

ACESSO ABERTO A DADOS DE PESQUISA NO BRASIL

Repositórios brasileiros
de dados de pesquisa

RELATÓRIO – 2018



RDP BRASIL
Rede de Dados de Pesquisa

Coordenador

Rafael Port da Rocha (Universidade Federal do Rio Grande do Sul)

Coordenadora adjunta

Sônia Elisa Caregnato (Universidade Federal do Rio Grande do Sul)

Pesquisadores

Universidade Federal do Rio Grande do Sul

Caterina Groposo Pavão
Paula Caroline Schifino Jardim Passos
Rene Faustino Gabriel Junior
Samile Andréa de Souza Vanz

Universidade Federal de Rio Grande

Eduardo Nunes Borges
Luís Alberto Barbosa Azambuja

Bolsista

Universidade Federal do Rio Grande do Sul

Iván Andrés Fornos Angues

Contato

e-mail: dadosdepesquisa@rnp.br
site: <https://dadosdepesquisa.rnp.br>

Rua Ramiro Barcelos, 2705 - Campus Saúde
Sala: 106 - Anexo 1
Brasil - Porto Alegre - RS - CEP 90.035-007
Telefone: +55(51)3308.5942



Este relatório é licenciado sobre a licença CC BY - Creative Commons Attribution 4.0 International License.

Como citar:

PAVÃO, Caterina Groposo; VANZ, Samile Andrea de Souza; PASSOS, Paula Caroline Schifino Jardim; CAREGNATO, Sônia Elisa; AZAMBUJA, Luís Alberto Barbosa; BORGES, Nunes Borges; GABRIEL JUNIOR, Rene Faustino; ROCHA, Rafael Port da. Acesso aberto a dados de pesquisa no Brasil: repositórios brasileiros de dados de pesquisa: relatório 2018. Disponível em: <<http://hdl.handle.net/10183/185138>>.

P337 PAVÃO, Caterina Groposo

Acesso aberto a dados de pesquisa no Brasil : repositórios brasileiros de dados de pesquisa : relatório 2018 / Caterina Groposo Pavão et al. - Porto Alegre, RS : UFRGS, 2018.
19 p.

Relatório do projeto RDP Brasil: Rede de Dados de Pesquisa Brasileira, – Universidade Federal do Rio Grande do Sul; Universidade Federal do Rio Grande.

1. Dados abertos de pesquisa. 2. Repositório de dados de pesquisa. I. Título



Resumo

Relatório da primeira etapa do projeto Rede de Dados de Pesquisa Brasileira (RDP Brasil), o qual objetivou tanto identificar repositórios brasileiros que disponibilizam dados de pesquisa, quanto levantar características relacionadas à abrangência, à temática, aos padrões de metadados e aos softwares utilizados. Neste levantamento, foi utilizado como fonte de informação principal o Re3data. Ademais, fez-se uso de buscas na internet e de indicações feitas pelos respondentes da *survey*, realizada na etapa dois do projeto RDP Brasil. Foram identificados sete repositórios registrados no Re3data e oito repositórios de universidades e de instituições de pesquisa disponíveis na internet. As indicações de repositórios utilizados pelos respondentes do *survey* não acrescentaram novidades em relação ao conjunto identificado inicialmente. Dentre os 15 repositórios identificados, quatro são de abrangência internacional (envolvendo várias instituições), sete são de abrangência nacional e cinco são de abrangência multi-institucional. Nesses repositórios, predominam as temáticas voltadas às Ciências Sociais Aplicadas, às Ciências Agrárias, às Ciências Exatas e da Terra, às Ciências Biológicas e à Engenharia. Os padrões de metadados encontrados foram o Dublin Core, o DataCite Metadata Schema, a ISO 19115 e o DarwinCore. O identificador persistente DOI é empregado em quatro repositórios e o padrão de interoperabilidade OAI-PMH é citado em cinco repositórios. O Dspace é utilizado em dois repositórios, assim como o Metacat. O Dataverse é utilizado em um repositório. Em três repositórios, são utilizados softwares desenvolvidos localmente. Já em outros sete, esta informação não é oferecida. Grande parte dos dados depositados são dos tipos texto, bases de dados, imagens, planilhas e gráficos. As políticas dos repositórios foram encontradas em apenas oito dos 15 analisados. Conclui-se não somente que iniciativas institucionais de compartilhamento de dados abertos de pesquisa no Brasil ainda são poucas, mas também que esforços públicos ainda são necessários para consolidar tais iniciativas.

