 - Página Inicial da Plataforma BrCris



Fonte: Ibict (2024).

A idealização do Projeto do Sistema BrCris, que é o CRIS no contexto da ciência brasileira, data de 2014. Inspirado no modelo proposto por Portugal de um CRIS nacional (o PTCRIS - <https://ptcris.pt>), o Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) iniciou uma série de estudos e parcerias interinstitucionais para a execução do Projeto. Em 2020, houve a implementação formal de um Projeto de Pesquisa para a construção do BrCris, com o intuito de fornecer ferramentas tecnológicas para a comunidade acadêmica brasileira com dados consolidados da produção científica nacional, tomando como base outros projetos CRIS e padrões internacionais disponibilizados pelo OpenAire e COAR.

O BrCris adota um esquema de representação de dados em dois níveis. O primeiro é o nível lógico, materializado como um modelo de entidades e relacionamentos baseado no CERIF (Jörg, 2010), também chamado de metamodelo. Este modelo é traduzido para um esquema relacional físico na plataforma LA Referencia, onde os dados são carregados e processados. Trata-se, portanto, de um modelo que atende às necessidades internas de armazenamento de dados, sendo responsável pela organização e integração das informações coletadas para que se tornem insumos para os objetivos do projeto (Pinto *et al.*, 2021).

Em se tratando do modelo de dados para este projeto, iniciou-se pela adoção de nove entidades de dados, seguindo padrões amplamente utilizados na comunidade científica internacional.

Após todo o processo de tratamento, independentemente da fonte de dados, os dados gerados

como saída, são importados em um único banco de dados, e utilizando-se dos identificadores únicos gerados ou identificados, os conjuntos de dados são vinculados e deduplicados, viabilizando dessa forma a interoperabilidade dos dados, independentemente de sua fonte e formato.

A criação do metamodelo de entidades e relacionamento levou em consideração as características dos diversos conjuntos de dados coletados, visando a facilitar as rotinas de tratamento dos dados, em especial a deduplicação de entidades. Porém, neste formato, as informações resultantes não seriam facilmente acessadas e reutilizadas por agentes externos. Para resolver este problema, foi desenvolvido o segundo nível de representação, o nível semântico, implementado como uma ontologia para permitir a visualização e navegação dos dados como um grafo de conhecimento.

Na criação do modelo semântico foram adotadas como premissas a maximização do reuso de recursos existentes e a utilização de padrões internacionais para representação de dados na área, para que o BrCris fosse compatível com sistemas similares ao redor do mundo.

O modelo semântico do BrCris é composto por um subconjunto da ontologia VIVO, composta pelas classes e propriedades equivalentes às entidades, atributos e relacionamentos do metamodelo lógico, acrescido de uma extensão local que cobre informações específicas do contexto brasileiro.

Atualmente, o BrCris oferece três métodos distintos para navegação e visualização de seus dados. O primeiro envolve a utilização de sua interface de busca, que permite

pesquisas abrangentes em todas as entidades da plataforma, inclusive com a aplicação de conectores lógicos para buscas mais específicas. Além disso, a visualização e exportação dos dados é facilitada pelos *dashboards* atualmente disponíveis. Por fim, o BrCris também possibilita a navegação entre as entidades utilizando a plataforma VIVO.

Como parte do sistema de recuperação de informações, foi desenvolvida uma interface gráfica web baseada na *Search-UI* da *Elastic*. A *Search-UI* é uma biblioteca de código aberto escrita em Typescript, que fornece uma variedade de componentes web personalizáveis compatíveis com aplicativos desktop e móveis. Essa biblioteca integra-se ao Elasticsearch, oferecendo uma interface abrangente para busca e visualização de informações na web.

A *Search-UI* disponibiliza um conjunto de componentes de busca configuráveis, simplificando a criação de interfaces altamente personalizáveis com funcionalidades avançadas de filtragem. Entre essas funcionalidades estão recursos como paginação, digitação preditiva, autocompletar, filtros, classificação e geração de gráficos de indicadores. Essa abordagem visa facilitar aos usuários a localização precisa do que necessitam, promovendo uma experiência de busca eficiente e aprimorada (Figura 2).

A interface de busca do BrCris oferece uma série de funcionalidades avançadas para a pesquisa e visualização de informações. Ao acessar a interface, os usuários podem realizar buscas detalhadas utilizando a caixa de pesquisa,

que aceita termos específicos ou operadores lógicos para refinar os resultados.

Figura 2 - Interface de Busca de Publicações do BrCris



Fonte: Ibict (2024).

A página também permite a personalização da quantidade de resultados exibidos por página e a ordenação das publicações por diferentes critérios, como a data de publicação em ordem decrescente, garantindo que as informações mais recentes sejam facilmente acessíveis.

Além das funcionalidades de busca, a página de publicações do BrCris apresenta opções avançadas de filtragem e visualização. Os usuários podem aplicar filtros para restringir os resultados a critérios específicos, como autor, instituição ou tipo de documento.

A interface também inclui a opção de exportar os dados das publicações, permitindo que os usuários baixem as informações em formatos compatíveis para análise externa. A apresentação dos resultados é clara e organizada, facilitando a navegação entre as diferentes

publicações e proporcionando uma experiência de usuário eficiente e intuitiva.

A implementação de diversas visualizações em um conjunto de *dashboards* se caracteriza como um importante mecanismo de exportação de dados, oferecendo uma perspectiva abrangente sobre a produção científica nacional.

Essas visualizações permitem a aplicação de filtros e outros métodos de personalização em todos os atributos de uma entidade específica, proporcionando aos usuários a capacidade de realizar análises detalhadas tanto em termos temporais quanto em áreas específicas ou tópicos de interesse (Figura 3).

Além disso, a aplicação desses filtros gera um conjunto de dados específicos, resultantes do processo de filtragem, que podem ser exportados em formatos padronizados, como .csv. Isso facilita a manipulação e análise dos dados fora da plataforma, permitindo que pesquisadores e analistas explorem as informações de maneira mais aprofundada e integrada com outras ferramentas e sistemas.

Com essa funcionalidade, o BrCris se destaca como uma ferramenta flexível para o mapeamento e análise da produção científica, atendendo às necessidades de uma comunidade científica diversificada e em constante evolução.

Os painéis de indicadores do BrCris proporcionam uma visualização detalhada e interativa da produção científica nacional, oferecendo diversas possibilidades de análise e personalização. Os usuários podem explorar os dados por meio de gráficos, tabelas e outras

representações visuais que facilitam a compreensão de padrões e tendências. A interface permite a aplicação de múltiplos filtros, como por período de tempo, instituição, área do conhecimento e tipo de publicação, possibilitando uma análise refinada e direcionada conforme os interesses específicos de cada usuário.

Figura 3 - Painéis de Visualização e Exportação de Dados



Fonte: Ibict (2024).

Além das opções de visualização, os painéis de indicadores do BrCris oferecem funcionalidades de exportação de dados. Os usuários podem exportar os conjuntos de dados filtrados em formatos padronizados, como .csv, o que permite uma análise mais aprofundada, utilizando outras ferramentas e softwares de análise. Essa capacidade de exportação facilita a integração dos dados com diferentes sistemas e apoia a realização de pesquisas e estudos mais detalhados, tornando o BrCris uma ferramenta essencial para pesquisadores, gestores e formuladores de políticas na área de ciência e tecnologia.

E ainda, a plataforma VIVO oferece diversas visualizações que desempenham um papel crucial na exploração detalhada dos dados. Essas visualizações não apenas facilitam a navegação pelas informações de uma entidade, mas também revelam seus relacionamentos com outras entidades, proporcionando uma análise interativa dos dados. Através da plataforma, é possível acessar todas as informações pertinentes à entidade de interesse e seus vínculos. Detalhes como publicações, áreas de atuação e redes de coautoria de um pesquisador específico são prontamente acessíveis, permitindo uma visão abrangente de sua contribuição acadêmica e colaborativa (Figura 4).

**Figura 4 - Visualização de uma entidade na Plataforma VIVO**

### Thiago Magela Rodrigues Dias

Doutor em Modelagem Matemática e Computacional pelo CEFET-MG (2016) tendo trabalhado com Bibliometria, Extração de Dados Científicos e Análise de Redes de Colaboração Científica. Mestre em Modelagem Matemática e Computacional pelo CEFET-MG (2008), trabalhando com Arquitetura Orientada a Serviços e Web Mining com o tema: Uma Arquitetura Orientada a Serviços para emprego em Sistemas de Mineração de Dados na Web. Possui graduação em Ciência da Computação pelo Centro Universitário de Formiga ([... more](#))

PhD in Mathematical and Computational Modeling from CEFET-MG (2016) having worked with Bibliometrics, Scientific Data Extraction and Analysis of Scientific Collaboration Networks. Master in Mathematical and Computational Modeling from CEFET-MG (2008), working with Service Oriented Architecture and Web Mining with the theme: A Service Oriented Architecture for use in Data Mining Systems on the Web. Graduated in Computer Science from Centro Universitário de Formiga - UNIFOR (2004), in addition ([... more](#))

#### Áreas De Investigação

[Banco de Dados](#) | [Ciéncia da Computação](#) | [Engenharia de Software](#) | [Métodos Quantitativos. Bibliometria](#) |  
[Sistemas de Informação](#) | [Técnicas de Recuperação de Informação](#) |



Fonte: Ibict (2024).