Palavras-chave:

Dados abertos de pesquisa. Repositório de dados de pesquisa.

Sumário

1 Introdução.....	5
2 Procedimentos metodológicos	5
3 Análise dos dados.....	6
4 Conclusão	18

1 Introdução

Atendendo ao objetivo 1 do projeto Rede de Dados de Pesquisa Brasileira (RDP Brasil), realizou-se um levantamento tanto para identificar repositórios de Acesso Aberto a Dados de Pesquisa (AADP) em universidades e em instituições de pesquisa brasileiras, quanto para identificar outras experiências de soluções e de práticas de armazenamento e de disseminação de dados de pesquisa. Foram relacionadas na pesquisa também aquelas iniciativas que não se reconhecem como repositórios de dados, mas que apresentam características semelhantes aos repositórios relacionados no que diz respeito ao seu conteúdo. A pesquisa se justifica pela ausência de estudos que consolidem essas práticas em âmbito nacional. A próxima seção apresenta os procedimentos metodológicos utilizados para o levantamento e, em seguida, mostra a apresentação dos repositórios e das iniciativas brasileiras identificadas.

2 Procedimentos metodológicos

A pesquisa de identificação das práticas de AADP em instituições brasileiras ocorreu em três etapas, entre março e junho de 2018. Na primeira delas, os dados foram coletados a partir do diretório internacional Research Data Repositories Information (Re3data). Para busca, foi utilizada a opção *Browser by country* do menu principal. Em seguida, foi selecionada a opção Brasil.

Na segunda etapa, foram buscados repositórios de dados em publicações na web sobre o tema e em sites de instituições de pesquisa reconhecidas nacionalmente pelo seu trabalho com dados de pesquisa. Essa etapa permitiu verificar a utilização de nomenclatura alternativa para o termo repositório como, por exemplo, banco ou base de dados. Nos casos em que isso ocorreu e em que foram encontradas características e informações similares às dos repositórios de dados propriamente ditos – como no Instituto Nacional de Meteorologia (INMET), o serviço foi incorporado.

Na última etapa, foi utilizada uma das questões da *survey* realizada principalmente para contemplar o objetivo 2 do projeto RDP Brasil. A questão deixava espaço livre para o respondente especificar o nome dos repositórios de dados por ele utilizados. Nessa etapa, nenhum repositório diferente dos já identificados nas etapas 1 e 2 foi mencionado.

O levantamento foi realizado entre os meses de maio e junho de 2018. Nos repositórios de dados encontrados, procurou-se identificar características quanto aos seguintes itens:

- a) âmbito (nacional ou cooperação internacional);
- b) áreas temáticas de abrangência;
- c) número de instituições envolvidas na coleta e disponibilização de dados;
- d) políticas do repositório (institucional, direitos autorais, licenças de uso);

- e) formato dos dados;
- f) interoperabilidade e padrões de metadados;
- g) software utilizado, aplicações e interfaces;
- h) existência de informações quanto às orientações para elaboração do plano de gestão de dados e ao ciclo de vida da pesquisa.

Em alguns casos, não foi possível identificar todas as características elencadas acima, pois as páginas desses repositórios não fornecem tais informações, mesmo quando o registro como usuário da plataforma é realizado. Em outros casos, encontram-se informações em notícias ou em documentos veiculados pela instituição, porém não foi possível acessar o repositório como, por exemplo, a Base Tuiuiu - Tecnologia de Gestão e Compartilhamento de Dados Primários e Secundários de Projetos, da EMBRAPA. Por esse motivo, alguns campos das tabelas apresentadas neste relatório são preenchidos como “informação não disponível” e a descrição de algumas características não é uniforme para todos os repositórios.

O Re3data é um registro global de repositórios de dados de pesquisa que abrange diferentes áreas do conhecimento. Os repositórios registrados devem prestar-se para o armazenamento e para o acesso permanente a conjuntos de dados de pesquisa, a órgãos financiadores, a editores e a instituições acadêmicas. O Re3data promove uma cultura de compartilhamento, maior acesso e visibilidade dos dados da pesquisa. Foi lançado em 2012 e é financiado pela German Research Foundation (DFG).