A visualização oferecida pela plataforma VIVO é importante para a análise e compreensão da produção científica, tecnológica e das redes de colaboração científica. Ela permite obter uma visão detalhada e integrada das contribuições individuais e coletivas dentro do ambiente acadêmico. Ao apresentar informações de forma gráfica e interativa, a plataforma facilita a

identificação de padrões, tendências e conexões que, de outra forma, não poderiam ser facilmente identificadas, principalmente em detrimento de análises tradicionais baseadas apenas em dados tabulares.

A representação gráfica de dados, como coautoria e áreas de atuação, não só revela as interações existentes, mas também pode sugerir novas oportunidades de parceria e colaboração. Isso é particularmente valioso em um contexto de pesquisa cada vez mais globalizado e interconectado, onde a sinergia entre diferentes áreas do conhecimento pode levar a inovações significativas. Assim, a plataforma VIVO não apenas melhora a transparência e a acessibilidade das informações científicas, mas também potencializa o avanço do conhecimento ao fomentar um ambiente colaborativo e bem informado.

Um exemplo de análise específica proporcionada pela plataforma VIVO é a exploração da rede de colaboração de um pesquisador, onde informações detalhadas sobre a rede são apresentadas, acompanhadas por uma representação gráfica em forma de grafo (Figura 5).

Figura 5 - Rede de Colaboração Científica de um Pesquisador



Fonte: Ibict (2024).

Esse grafo visualiza as conexões entre pesquisadores, destacando colaborações e coautorias, o que facilita a identificação de parcerias. Além disso, essas visualizações interativas permitem aos usuários explorar dados de forma intuitiva, ajustando os filtros e parâmetros conforme necessário para obter uma compreensão mais profunda e personalizada das redes de pesquisa e dos impactos acadêmicos. Assim, a plataforma VIVO se torna uma ferramenta poderosa para pesquisadores, gestores e instituições acadêmicas na análise e valorização da produção científica.

## **CONSIDERAÇÕES FINAIS**

O BrCris configura-se como um espaço vital para a pesquisa e análise de dados. As informações agregadas e organizadas segundo um modelo de dados semântico permitem a geração de serviços para diversos atores, nos contextos de gestão e pesquisa acadêmica, bem como na área de informação para a inovação. Esta plataforma é uma iniciativa que coleta e enriquece dados de repositórios e bases de dados abertos, sendo uma proposta singular no mundo. Ela facilita a obtenção de um panorama abrangente da produção e atuação dos atores acadêmicos e científicos brasileiros. No entanto, a operação do BrCris requer significativos recursos computacionais e humanos para o tratamento e padronização desses dados.

Diante do exposto, os resultados obtidos com o BrCris não apenas complementam, mas também expandem significativamente a compreensão da produção científica brasileira. Destaca-se a relevância e o potencial

do BrCris na promoção de uma visão abrangente e aprofundada da pesquisa no Brasil, evidenciando como suas funcionalidades avançadas de visualização e exportação de dados possibilitam análises detalhadas e personalizadas. Ao integrar diversas fontes de dados e oferecer ferramentas para a exploração e interpretação dessas informações, o BrCris se consolida como uma importante plataforma, contribuindo para um entendimento mais completo e estratégico da ciência e tecnologia no país.

## REFERÊNCIAS

COLLAZO-REYES, Francisco. Growth of the number of indexed journals of Latin America and the Caribbean: the effect on the impact of each country. **Scientometrics**, Budapest, v. 98, p. 197-209, 2014.

DIAS, Thiago Magela Rodrigues *et al.* BrCris: plataforma para integração, análises e visualização de dados técnicos-científicos. p. 622-638, **Informação e Informação**, Londrina, v. 27, n. 3, 2022. DOI: 10.5433/1981-8920.2022v27n3p622.

EUROCRIS. **Directory of Research Information Systems (DRIS)**. Nijmegen, Netherlands. Disponível em: <https://eurocris.org/services/dris>. Acesso em: 23 out. 2023.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Ecossistema de Informação da Pesquisa Científica Brasileira, BrCris**. Disponível em: <https://brcris.ibict.br>. Acesso em: 23 out. 2024.

JÖRG, Brigitte. CERIF: The common European research information format model. **Data Science Journal**, London, v. 9, p. CRIS24-CRIS31, 2010.

- KONG, Xiangjie *et al.* Academic social networks: Modeling, analysis, mining and applications. **Journal of Network and Computer Applications**, London, v. 132, p. 86-103, 2019.
- LANE, Julia. Let's make science metrics more scientific. **Nature**, London, v. 464, n. 7288, p. 488-489, 2010.
- MEADOWS, Arthur Jack. **A comunicação científica**. Trad. A. A. B. de Lemos. Brasília: Briquet de Lemos, 1999.
- PINTO, Adilson Luiz *et al.* The Brazilian current research information system: BrCris. *In:* SILVA, Carlos Guardado da Silva; REVEZ, Jorge; CORUJO, Luis (coord.). **Organização do conhecimento no horizonte 2030: desenvolvimento sustentável e saúde**. Lisboa: Universidade de Lisboa, 2021. ISBN 9789895661374. p. 319-330. DOI: 10.51427/10451/50067.
- SARAVIA, Nancy Gore; MIRANDA, Juan Francisco. Plumbing the brain drain. **Bull World Health Organ**, Genebra, v. 82, n. 8, set. 2004.
- SIVERTSEN, Gunnar. Developing Current Research Information Systems (CRIS) as data sources for studies of research. *In:* GLÄNZEL, Wolfgang *et al.* (ed.). **Springer handbook of science and technology indicators**. Amsterdã: Springer, Cham. 2019. p. 667-683. DOI: 10.1007/978-3-030-02511-3\_25.
- Como citar este capítulo
- DIAS, Thiago Magela Rodrigues; CARVALHO SEGUNDO, Washington Luís R. de; SOUZA, Marcel Garcia de. Ecossistema de Informação da Pesquisa Científica Brasileira (BrCris). *In:* AMARO, Bianca; CAMPOS, Phillip de Freitas; BARCELOS, Janinne. (org.). **Infraestruturas de Ciência e de Acesso Aberto no Brasil: iniciativas do Instituto Brasileiro de Informação em Ciência e Tecnologia**. Brasília, DF: Editora Ibict, 2025. Cap. 14, p. 233-248. DOI: 10.22477/9788570132543.cap14

# **REDE BRASILEIRA DE REPOSITÓRIOS DIGITAIS (RBRD)**

Juliana Araujo Gomes de Sousa  
Bianca Amaro



## INTRODUÇÃO

O Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), desde o princípio dos anos 2000, vem empunhando a bandeira do Acesso Aberto/Ciência Aberta no Brasil. Comprometido com essa nova maneira de fazer ciência, o Ibict leva a cabo uma série de ações para auxiliar na estruturação da infraestrutura da chamada via verde, ou seja, na construção de repositórios digitais nas instituições de ensino e pesquisa nacionais.

Uma ação estruturante do Ibict para impulsionar a criação de repositórios foi o lançamento do Edital de Chamada Finep/PCAL/XBDB nº 001/002/2009, que previa a doação de servidores às instituições públicas de ensino e pesquisa, atrelada a um compromisso de criar a infraestrutura necessária para a implementação de um repositório institucional da produção científica.

Após o lançamento do edital e com a disseminação e reconhecimento dos benefícios promovidos pelos preceitos e estratégias do Movimento de Acesso Aberto (MAA) várias instituições de ensino e pesquisa, em todas as regiões do Brasil iniciaram a implementação de repositórios institucionais. Por tratar-se da necessidade de absorção de uma nova tecnologia, como também a criação de uma estrutura informacional, que deveria estar acorde com os padrões internacionais de interoperabilidade, o Ibict passou a ser fortemente demandado a prestar apoio por parte das instituições.

As demandas ao Ibict de auxílio e suporte para a criação de repositórios aumentaram exponencialmente e então o Instituto,

visando atuar de maneira coordenada e otimizar o atendimento às instituições, decidiu pela criação da Rede Brasileira de Repositórios Institucionais de Publicações Científicas em Acesso Aberto – RIAA (Amaro, 2019).

A criação da Rede Brasileira de Repositórios Institucionais de Publicações Científicas em Acesso Aberto (RIAA) foi também motivada por iniciativas de fortalecimento do movimento de acesso aberto à informação científica no Brasil. Os principais objetivos da Rede incluíam: impulsionar a criação de subredes regionais; definir e disseminar padrões; estimular a colaboração e compartilhamento; disseminar critérios de qualidade; promover e impulsionar a capacitação de instituições; realizar o acompanhamento/monitoramento dos repositórios. Todos os objetivos tinham a finalidade de ampliar a visibilidade da produção científica nacional, garantir acesso público e gratuito a publicações financiadas por recursos públicos e promover a interoperabilidade entre repositórios institucionais.

A gênese de criação da RIAA foi a I Reunião sobre Repositórios Institucionais da Região Norte, promovida pela Biblioteca Central da Universidade Federal do Pará (UFPA). Na ocasião, integrantes de instituições da região discutiram, entre outros temas, a necessidade de potencializar a visibilidade da produção científica e o resultado foi a elaboração da chamada *Carta de Belém* (2014). A Carta de Belém, que deu origem à criação da Rede Norte de Repositórios Digitais, foi motivada pela necessidade de aumentar a visibilidade e acessibilidade da

produção científica da Região Norte. O documento surgiu de reflexões sobre os desafios regionais, como a limitada infraestrutura tecnológica e a ampla dispersão geográfica, buscando promover a ciência aberta e a integração da produção científica regional em iniciativas nacionais e internacionais. Seus objetivos incluem fortalecer a governança regional, criar repositórios digitais, capacitar profissionais e garantir acesso aberto à informação científica, tecnológica e artística da região.

A partir da criação da Rede Norte, o Ibict impulsionou que as demais regiões do Brasil fossem constituindo as suas redes regionais. A segunda rede criada foi a Rede Nordeste de Re却itórios Digitais. Segundo o regimento da Rede, seus objetivos são: reunir as IES e EPT da Região Nordeste que possuem Re却itórios Digitais, para consolidar parcerias com ações em prol do desenvolvimento e da sustentabilidade de seus respectivos Re却itórios; atuar, colaborativamente, desenvolvendo ações práticas, visando empreender melhorias nos Re却itórios Digitais das Entidades parceiras; contribuir para a visibilidade, acessibilidade, usabilidade e difusão dos conteúdos digitais armazenados nos re却itórios digitais das entidades parceiras; promover a construção de fóruns de discussão através das instituições parceiras, garantindo a adoção de padrões comuns e melhores práticas em re却itórios locais que pertencem à Rede (Rede Nordeste [...], 2018).

A terceira rede criada foi a Rede Sudeste de Re却itórios Digitais, com a assinatura da Carta do Rio (Rede Sudeste [...], 2017). Os compromissos constantes na Carta são: criar, com o apoio dos profissionais de

informação, repositórios digitais, institucionais e temáticos, observando os pressupostos do acesso aberto à informação científica; incentivar o depósito da produção científica nos seus repositórios institucionais; buscar o estabelecimento de uma política de ciência aberta e de funcionamento dos repositórios; organizar ações contínuas de capacitação de usuários; proporcionar ajuda mútua na implantação e no gerenciamento de repositórios; buscar novas instituições que possuam perfil para integrar a Rede; criar um canal de comunicação para a Rede; apoiar as ações dos grupos de trabalho criados no âmbito da Rede.

A quarta rede regional, formalizada por meio da Carta de Bagé (Pavão, 2024), foi a Rede Sul de Re却itórios. Os principais compromissos firmados foram: criar, por meio das bibliotecas, repositórios institucionais baseados no princípio do acesso aberto à informação científica; incentivar o depósito da produção científica nos seus repositórios institucionais; editar, revisar e disponibilizar a produção científica depositada via autoarquivamento; estabelecer a política de funcionamento dos repositórios institucionais; organizar ações contínuas de capacitação de usuários; proporcionar ajuda mútua na implantação e no gerenciamento de repositórios; identificar instituições que possuam perfil para integrar a Rede Sul de Re却itórios; criar um portal agregador para os repositórios da rede; divulgar no portal da instituição o link para o re却itório institucional; instituir diretrizes para a criação do consórcio de Re却itórios da Região Sul.

A caçula das redes é a Rede de Re却itórios Institucionais do Centro-Oeste. Para a sua criação foi lançada a Carta de Brasília, onde conclamava às

universidades, institutos de pesquisa, institutos federais e instituições congêneres da região a se comprometerem às seguintes ações: criar, com o apoio dos profissionais de informação, repositórios digitais, institucionais e/ou temáticos, em consonância com os pressupostos do acesso aberto à informação científica; adotar medidas de incentivo ao depósito da produção técnico-científica em seus repositórios institucionais; promover o estabelecimento de política de ciência aberta e de gestão dos repositórios digitais; atuar em colaboração na implantação e no gerenciamento de repositórios; incentivar novas instituições da região, com perfil de atuação em prol do acesso aberto, a integrar a Rede e interagirem entre si; estabelecer e motivar a utilização de canal de comunicação para a Rede; apoiar iniciativas que promovam a integração de seus sistemas de informação em nível regional e nacional.

Todas as coordenações das redes regionais têm tido uma importante atuação junto às instituições de pesquisa. As redes promovem reuniões frequentes com seus membros e organizam palestras, cursos, materiais de apoio e buscam continuamente ampliar o número de instituições participantes. Neste sentido, foi organizado, em 2022, o I Encontro da Rede Brasileira de Re却itórios Digitais: Re却itórios Digitais e Ciência Aberta, evento online promovido em parceria entre o Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), Icict/Fiocruz e as Redes Regionais de Re却itórios e o II Encontro da Rede Brasileira de Re却itórios Digitais: Re却itórios Digitais e Ciência Aberta foi realizado em 2023 e sediado pela Fundação Getúlio Vargas. Nestes eventos, a programação

foi constituída de palestras e mesas de debate voltadas para a comunidade de repositórios digitais de publicações científicas e de dados de pesquisa. Destaque-se também a criação do sítio web da RBRD, onde estão reunidas informações de cada subrede regional (Ibict, 2025).

Com a evolução da ideia do compartilhamento na Ciência e a consequente criação do movimento da Ciência Aberta, as instituições de pesquisa estão cada dia mais compreendendo a necessidade de criarem repositórios de dados de pesquisa. Já não mais se trata de compartilhar somente as publicações científicas e passou a ser necessário compartilhar também os dados oriundos das pesquisas. Neste sentido, foi necessário ampliar o escopo da Rede Brasileira de Re却itórios Institucionais de Publicações Científicas em Acesso Aberto. Foi então debatido pelo Ibict, em conjunto com as coordenações das redes regionais, a mudança de seu nome. No ano de 2022, decidiu-se passar a nomeá-la de Rede Brasileira de Re却itórios Digitais (RBRD) e as subredes também substituíram os termos “repositórios institucionais” por “repositórios digitais”.

Compartilhar os dados de pesquisa foi um novo desafio para a RBRD. Além de continuar disseminando a importância da criação de repositórios institucionais da produção científica, passou a ser necessário propalar a relevância de que as instituições de pesquisa criassem seus repositórios de dados de pesquisa para estarem acordes com as novas práticas da Ciência internacional. Para que a RBRD pudesse atuar, de forma coordenada, na criação de repositórios de dados de pesquisa foi criado, de comum acordo entre as coordenações das subredes

regionais, o Ibict e a Rede Nacional de Ensino e Pesquisa (RNP), o Núcleo de Dados de Pesquisa (NDP).

## NÚCLEO DE DADOS DE PESQUISA (NDP)

A parceria entre o Ibict e a RBRD em prol da consolidação do desenvolvimento e da sustentabilidade dos repositórios digitais de publicações fortaleceu e evidenciou o sentido de redes de colaboração, visto que todo o trabalho exercido pela RBRD não envolve recursos financeiros diretos, mas sim um investimento coletivo em conhecimento e práticas compartilhadas. O compartilhamento colaborativo do conhecimento dentro de uma rede não apenas fortalece as capacidades coletivas, mas também aumenta a inovação e o desenvolvimento sustentável, uma vez que diferentes perspectivas e expertises podem ser combinadas de forma mais eficiente.

É nesse cenário que a RNP, ao observar o alcance da RBRD em território nacional, enxerga a oportunidade de se integrar à rede compartilhando sua experiência em redes de comunicação e infraestrutura tecnológica para fomentar a implementação de repositórios de dados de pesquisa pelas instituições de ensino e pesquisa brasileiras.

A partir da integração da RNP a RBRD, desenvolveu-se então o Núcleo de Dados de Pesquisa (NDP) que passa a funcionar no âmbito da RBRD com o objetivo de impulsionar a criação de um ecossistema robusto e interconectado no desenvolvimento e manutenção de repositórios de dados de pesquisa.