Para se registrar no Re3data, um repositório de dados de pesquisa deve obedecer a certos requisitos, a saber: ser administrado por uma entidade legal como, por exemplo, uma biblioteca, uma universidade etc.; esclarecer as condições de acesso aos dados e ao repositório, bem como aos termos de uso; e ter foco em dados de pesquisa.

O Re3data considera um repositório de dados de pesquisa como uma das partes de uma infraestrutura de informação sustentável que fornece armazenamento a longo prazo e acesso a dados de pesquisa, considerados a base para uma publicação científica. Por sua vez, dados de pesquisa significam objetos de informação gerados por projetos de pesquisa, por exemplo, experimentos, medições, pesquisas ou entrevistas.

3 Análise dos dados

No Re3data foram identificados, entre maio e junho de 2018, os seguintes repositórios, mantidos ou nos quais há participação de instituições brasileiras: IBICT Dataverse Network, BDEP - Banco de Dados de Exploração e Produção, GLOBE - Global Collaboration Engine, International Ocean Discovery Program (IODP), PPBio Data Repository, WorldClim - Global Climate Data e a Base de Dados Científicos da Universidade Federal do Paraná. A seguir serão relatadas as informações coletadas sobre estes repositórios.

■ **IBICT Dataverse Network**

O IBICT Dataverse Network disponibiliza um repositório de dados de pesquisa que cuida da preservação a longo prazo e de boas práticas de arquivamento, para que os pesquisadores possam compartilhar, manter o controle e obter o reconhecimento de seus dados. O repositório suporta o compartilhamento de dados de pesquisa com citação de dados persistentes, permitindo que eles sejam reproduzidos.

URL	https://repositoriopesquisas.ibict.br
Âmbito	Nacional
Instituições participantes	Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC), Brasil; Dataverse Network Project, Estados Unidos.
Temática	Ciências Sociais e Humanas, Ciências da Vida e Ciências Naturais
Formatos dos dados	Formatos de dados científicos e estatísticos, texto, dados brutos e bases de dados
Políticas	As políticas, normas da comunidade, termos e contratos de uso seguem o padrão da Harvard Dataverse Policies https://dataverse.org/best-practices/harvard-dataverse-policies
Metadados	Data Documentation Initiative (DDI) http://www.dcc.ac.uk/resources/metadata-standards/ddi-data-documentation-initiative e o DataCite Metadata Schema for the Publication and Citation of Research Data http://schema.datacite.org/meta/kernel-2.1/doc/DataCite-MetadataKernel_v2.1.pdf , que por sua vez utiliza o padrão DublinCore
Identificador persistente	Informação não disponível
Software	Dataverse
Interoperabilidade	OAI-PMH

■ **Banco de Dados de Exploração e Produção (BDEP)**

O BDEP possui em seu acervo mais de seis petabytes de dados técnicos gerados nas atividades de prospecção petrolífera em todo território brasileiro. Desde 2000, o BDEP recebe, avalia, armazena e disponibiliza esses dados aos solicitantes, seguindo as regras de sigilosidade estipuladas pelas normas regulatórias vigentes. A aquisição e o gerenciamento de dados dessa coleção garantem ao Brasil o domínio sobre o potencial de conhecimento gerado em hidrocarbonetos.

URL	http://www.anp.gov.br/exploracao-e-producao-de-oleo-e-gas/dados-tecnicos/acesso-aos-dados-tecnicos
Âmbito	Nacional
Instituições participantes	Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP) e Serviço Geológico do Brasil.
Temática	Geociências, Geografia, Mineralogia, Geofísica, Geodésia, Geoquímica, Ciências Naturais
Formatos dos dados	Imagens, padrões e normas, gráficos estruturados, documentos textuais, dados estatísticos e científicos. Os dados técnicos devem ser formatados conforme o tipo: dados sísmicos, coordenadas, geoquímicos etc. As especificações encontram-se em http://www.anp.gov.br/exploracao-e-producao-de-oleo-e-gas/dados-tecnicos/padroes-tecnicos
Políticas	Sobre confidencialidade dos dados Resolução ANP nº 11/2011, sobre a solicitação de dados Resolução ANP nº 1/2015,
Metadados	Data Documentation Initiative (DDI) http://www.dcc.ac.uk/resources/metadata-standards/ddi-data-documentation-initiative
Identificador persistente	Informação não disponível
Software	Informação não disponível
Interoperabilidade	Informação não disponível