Nesse sentido o NDP tem como essência disseminar as boas práticas para criação de repositórios de dados de pesquisa de acordo com os preceitos da Ciência Aberta que podem ser destacados como:

1. Acesso Aberto e Transparência: os repositórios de dados de pesquisa promovem o acesso aberto aos dados, o que amplia a transparência e a reproduzibilidade dos estudos. Quando os dados são disponibilizados publicamente, outros pesquisadores podem verificar os resultados, reproduzir estudos e construir novos trabalhos com base em dados confiáveis.
2. Avanço Científico: os repositórios de dados de pesquisa facilitam o acesso a dados de pesquisa permitindo a outros cientistas explorar, combinar e reinterpretar informações, impulsionando descobertas inovadoras e promovendo o avanço do conhecimento científico.
3. Eficiência e Redução de Custos: a reutilização de dados de pesquisas anteriores reduz a necessidade de novos levantamentos de dados para as mesmas perguntas, economizando tempo e recursos. Isso é especialmente relevante em áreas onde a coleta de dados é cara e trabalhosa, como ciências ambientais e biomédicas.
4. Colaboração e Interdisciplinaridade: os repositórios de dados abertos facilitam a colaboração entre pesquisadores de diferentes áreas e instituições, estimulando projetos interdisciplinares que podem

abordar problemas complexos de forma mais abrangente.

5. Preservação e Segurança dos Dados: os repositórios garantem a preservação e segurança dos dados a longo prazo. Isso evita a perda de dados valiosos e proporciona um espaço seguro para armazenar grandes volumes de dados de pesquisa, que podem ser perdidos ou danificados com o tempo se forem guardados de forma individual ou local.
6. Impacto e Visibilidade: ao tornar os dados disponíveis, os pesquisadores aumentam o impacto e a visibilidade de seu trabalho. Dados depositados em repositórios confiáveis podem ser citados e reutilizados, contribuindo para o reconhecimento e influência de suas pesquisas. Importante também destacar o aumento da visibilidade das instituições dos pesquisadores, que depositam seus dados em repositórios de dados.
7. Conformidade com Políticas e Financiamentos: muitas agências de fomento já estão exigindo, como pré-requisito para o financiamento, a disponibilização de dados em repositórios abertos como parte de suas políticas de transparência e impacto social. Isso também atende a diretrizes éticas de acessibilidade e compartilhamento público do conhecimento, particularmente em pesquisas financiadas por recursos públicos.

Para disseminar de forma organizada e responsável os preceitos da Ciência Aberta em cada nó da RBRD, o Ibict

assumiu a responsabilidade de treinar e compartilhar conhecimentos relacionados à Ciência da Informação. Esses esforços foram realizados por meio de treinamentos on-line simultâneos. Por sua vez, ao ingressar no projeto, a RNP contribuiu não apenas com sua expertise em tecnologia da informação e infraestrutura de rede, mas também desempenhou um papel fundamental no incentivo ao desenvolvimento de pesquisas, alocando recursos humanos e financeiros.

Todas as atividades foram realizadas no prazo de seis meses, o que só foi possível mediante à significativa contribuição da RNP, que, além de sua expertise em TI e infraestrutura de rede, destinou recursos humanos e financeiros para viabilizar o projeto. Esse investimento foi crucial para o mapeamento das infraestruturas de cada instituição participante da rede.

O apoio financeiro incluiu a concessão de bolsas de pesquisa, reconhecendo a importância do capital humano qualificado para o sucesso do projeto. Foram disponibilizadas bolsas para seis bolsistas, um em cada região do Brasil, garantindo que o processo de mapeamento fosse conduzido de forma descentralizada e abrangente. Cada bolsista foi responsável por estudar as necessidades e peculiaridades das instituições de sua região, identificar práticas locais eficazes e apoiar equipes no aprimoramento de capacidades relacionadas ao armazenamento e gestão de dados de pesquisa.

Além dos seis bolsistas regionais, a RNP financiou um coordenador geral, cuja função foi supervisionar as atividades e promover a articulação entre as diferentes regiões. Esse coordenador desempenhou um papel

estratégico ao integrar as informações coletadas pelos bolsistas, assegurando que as melhores práticas e soluções fossem compartilhadas entre as regiões e alinhadas aos objetivos do projeto. A coordenação também foi essencial para padronizar os processos e garantir a conformidade das soluções com as políticas nacionais e internacionais.

Além do apoio financeiro, a RNP, por meio da Escola Superior de Redes, disponibilizou um curso em ambiente virtual voltado para profissionais de tecnologia da informação. O objetivo do curso foi capacitar os participantes em aspectos tecnológicos relacionados ao software Dataverse, abrangendo tópicos como instalação, configuração, uso de identificadores persistentes, integração com APIs e utilização de painéis para visualização de dados.

As ações conjuntas entre as instituições parceiras resultaram em avanços significativos para o NDP, especialmente no compartilhamento de conhecimento técnico e na conscientização dos pesquisadores sobre os benefícios da Ciência Aberta, promovidos por meio de repositórios de dados de pesquisa em acesso aberto.

## **CONCLUSÕES**

A Ciência Aberta busca aumentar a transparência, a acessibilidade e a colaboração na pesquisa científica, garantindo que os dados, publicações e outros resultados de pesquisa estejam disponíveis para todos, incluindo pesquisadores, estudantes e o público em geral. Nesse contexto, a RBRD se apresenta como uma infraestrutura

essencial, que não apenas contribui para a visibilidade da produção acadêmica nacional, mas também fomenta a troca de informações entre diferentes profissionais e diferentes instituições brasileiras.

Além disso, a RBRD impulsiona a conformidade com práticas globais de compartilhamento de dados e publicações científicas em acesso aberto, incentivando os pesquisadores a disponibilizarem seus resultados de forma mais rápida e acessível, sem as barreiras das publicações tradicionais. Isso facilita a colaboração interdisciplinar, o que é crucial para o avanço da ciência, além de democratizar o acesso ao conhecimento e possibilitar que diferentes setores da sociedade se beneficiem das descobertas científicas.

Por meio dessa iniciativa, o Brasil se posiciona de maneira mais sólida no cenário internacional da pesquisa científica, contribuindo para o fortalecimento da sua ciência e tecnologia.

## REFERÊNCIAS

AMARO, Bianca. A Via Verde do Brasil e a Rede Brasileira de Repositórios Institucionais de Publicações Científicas em Acesso Aberto (RIAA). In: BARBALHO, Célia Regina Simonetti; INOMATA, Danielly Oliveira; GALVES, Jeane Macelino (org.). **A ciência aberta e seus impactos na Região Norte do Brasil**. Manaus: Edua, 2019. cap. 1.

CARTA DE BELÉM. Belém: [s. n.], 2014.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Rede Brasileira de Repositórios Digitais**

(RBRD). Brasília: Ibict, 2025. Disponível em: <http://rbrd.ibict.br>. Acesso em; 11 ago. 2025.

PAVÃO, Caterina Groposo. **Rede Sul de Repositórios Digitais**. Rio de Janeiro. FGV. 2024. (Slides).

REDE NORDESTE DE REPOSITÓRIOS DIGITAIS. **Regimento da Rede Nordeste de Repositórios Digitais**. Recife: [s. n.], 2018.

REDE SUDESTE DE REPOSITÓRIOS DIGITAIS. **Carta do Rio**. Rio de Janeiro: Fiocruz/Icict, 2017.

### Como citar este capítulo

SOUZA, Juliana Araujo Gomes de; AMARO, Bianca. Rede Brasileira de Repositórios Digitais (RBRD). In: AMARO, Bianca; CAMPOS, Phillippe de Freitas; BARCELOS, Janinne. (org.). **Infraestruturas de Ciência e de Acesso Aberto no Brasil: iniciativas do Instituto Brasileiro de Informação em Ciência e Tecnologia**. Brasília, DF: Editora Ibict, 2025. Cap. 15, p. 249-262. DOI: 10.22477/9788570132543.cap15

# LAGUNA: INFRAESTRUTURA INFORMATACIONAL ABERTA E LAGO DE DADOS

Patricia da Silva Neubert

Janinne Barcelos

Fabio Lorensi do Canto

Priscila Machado Borges Sena



## INTRODUÇÃO

O projeto Laguna é uma iniciativa coordenada pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) em colaboração com diversas instituições federais brasileiras. Seu objetivo principal é desenvolver uma Infraestrutura Informacional Aberta (IIA) para consolidar e disponibilizar dados do ecossistema brasileiro de Ciência, Tecnologia e Inovação (CT&I), baseada na arquitetura de um lago de dados (*data lake*) (Carvalho Segundo *et al.*, 2023a e 2023b).

Um *data lake* é um repositório centralizado que armazena dados em seu formato bruto, sem restrições de estrutura ou tipo. Essa abordagem permite a integração de dados de diversas fontes, sejam elas estruturadas (como tabelas de bancos de dados), semiestruturadas (como arquivos JSON ou XML), ou não estruturadas (como e-mails, documentos, imagens, áudio e vídeo) (Alserafi *et al.*, 2016; Khine; Wang, 2018; O’Leary, 2014). Diferente dos sistemas de armazenamento tradicionais que requerem uma estrutura rígida de dados, um lago de dados mantém os dados em seu estado original até que sejam necessários para análises ou processamento, permitindo maior flexibilidade para o uso de tecnologias avançadas, como aprendizado de máquina e inteligência artificial (Derakhshannia *et al.*, 2020).

Entre as vantagens do lago de dados está a capacidade de coletar e atualizar dados diretamente em seu formato nativo nas fontes selecionadas, eliminando etapas preliminares de padronização, o que reduz custos e agiliza o processamento (Fang, 2015). Além disso, o lago de

dados permite que grandes volumes de informações sejam armazenados de maneira econômica, utilizando armazenamento em nuvem escalável, como o *Amazon Web Services (AWS)*, *Google Cloud* ou *Microsoft Azure*, que oferecem suporte a petabytes de dados com baixo custo. Isso facilita a realização de análises extensivas e o desenvolvimento de modelos de aprendizado de máquina, proporcionando insights valiosos para a tomada de decisões estratégicas (Fadnis, 2024).

O Laguna foi inicialmente projetado para atender a duas frentes de serviços estratégicos do Ibict: o BrCris (*Current Research Information System do Brasil*) e a construção de repositórios de dados de pesquisa, uma demanda crescente de instituições de ensino e pesquisa no país. O BrCris (Ibict, 2023) integra informações de repositórios e bases de dados nacionais e internacionais, como a Plataforma Lattes, a Plataforma Sucupira, a Wikidata e a CrossRef, entre outras, criando uma infraestrutura unificada que facilita a gestão e a análise da produção científica brasileira. Paralelamente, a construção de repositórios de dados de pesquisa visa oferecer suporte às instituições que necessitam armazenar, organizar e disseminar dados científicos de maneira eficiente, promovendo a transparência e o acesso aberto conforme os princípios FAIR (*Findable, Accessible, Interoperable, Reusable*).

A fragmentação dos registros de atividades científicas em diferentes sistemas de informação representa um desafio significativo para a obtenção de um panorama claro e abrangente do cenário de Ciência, Tecnologia e Inovação (CT&I). Muitos desses sistemas

operam de forma isolada, com baixa interoperabilidade, o que dificulta a integração e o uso eficiente dos dados para análises robustas e para a formulação de políticas públicas baseadas em evidências. A falta de integração e a existência de barreiras à interoperabilidade tornam o uso desses dados limitado, prejudicando sua utilidade como fontes confiáveis de informação sobre o desenvolvimento científico e tecnológico de um país.

O Laguna visa superar essas limitações adotando uma abordagem de lago de dados, que centraliza dados provenientes de diversas fontes, permitindo seu tratamento, análise e reutilização para múltiplos propósitos (Ibict, 2023). Ao consolidar esses dados em um único repositório, o Laguna facilita a criação de indicadores e relatórios detalhados sobre a produção científica brasileira, essencial para o desenvolvimento e a avaliação de políticas públicas eficazes em CT&I. Além disso, ao promover a convergência de dados e informações, o Laguna se alinha com o movimento de Ciência Aberta, ampliando o acesso e a transparência dos dados científicos. Para tanto, o Laguna baseia-se na criação de uma Infraestrutura Informacional Aberta (IIA).

## **INFRAESTRUTURA INFORMACIONAL ABERTA (IIA) E DO LAGO DE DADOS**

O desenvolvimento da Infraestrutura Informacional Aberta (IIA) e do lago de dados no âmbito do Projeto Laguna conta com uma equipe multidisciplinar responsável por diversos procedimentos técnicos e operacionais. Esses procedimentos envolvem a aplicação de métodos

computacionais avançados para a coleta, tratamento, organização, análise e disponibilização de dados sobre as atividades científicas, tecnológicas e de inovação no Brasil.

Dado que os registros das atividades científicas brasileiras estão fragmentados em vários sistemas públicos e institucionais, como as plataformas Lattes e Sucupira, é necessário um esforço considerável para coletar esses dados. Além desses sistemas, são amplamente utilizados mecanismos de busca, bases de dados e diretórios que agregam registros de produções científicas brasileiras em diversos formatos e tipos. Entre essas fontes, destacam-se a Biblioteca Digital de Teses e Dissertações (BDTD), CrossRef, *Directory of Open Access Journals* (DOAJ), Sistema Regional de Información en Línea para Revistas de América Latina, el Caribe, España y Portugal (Latindex), *Google Scholar Metrics*, *OpenAlex*, *OpenCitations*, OpenAIRE, Portal ISSN e Wikidata.

Para alimentar o lago de dados, é fundamental que essas fontes de informação ofereçam os dados em formatos que permitam exportação, seja por meio de acesso aberto ou por extração de dados a partir de acesso restrito, utilizando VPN institucional ou mediante acordos específicos. Em todos os casos, é essencial agregar valor aos dados integrados, garantir a continuidade na atualização da IIA e promover seu uso nos serviços desenvolvidos a partir dessa integração. Ademais, as diferentes fontes e conjuntos de dados requerem da equipe estratégias variadas para estudo e tratamento, incluindo o desenvolvimento de soluções próprias. Nesse contexto, o projeto tem dado ênfase ao uso de fontes abertas e acessíveis, como o OpenAlex (Neubert et al., 2024).

Uma vez coletados, os dados passam por diversos métodos de tratamento e processamento para maximizar seu valor (Fang, 2015; Giebler *et al.*, 2019). Esse processamento é dividido em várias etapas: (a) seleção e separação dos dados relevantes; (b) transformação e integração para assegurar compatibilidade e conectividade entre diferentes fontes; (c) organização, classificação e indexação para facilitar a recuperação eficiente; e (d) recuperação e visualização, visando disponibilizar os dados de maneira clara e acessível para os usuários finais (Carvalho Segundo *et al.*, 2023b). Tais métodos são descritos no Quadro 1, que destaca as práticas e as técnicas adotadas para garantir a qualidade, a integridade e a interoperabilidade dos dados integrados no sistema.

Quadro 1 - Descrição dos procedimentos para criação da IAA e lago de dados

Procedimento	Descrição
Seleção das fontes dos dados	Identificação, estudo e seleção das fontes sobre atividade em CT&I brasileira, nas quais são priorizados os repositórios e bases de dados de CT&I que atendam total ou parcialmente os princípios FAIR.
Coleta dos dados	Os dados serão coletados por meio de API's públicas ou ferramentas de busca e extração disponíveis em cada uma das fontes selecionadas; se necessário, são utilizadas ferramentas de extração de dados web (web scraping tools) existentes, ou desenvolvidas especificamente para o projeto.

Procedimento	Descrição
Tratamento dos dados	Após a coleta, é realizada a seleção e a separação dos dados, por meio de filtragens e categorizações. É por meio desta etapa que as informações auxiliares (a chamada informação de overhead) são eliminadas, e os arquivos coletados são desmembrados entre os diversos tipos de entidades descritas em seu conteúdo (também chamado de payload).
Integração	Os dados já desmembrados, classificados e categorizados, são adaptados e validados, formando relações com registros de outras fontes. Um registro coletado de uma fonte A tem um atributo comum com o registro coletado da fonte B, podendo ser estabelecida uma vinculação entre ambas, com um determinado grau de confiabilidade. Os demais atributos dos registros podem ser mesclados de forma a resultar em um só registro enriquecido, eliminando-se réplicas. Um esquema de validação pode ser criado de modo a se descartar registros malformados, redundantes, inconsistentes ou ambíguos.
Estruturação	Os registros são organizados, classificados e indexados, de acordo com os seus atributos. As classificações servem de base para construção de interfaces de busca, webservices e dashboards de visualização.
Disponibilização	Com base nos dados tratados, podem ser construídas novas métricas e indicadores,

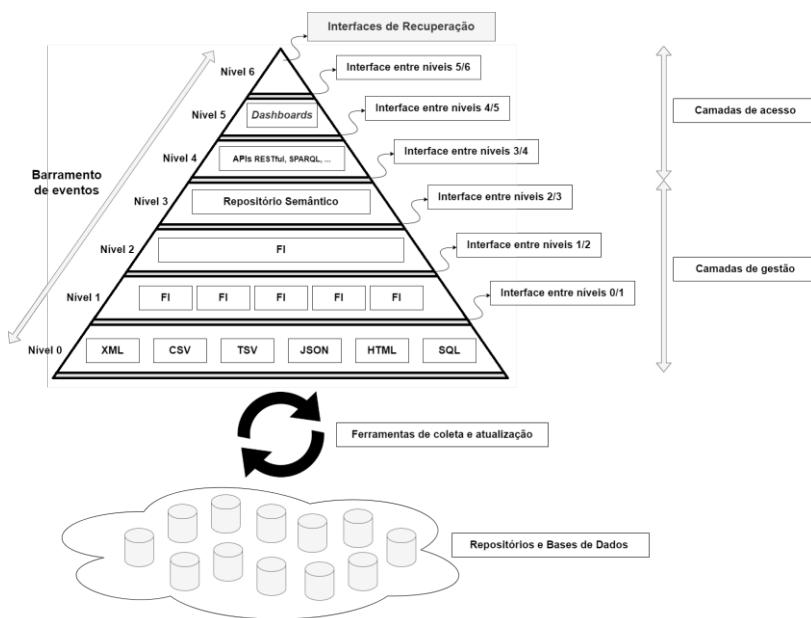
Procedimento	Descrição
	adaptados à produção nacional. Para visualização, serão utilizadas ferramentas de exibição de redes de colaboração, de dados geoespaciais, de séries temporais, de esquemas dinâmicos de tabulação, entre outros.