■ **Global Collaboration Engine (GLOBE)**

O Globe é um ambiente colaborativo on-line que permite aos pesquisadores compartilhar, comparar e integrar estudos locais e regionais com dados globais para avaliar a relevância global de seu trabalho sobre mudança de uso e cobertura de solo. Usa paradigmas da computação social colaborativa e modelos estatísticos avançados para agregar e analisar estudos de caso locais e regionais e permitir a síntese e a integração de informações globalmente relevantes.

URL	http://globe.umbc.edu
Âmbito	Internacional
Instituições participantes	Instituto Nacional de Pesquisas Espaciais (INPE), Brasil; International Network of Research on Coupled Human and Natural Systems (ICHANS-NET), University of Maryland, The Other Firm e National Science Foundation (NSF), Estados Unidos.

Temática	Biologia, Geociências, Oceanografia, Ciências da Vida, Ciências Naturais
Formatos dos dados	Dados baseados em rede, gráficos estruturados, Dados brutos As especificações para descrição encontram-se em http://globe.umbc.edu/tutorials/tutorial-data-sheets/
Políticas	Informação não disponível
Metadados	Disponível em: http://www.dcc.ac.uk/resources/metadata-standards/abcd-access-biological-collection-data
Identificador persistente	DOI
Software	Informação não disponível
Interoperabilidade	Informação não disponível

■ **International Ocean Discovery Program (IODP)**

O IODP é um programa de colaboração internacional de pesquisa marinha que explora a história e a dinâmica da Terra usando plataformas de pesquisa oceânica para recuperar dados registrados em sedimentos e rochas do fundo do mar e monitorar ambientes de subsolos. Reúne 26 nações, os cientistas são selecionados para realizar expedições de pesquisa nos oceanos do mundo.

URL	http://www.iodp.org
Âmbito	Internacional
Instituições participantes	Integrated Ocean Discovery Program (IODP), National Science Foundation (NSF), Estados Unidos; IODP-Koordinationsbüro an der Bundesanstalt für Geowissenschaften und Rohstoffe BGR e MARUM - Zentrum für Marine Umweltwissenschaften der Universität Bremen, Alemanha; Japan Agency for Marine-Earth Science and Technology, Center for Deep Earth Exploration e Japan's Ministry of Education, Culture, Sports, Science and Technology, Japão; India Ministry of Earth Science, Índia; Australian-New Zealand IODP Consortium, Austrália; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brasil; Korea Institute of Geology, Mining and Materials, Coreia do Sul; The European Consortium for Ocean Research Drilling, Comunidade Europeia; The People's Republic of China Ministry of Science and Technology, China.

Temática	Geociências (incluindo Geografia), Oceanografia, Geofísica e Geodesia, Geologia e Paleontologia, Geoquímica, Mineralogia e Cristalografia, Ciências Naturais e Ciências Atmosféricas
Formatos dos dados	Informação não disponível
Políticas	Distribuição de Amostra do ODP, Distribuição de Dados e Política de Publicações (Aprovado em março de 1999; revisado em 1º de junho de 2001, 15 de abril de 2002, 19 de agosto de 2002). Este documento descreve a política e os procedimentos para distribuição de amostras e dados do Programa de Perfuração Oceânica (ODP) e do Projeto Deep Sea Drilling (DSDP) para pesquisadores cientistas, curadores e educadores. http://www-odp.tamu.edu/publications/policy/policy.pdf Ao receber amostras, os pesquisadores concordam implicitamente em cumprir a Política de Amostra, Dados e Obrigações do IODP. Um Conselho Consultivo de Curadoria faz as decisões finais sobre a distribuição de amostras de IODP. https://www.iodp.org/resources/access-data-and-samples
Metadados	http://www.dcc.ac.uk/resources/metadata-standards/iso-19115
Identificador persistente	Informação não disponível
Software	Informação não disponível
Interoperabilidade	OAI-PMH

■ **PPBio Data Repository**

O PPBio Data Repository foi criado em 2004 com o objetivo de fomentar estudos sobre biodiversidade no Brasil, descentralizando a produção científica de centros acadêmicos já desenvolvidos, integrando atividades de pesquisa e disseminando resultados em diversas finalidades, incluindo gestão ambiental e educação. O PPBio contribui com seus dados para a rede DataONE como um nó membro: <https://search.dataone.org/#profile/PPBIO>.