Fonte: adaptado de Carvalho Segundo (2023).

Após a obtenção dos dados a partir das fontes de informação externas, todos os procedimentos são realizados por meio de interfaces de processamento organizadas em seis níveis, conforme ilustrado na estrutura de processamento do lago de dados na Figura 1.

No Nível 0 da pirâmide, encontram-se os conjuntos de dados na forma em que foram coletados, nos mais variados formatos (como XML, CSV, entre outros). A interface entre os Níveis 0 e 1 transforma esses dados em Formato Intermediário (FI). Desta forma, no Nível 0 há uma normalização estrutural de representação de dados. No entanto, ainda podem existir redundâncias e falta de normalização ou validação quanto ao preenchimento dos campos mapeados. A interface entre os Níveis 1 e 2 realiza a remoção das redundâncias por meio de ações de deduplicação, mantendo a representação via FI. Nesta etapa, também são realizadas ações de normalização e validação dos campos, garantindo a consistência e precisão dos dados à medida que avançam para o próximo nível.

Figura 1 - Estrutura de processamento de dados em níveis do lago de dados



Fonte: Carvalho Segundo (2023).

Já a interface entre os Níveis 2 e 3 exerce a transformação do FI para um padrão de *dados ligados*, onde todos os dados são armazenados no formato de Triplas, na configuração [sujeito, predicado, objeto]. Nesse formato, o sujeito é sempre um identificador, o predicado é uma propriedade e o objeto pode ser um valor literal ou um novo identificador. No nível 3, a representação por meio de Triplas recebe um tratamento fino de representação das propriedades e preenchimento dos valores dos campos,

que são balizados via vocabulários semânticos (Hatch; Brown, 1995).

Por fim, nos Níveis 4, 5 e 6, são desenvolvidas, respectivamente: Interfaces de Programação de Aplicações (APIs) para o consumo dos dados tratados, que permitem o acesso estruturado e seguro aos dados por diferentes aplicações e sistemas; painéis de indicadores (*dashboards*), que fornecem uma interface visual para a visualização e monitoramento dos dados processados, facilitando a análise e a tomada de decisões com base em *insights* e métricas derivadas dos dados; e interfaces de recuperação da informação agregada, projetadas para permitir buscas complexas, consultas avançadas e o cruzamento eficiente de grandes volumes de dados de diferentes fontes.

## ACESSO AO LAGO DE DADOS

Na primeira fase do projeto, o acesso ao lago de dados está limitado aos pesquisadores do Laguna e de outros projetos mantidos pelo Ibict, tais como o BrCris, o OasisBR, a Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), o Diretório das revistas científicas eletrônicas brasileiras (Miguilim) e o indicador persistente *Decentralized Archival Resource Key* (dARK). Estão sendo realizadas diversas atividades de integração entre esses projetos, visando a testar metodologias de processamento e de análise a partir de conjuntos de dados relevantes em uma perspectiva nacional.

Nesse contexto, foi desenvolvido em conjunto com o projeto BrCris um modelo de integração de dados das

Plataformas Lattes e OpenAlex por meio de identificadores persistentes de publicações (DOI) e de autores (ORCID), bem como do registro ISSN de periódicos. Esse modelo possibilitou o enriquecimento dos registros disponibilizados no BrCris, mantendo a origem dos dados do Lattes (fonte principal), mas agregando dados do modelo do OpenAlex (coautoria, fontes das publicações, vinculação institucional) enquanto fonte secundária.

Outra atividade de integração foi realizada com o projeto Miguilim, diretório de revistas científicas brasileiras. Os dados das revistas foram cruzados com diversas bases de dados de periódicos, visando a extração de dados relacionados às publicações nacionais. Foram identificados aspectos como indexação, identificadores de impacto, de produtividade e de citação, classificação em áreas do conhecimento, entre outros.

Além disso, a equipe do Laguna tem participado de reuniões técnicas do Ibict com outras instituições de ensino e pesquisa do país, com a finalidade de estabelecer colaborações nas áreas de pesquisa e avaliação científica, indicadores, repositórios e Ciência Aberta. Essas colaborações podem ser informais, visando a extração de conjuntos de dados do Laguna para utilização em pesquisas. Podem ainda ser formalizadas por meio de protocolos de intenção e acordos de cooperação técnica.

Está prevista também no projeto a disponibilização de conjuntos de dados tratados e enriquecidos no repositório de dados de pesquisa do Ibict, o Alea. Serão selecionados conjuntos de dados relativos à produção brasileira em CT&I, especialmente os conjuntos extraídos de fontes abrangentes tais como OpenAlex, Lattes,

Plataforma Sucupira, OpenAIRE, DOAJ, entre outras. A disponibilização desses conjuntos terá a finalidade de incentivar a utilização dos dados por pesquisadores brasileiros.

## RESULTADOS ESPERADOS

O principal resultado esperado com a execução desse projeto é o desenvolvimento de um amplo repositório de dados abertos relativos à CT&I brasileira, com infraestrutura de lago de dados, armazenamento e processamento distribuídos. Essa infraestrutura permitirá o tratamento e a análise de conjuntos de dados, gerando novos conhecimentos e aplicações estratégicas. O repositório será composto por diversas fontes integradas e padronizadas, possibilitando consultas e análises com ferramentas de visualização, além de exportações por meio de APIs em formatos padronizados e certificados.

As informações serão classificadas e padronizadas, permitindo sua importação por outras ferramentas de análise ou repositórios locais. Dessa forma, o projeto busca consolidar uma base de dados de alta qualidade, integrada e acessível, promovendo o uso eficiente dos dados para o avanço da CT&I no Brasil.

Sob o aspecto tecnológico, espera-se que a equipe adquira *know-how* em tecnologias computacionais de alto desempenho, como inteligência artificial, big data e computação em nuvem. Esse conhecimento será estratégico para aplicação em outros contextos nacionais, ampliando a capacidade técnica e operacional em projetos futuros.

O projeto também visa gerar impactos sociais em médio e longo prazo, principalmente por meio do desenvolvimento de uma infraestrutura informational robusta para CT&I no Brasil. Essa infraestrutura otimizará o aproveitamento dos resultados de pesquisa e subsidiará a formulação de políticas públicas em áreas estratégicas para o desenvolvimento nacional.

Os impactos econômicos esperados incluem a otimização da distribuição de recursos públicos — humanos e materiais — em projetos de CT&I, o aumento da visibilidade e do aproveitamento de resultados de pesquisa, bem como a redução de custos com assinaturas de serviços informacionais oferecidos por empresas comerciais.

Por fim, a partir dos indicadores gerados pelo lago de dados e do tratamento e agregação dos dados, prevê-se o desenvolvimento de produtos e serviços de apoio à pesquisa em CT&I brasileira, como sistemas de recomendação e identificação. Esses produtos e serviços deverão potencializar a inovação e o acesso a dados estratégicos no país.

## REFERÊNCIAS

- ALSERAFI, Ayman *et al.* Towards Information Profiling: Data Lake Content Metadata Management. In: INTERNATIONAL CONFERENCE ON DATA MINING WORKSHOPS (ICDMW), 16., 2016. **Anais** [...]. Barcelona: IEEE, 2016. p. 178-185. DOI: 10.1109/ICDMW.2016.0033.
- CARVALHO SEGUNDO, Washington Luís Ribeiro de *et al.* Fontes de dados sobre periódicos científicos: a proposta do projeto

laguna de dados. *In: CONFERÊNCIA LUSÓFONA DE CIÊNCIA ABERTA*, 14., 2023. **Anais** [...]. Natal, UFRN, 2023c.

CARVALHO SEGUNDO, Washington Luís Ribeiro de *et al.* Tratamento de dados FAIR no Projeto Laguna. *In: CONFERÊNCIA LUSÓFONA DE CIÊNCIA ABERTA*, 14., 2023. **Anais** [...]. Natal, UFRN, 2023a.

CARVALHO SEGUNDO, Washington Luís Ribeiro de. Construindo uma infraestrutura aberta de dados de pesquisa no Brasil. *In: ENCONTRO DA REDE BRASILEIRA DE REPOSITÓRIOS DIGITAIS*, 1., 2022. **Anais** [...]. Rio de Janeiro: Fiocruz/Icict; Ibict, 2022.

DERAKHSHANIA, Marzieh *et al.* Data Lake Governance: towards a systemic and natural ecosystem analogy. **Future Internet**, Basel, v. 12, n. 8, p. 1-16, 27 jul. 2020. DOI: 10.3390/fi12080126.

FADNIS, Bhushan. The Data Lakes: a leap forward future of data warehousing. **International Journal of Innovative Science And Research Technology**, Jaipur, p. 3063-3067, 15 jun. 2024. DOI: 10.38124/ijisrt/IJISRT24MAY2158.

FANG, Huang. Managing data lakes in the big data era: What's a data lake and why has it became popular in data management ecosystem. *In: INTERNATIONAL CONFERENCE ON DATA MINING WORKSHOPS (IEEE)*, 2015. **Anais** [...]. Shenyang, China: IEEE computer society, 2015.

GIEBLER, Corinna *et al.* Leveraging the Data Lake: current state and challenges. **Lecture Notes In Computer Science**, Cham, p. 179-188, 2019. DOI: 10.1007/978-3-030-27520-4\_13.

HATCH, Evelyn; BROWN, Cheryl. **Vocabulary, semantics, and language education**. New York: Cambridge University Press, 1995.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Ecossistema de informação da pesquisa científica brasileira, BrCris.** Brasília, Ibitc, 2023. Disponível em: <https://brcris.ibict.br>. Acesso em: 09 ago. 2024.

KHINE, Pwint Phyu; WANG, Zhao Shun. Data lake: a new ideology in big data era. In: ANNUAL INTERNATIONAL CONFERENCE ON WIRELESS COMMUNICATION AND SENSOR NETWORK (WCSN), 4., 2017. **ITM Web Of Conferences**, Wuhan, v. 17, p. 1-10, 2018. DOI: 10.1051/itmconf/20181703025.

NEUBERT, Patricia *et al.* OpenAlex como fonte de dados para sistemas nacionais de informação científica: a experiência do projeto laguna. In: WORKSHOP DE INFORMAÇÃO, DADOS E TECNOLOGIA - WIDAT, 7., 2024. **Anais** [...]. Porto Velho: Ibitc, 2024. DOI: 10.22477/vii.widat.184.

O'LEARY, Daniel E. Embedding AI and Crowdsourcing in the Big Data Lake. **IEEE Intelligent Systems**, [s. l.], v. 29, n. 5, p. 70-73, set. 2014. 2014.82. DOI: 10.1109/MIS.2014.82.

### Como citar este capítulo

NEUBERT, Patricia da Silva; BARCELOS, Janinne; CANTO, Fabio Lorensi do; SENA, Priscila Machado Borges. Laguna: Infraestrutura Informacional Aberta e lago de dados. In: AMARO, Bianca; CAMPOS, Phillip de Freitas; BARCELOS, Janinne. (org.). **Infraestruturas de Ciência e de Acesso Aberto no Brasil: iniciativas do Instituto Brasileiro de Informação em Ciência e Tecnologia**. Brasília, DF: Editora Ibitc, 2025. Cap. 16, p. 263-277. DOI: 10.22477/9788570132543.cap16



# SOBRE OS ORGANIZADORES



### **Bianca Amaro**

Doutora em Linguística Aplicada pela Universidade Pompeu Fabra (Espanha), com formação em Biblioteconomia, Direito e Letras. Tecnologista no Ministério da Ciência, Tecnologia e Inovação, onde atua como Coordenadora de Governança de Dados.

Curriculum Lattes: <http://lattes.cnpq.br/1445782939373313>.

### **Fhillipe de Freitas Campos**

Mestrando em Ciência da Informação (2023-) e Bacharel em Biblioteconomia pela Universidade de Brasília. Atua como Analista de Gestão da Informação (Bibliotecário) e Pesquisador no Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict).

Curriculum Lattes: <http://lattes.cnpq.br/2076669848354453>.

### **Janinne Barcelos**

Doutora em Ciência da Informação pela Universidade de Brasília (UnB), mestre em Comunicação, Cultura e Cidadania pela Universidade Federal de Goiás (UFG) e bacharel em Comunicação Social - Jornalismo pela Universidade Federal do Tocantins (UFT). Tecnologista no Instituto Brasileiro de Informação para a Ciência e Tecnologia (Ibict).

Curriculum Lattes: <http://lattes.cnpq.br/7729780084365345>.

# SOBRE OS AUTORES



### **Alexandre Faria de Oliveira**

Mestrando em Gestão Estratégica de Organizações, com graduação em Processamento de Dados e pós-graduação em Sistemas Orientados a Objetos. Tecnologista no Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) onde atua como Coordenador-Geral de Tecnologias de Informação e Informática.

Curriculum lattes: <http://lattes.cnpq.br/7305173092285929>.

### **Bernardo Dionízio Vechi**

Bacharel em Biblioteconomia pela Universidade de Brasília (UnB). Atua como bibliotecário no Centro de Ensino Unificado de Brasília (CEUB) e como pesquisador no Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict).

Curriculum lattes: <http://lattes.cnpq.br/9243400892850329>.

### **Bianca Amaro**

Doutora em Linguística Aplicada pela Universidade Pompeu Fabra (Espanha), com formação em Biblioteconomia, Direito e Letras. Tecnologista no Ministério da Ciência, Tecnologia e Inovação, onde atua como Coordenadora de Governança de Dados.

Curriculum lattes: <http://lattes.cnpq.br/1445782939373313>.

### **Blena Estevam dos Santos**

Bacharel em Biblioteconomia pela Universidade de Brasília (UnB). Atua como Pesquisadora e Bibliotecária no Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict).

Curriculum lattes: <http://lattes.cnpq.br/9388711137343097>.

### **Cássio Teixeira de Moraes**

Bacharel em Biblioteconomia e Arquivologia pela Universidade de Brasília (UnB) e tecnólogo em Gestão da Tecnologia da Informação pela Universidade Estácio. Atua como Bibliotecário e Pesquisador no Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict).

Curriculum lattes: <http://lattes.cnpq.br/3368268946691719>.

### **Denise Aparecida Freitas de Andrade**

Bacharel em Biblioteconomia pela Universidade de Brasília (UnB). Bibliotecária e Pesquisadora no Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), com experiência nas áreas de revistas científicas, ciência aberta e acessibilidade digital.

Curriculum lattes: <http://lattes.cnpq.br/6698900487294293>.

### **Diego José Macedo**

Mestre em Ciência da Informação pela Universidade de Brasília (UnB), e graduado em Sistemas de Informação pela Universidade Católica de Brasília (UCB). Tecnologista no Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict).

Curriculum lattes: <http://lattes.cnpq.br/2205539000237712>.

### **Fabiano Couto Corrêa da Silva**

Doutor em Información y Documentación pela Universitat de Barcelona (Espanha), Mestre em Ciência da Informação pela Universidade Federal de Santa Catarina (UFSC) e graduado em Biblioteconomia pela Universidade Federal do Rio Grande do Sul (UFRGS). Professor Adjunto e chefe do Departamento de Ciência da Informação da Universidade Federal do Rio Grande do Sul.

Curriculum Lattes: <http://lattes.cnpq.br/4635807083312321>.

### **Fabio Lorensi do Canto**

Doutor e mestre em Ciência da Informação pela Universidade Federal de Santa Catarina (UFSC). Graduado em Biblioteconomia pela Universidade do Estado de Santa Catarina (UDESC) e em Direito pela Faculdade CESUSC. Atua como Bibliotecário-Documentalista na Biblioteca Central da UFSC.

Curriculum Lattes: <http://lattes.cnpq.br/5914776544385758>.

### **Phillipe de Freitas Campos**

Mestrando em Ciência da Informação (2023-) e Bacharel em Biblioteconomia pela Universidade de Brasília. Atua como Analista de Gestão da Informação (Bibliotecário) e Pesquisador no Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict).

Curriculum Lattes: <http://lattes.cnpq.br/2076669848354453>.

### **Gabriel Silveira Marques**

Mestre em Gestão do Conhecimento pela Universidade Católica de Brasília, com graduação em Biblioteconomia pela Universidade de Brasília e formação técnica em Análise e Desenvolvimento de Sistemas pela Faculdade Estácio. Atua como Bibliotecário e Pesquisador no Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict).

Curriculum Lattes: <http://lattes.cnpq.br/5114965459074231>.

### **Janinne Barcelos**

Doutora em Ciência da Informação pela Universidade de Brasília (UnB), mestre em Comunicação, Cultura e Cidadania pela

Universidade Federal de Goiás (UFG) e bacharel em Comunicação Social - Jornalismo pela Universidade Federal do Tocantins (UFT). Tecnologista no Instituto Brasileiro de Informação para a Ciência e Tecnologia (Ibict).

Curriculum Lattes: <http://lattes.cnpq.br/7729780084365345>.

### **Josir Cardoso Gomes**

Doutor em Ciência da Informação pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) e pela Universidade Federal do Rio de Janeiro (UFRJ). Mestre em Administração pelo IBMEC/RJ e graduado em Processamento de Dados. Pesquisador no Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) e professor na Escola de Comunicação, Mídia e Informação da Fundação Getúlio Vargas (FGV).