URL	https://ppbiodata.inpa.gov.br/metacatui/
Âmbito	Internacional
Instituições participantes	Centro de Estudos Integrados da Biodiversidade Amazônica (INCT-CENBAM), Brasil; Data Observation Network for Earth (DataONE), Estados Unidos e Long Term Ecological Research Network (LTER), internacional.

Temática	Ciências da Vida, Biologia Vegetal, Ecologia e Análise de Ecossistemas, Ecologia Animal, Biodiversidade e Pesquisa de Ecossistemas, Zootecnia, Botânica e Zoologia
Formatos dos dados	Imagens, dados audiovisuais, texto estruturado, texto simples, formatos de dados científicos e estatísticos, dados brutos
Políticas	Política de dados https://ppbio.inpa.gov.br/politicadados Política de uso CC https://creativecommons.org/licenses/by/3.0/ Licença de upload de dados https://ppbiodata.inpa.gov.br/metacatui/#about
Metadados	DCC http://www.dcc.ac.uk/resources/metadata-standards/eml-ecological-metadata-language
Identificador persistente	DOI e Uniform Resource Names (URN)
Software	Metacat https://knb.ecoinformatics.org/knb/docs/
Interoperabilidade	OAI-PMH

■ **Global Climate Data (WorldClim)**

O WorldClim é um conjunto de dados climáticos globais (grades de clima) com uma resolução espacial de cerca de um quilômetro quadrado. Os dados podem ser usados para mapeamento e modelagem espacial em um SIG ou com outros programas de computador.

URL	http://worldclim.org
Âmbito	Internacional
Instituições participantes	Centro de Referência em Informação Ambiental (CRIA), Brasil; University of California, Berkeley, Museum of Vertebrate Zoology, Estados Unidos; Ashoka Trust for Research in Ecology and the Environment, Índia e National Herbarium of the Netherlands, Holanda e Rainforest CRS, Austrália.
Temática	Ciências Naturais Climatologia, Ecologia, Oceanografia, Geociências (incluindo Geografia)
Formatos dos dados	Gráficos estruturados, dados científicos e estatísticos e dados brutos

Políticas	Direitos autorais e licenciamento http://osc.universityof-california.edu/scholarly-publishing/copyright-licensing/ Política de uso http://worldclim.org/current
Metadados	Informação não disponível
Identificador persistente	Informação não disponível
Software	Drupal https://www.drupal.org/
Interoperabilidade	Informação não disponível

■ **Base de Dados Científicos da Universidade Federal do Paraná (BDC/UFPR)**

A BDC/UFPR tem como objetivo reunir os dados científicos utilizados nas pesquisas que foram publicadas pela comunidade da UFPR em teses, dissertações, artigos de periódicos e outros materiais bibliográficos.

URL	https://bdc.c3sl.ufpr.br/
Âmbito	Nacional
Instituição participante	Universidade Federal do Paraná
Temática	Ciências Sociais e Humanas, Ciências Biológicas, Ciências Naturais e Engenharia
Formatos dos dados	Bancos de dados, formatos de dados científicos e estatísticos, códigos-fonte, gráficos estruturados, aplicativos de software, imagens, audiovisuais e texto
Políticas	Licenças de uso CC https://creativecommons.org/licenses/by/4.0/
Metadados	Dublin Core http://www.dcc.ac.uk/resources/metadata-standards/dublin-core
Identificador persistente	DOI e Handle System
Software	DSpace
Interoperabilidade	OAI-PMH

Encontram-se listados abaixo os repositórios que não estão registrados no Re3data, mas que foram identificados em universidades e/ou institutos de pesquisa do país. A pesquisa foi realizada entre os meses de maio e junho de 2018.