Curriculum Lattes: <http://lattes.cnpq.br/4652951094268508>.

### **Juliana Araujo Gomes de Sousa**

Mestranda em Ciência da Informação (2024-) e Bacharel em Biblioteconomia pela Universidade de Brasília (UnB). Atua na implementação e gestão de repositórios digitais científicos, com experiência nos softwares DSpace e Dataverse. Bibliotecária no Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep).

Curriculum Lattes: <http://lattes.cnpq.br/8711223158974435>.

### **Laura Vilela Rodrigues Rezende**

Doutora e mestre em Ciência da Informação pela Universidade de Brasília (UnB), com graduação em Ciência da Computação pela Universidade Católica de Goiás. Professora Associada da Universidade Federal de Goiás (UFG), onde atua na Faculdade de

Informação e Comunicação e no Programa de Pós-Graduação em Comunicação.

Curriculum lattes: <http://lattes.cnpq.br/1612227255633180>.

### **Leticia Guarany Bonetti**

Mestre em Ciência da Informação pela Universidade Federal de São Carlos (UFSCar) e graduada em Biblioteconomia pela Universidade de Brasília (UnB). Tecnologista no Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict).

Curriculum lattes: <http://lattes.cnpq.br/1895977717955732>.

### **Luana Rocha**

Mestre em Ciência da Informação pelo Programa de Pós-Graduação em Ciência da Informação do Instituto Brasileiro de Informação em Ciência e Tecnologia e da Universidade Federal do Rio de Janeiro (PPGCI/Ibict-UFRJ). Atua como pesquisadora no Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), integrando a equipe da plataforma de ciência cidadã Cívis.

Curriculum lattes: <http://lattes.cnpq.br/2058127520105488>.

### **Lucas Rodrigues Costa**

Doutor e Mestre em Informática pela Universidade de Brasília (UnB). Tecnologista no Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict).

Curriculum lattes: <http://lattes.cnpq.br/3133273170328412>.

### **Marcel Garcia de Souza**

Doutorando em Ciência da Informação pelo Programa de Pós-graduação em Ciência da Informação do Instituto Brasileiro de Informação em Ciência e Tecnologia (PPGCI), Mestre em Educação pela Universidade Federal do Rio Grande do Sul

(UFRGS), e graduado em Psicologia pela Universidade Católica de Brasília. Analista em Ciência e Tecnologia no Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), onde atua como Coordenador de Tratamento, Análise e Disseminação da Informação Científica.

Curriculum lattes: <http://lattes.cnpq.br/9517728665816047>.

### **Matheus Dantas**

Mestre em Ciência da Informação e Bacharel em Jornalismo pela Universidade Federal do Rio de Janeiro (2025). Atua como assistente de pesquisa e conteúdo no Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict).

Curriculum lattes: <http://lattes.cnpq.br/3166876006215214>.

### **Miguel Ángel Márdero Arellano**

Doutor e mestre em Ciência da Informação pela Universidade de Brasília, com graduação em Antropologia Social pelo Instituto Nacional de Antropología e Historia (México). Tecnologista no Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), onde coordena a Rede Brasileira de Serviços de Preservação Digital (Cariniana).

Curriculum lattes: <http://lattes.cnpq.br/7000017551659136>.

### **Millena Cordeiro Matos de Lima**

Graduada em Biblioteconomia pela Universidade de Brasília. Atua como Bibliotecária e Pesquisadora no Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict). Trabalha com repositórios digitais e ciência aberta, sendo responsável pela gestão da Biblioteca Digital Brasileira de Teses e Dissertações (BDTD).

Curriculum lattes: <http://lattes.cnpq.br/9404503653256237>.

### **Milton Shintaku**

Doutor e Mestre em Ciência da Informação pela Universidade de Brasília (UnB). Possui graduação em Ciências com habilitação em Matemática pelo Centro Universitário de Brasília e especialização em Análise de Sistemas pela Universidade Católica de Brasília. Tecnologista no Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), onde atua como Coordenador de Tecnologias para Informação.

Curriculum lattes: <http://lattes.cnpq.br/8605833104600600>.

### **Patricia da Silva Neubert**

Doutora e Mestra em Ciência da Informação pela Universidade Federal de Santa Catarina (UFSC), onde também se graduou em Biblioteconomia. Docente do Departamento de Ciência da Informação da UFSC, atuando nos cursos de graduação e no Programa de Pós-Graduação em Ciência da Informação (PGCIN).  
Curriculum lattes: <http://lattes.cnpq.br/8506732139258131>.

### **Priscila Machado Borges Sena**

Doutora e Mestra em Ciência da Informação pela Universidade Federal de Santa Catarina (PGCIN/UFSC). Graduada em Biblioteconomia pela Universidade Federal de Mato Grosso (UFMT). Pesquisadora no Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict).

Curriculum lattes: <http://lattes.cnpq.br/0155235005204514>.

### **Raphael Faria Vilas Boas**

Graduado em Biblioteconomia pela Universidade de Brasília (UnB). Bibliotecário na Rede Sirius da Universidade do Estado do

Rio de Janeiro (UERJ). Atuou como pesquisador no Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), Currículo lattes: <http://lattes.cnpq.br/0250661367752778>.

### **Rebeca dos Santos de Moura**

Mestre em Engenharia de Sistemas Eletrônicos e de Automação e Bacharel em Engenharia de Computação pela Universidade de Brasília. É tecnologista no Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict).

Currículo lattes: <http://lattes.cnpq.br/8677193043257356>.

### **Sandra de Albuquerque Siebra**

Doutora em Ciência da Computação pela Universidade Federal de Pernambuco (UFPE), é professora associada no Departamento de Ciência da Informação da mesma universidade, onde também integra o Programa de Pós-Graduação em Ciência da Informação.

Currículo lattes: <http://lattes.cnpq.br/4923627544089379>.

### **Sarita Albagli**

Doutora em Geografia pela Universidade Federal do Rio de Janeiro (UFRJ). Pesquisadora e docente do Programa de Pós-Graduação em Ciência da Informação do Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict). Coordena o Laboratório de Pesquisa em Ciência Aberta e Inovação Cidadã (CindaLab) e o Instituto Nacional de Ciência Cidadã (INCT).

Currículo lattes: <http://lattes.cnpq.br/8946328562221916>.

### **Tainá Batista de Assis**

Doutora em Ciência da Informação e da Comunicação pela Université de Toulon (França). Mestre em Ciência da Informação e graduada em Biblioteconomia pela Universidade de Brasília (UnB). É Tecnologista do Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict). Atua como presidente do Conselho do Governing Board da Rede ISSN.

Curriculum lattes: <http://lattes.cnpq.br/8018556894529318>.

### **Tatyane Guedes Martins da Silva**

Bacharel em Biblioteconomia pela Universidade de Brasília. Bibliotecária e pesquisadora no Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), atuando em serviços relacionados à ciência aberta e repositórios de dados.

Curriculum lattes: <http://lattes.cnpq.br/7310861285054095>.

### **Thiago Magela Rodrigues Dias**

Doutor em Modelagem Matemática e Computacional pelo Centro Federal de Educação Tecnológica de Minas Gerais (Cefet-MG), é professor nos programas de pós-graduação em Modelagem Matemática e Computacional do Cefet-MG e em Ciência da Informação da Universidade Federal de Santa Catarina (UFSC).

Curriculum lattes: <http://lattes.cnpq.br/4687858846001290>.

### **Washington Luís Ribeiro de Carvalho Segundo**

Doutor em Informática pela Universidade de Brasília (UnB), com período sanduíche no King's College London. Tecnologista no Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict),

onde atua como Coordenador-Geral de Informação Científica e Técnica.

Currículo lattes: <http://lattes.cnpq.br/9453481318889500>.



O livro "Infraestruturas de Ciência e de Acesso Aberto no Brasil: iniciativas do Instituto Brasileiro de Informação em Ciência e Tecnologia" mostra o resultado de décadas de investimento público na construção de estruturas informacionais capazes de apoiar um movimento mundial de abertura da ciência.

Atualmente, comprehende-se esse processo de abertura da ciência a partir de uma perspectiva sistêmica, denominada Ciência Aberta. Diversos motivos contribuíram para o sucesso desse movimento, em especial no Brasil. Dentre esses motivos, é necessário destacar a atuação vanguardista do Ibict, que propôs, planejou e implementou diversas ações que promoveram a adoção da ciência aberta em todo o país.

O livro que você irá ler sintetiza essas ações e as motivações por trás das iniciativas promovidas pelo Ibict ao longo das últimas décadas. Ele é um relato histórico! E também um reforço sobre a importância de se priorizar ações de promoção da ciência aberta no contexto da ciência brasileira.

***Tiago Emmanuel Nunes Braga***  
Diretor do Ibict

978-85-7013-254-3