■ **Base Tuiuiu**

A Base Tuiuiu é uma tecnologia desenvolvida em cooperação técnica entre a Embrapa Pantanal (Corumbá, MS) e o Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso do Sul (IFMS) para suporte ao acesso, manipulação e compartilhamento de dados. Com interface amigável, essa tecnologia permite inserir e acessar dados contínuos (no tempo) ou discretos (pontuais), de sensores automáticos ou coletados no campo.

URL	http://tuiuui.cpap.embrapa.br (Em 09/07/2018 o link não estava disponível)
Âmbito	Nacional
Instituições participantes	Empresa Brasileira de Pesquisa Agropecuária (Embrapa), Embrapa Pantanal, Brasil.
Temática	Ar, solo, água, planta, animal, microbiologia (podem ser criadas novas variáveis mediante solicitação do usuário para os responsáveis, coordenadores de projetos)
Formatos dos dados	O sistema disponibiliza quatro opções de formatos para download dos dados: XLS, PDF, CSV e XML
Políticas	Política do repositório, direitos autorais e plano de gestão de dados disponíveis em: https://ainfo.cnptia.embrapa.br/digital/bitstream/item/119152/1/DOC129.pdf
Metadados	Disponível em: https://ainfo.cnptia.embrapa.br/digital/bitstream/item/119152/1/DOC129.pdf
Identificador persistente	Informação não disponível
Software	Desenvolvimento local
Interoperabilidade	(Parece ser interoperável com repositório Alice)

Fonte: <https://ainfo.cnptia.embrapa.br/digital/bitstream/item/119152/1/DOC129.pdf>. Acesso em: 11 de jul de 2018

■ **Consórcio de Informações Sociais (CIS),**

O CIS é um sistema de compartilhamento de bancos de dados que surgiu em 2003 a partir de um projeto do Departamento de Sociologia da Universidade de São Paulo (USP), direcionado para a comunidade acadêmica de Ciências Sociais. Desde 2011, o CIS conta com apoio do CNPq.

URL	http://www.cis.org.br
Âmbito	Nacional
Instituições participantes	Universidade de São Paulo (USP) e Associação Nacional de Pós-graduação e Pesquisa em Ciências Sociais (ANPOCS), Brasil.
Temática	Ciências sociais

Formatos dos dados	Bancos de dados, livros de códigos e artigos disponíveis no acervo estão em diversos formatos, entre eles o SPSS, o MS-Excel, o MS-Access, o Win-SIS, o PDF e outros. Para envio de dados não serão aceitos arquivos em .exe, .com, .bat, .scr, .eml. Arquivos compactados podem ser enviados em 7-Zip, WinZip, RAR, ARJ.
Políticas	Informação não disponível
Metadados	Informação não disponível
Identificador persistente	Informação não disponível
Software	Desenvolvimento local em ASP
Interoperabilidade	Informação não disponível

Fonte: <http://www.cis.org.br>. Acesso em: 11 de jul de 2018

■ **Instituto Nacional de Meteorologia (INMET)**

O INMET possui um banco de dados para apoiar as atividades de ensino e pesquisa. Abriga dados meteorológicos diários em forma digital, de séries históricas das várias estações meteorológicas convencionais da rede de estações do INMET com milhões de informações, referentes às medições diárias, de acordo com as normas técnicas internacionais da Organização Meteorológica Mundial. Estão acessíveis os dados diários a partir de 1961 das estações para as quais se disponha, em forma digital, de pelo menos 80% dos dados que foram registrados naquele período. Os dados históricos referentes a períodos anteriores a 1961 ainda não estão em forma digital. As variáveis atmosféricas disponibilizadas para consultas são: precipitação ocorrida nas últimas 24 horas; temperatura do bulbo seco; temperatura do bulbo úmido; temperatura máxima; temperatura mínima; umidade relativa do ar; pressão atmosférica ao nível da estação; insolação; direção e velocidade do vento.

URL	http://www.inmet.gov.br/portal/
Âmbito	Nacional
Instituição participante	Ministério da Agricultura Pecuária e Abastecimento (MAPA), Brasil.
Temática	Meteorologia, hidrologia, recursos hídricos, saúde pública, meio ambiente etc.
Formatos dos dados	Texto estruturado, gráficos dinâmicos
Políticas	Informação não disponível
Metadados	Informação não disponível
Identificador persistente	Informação não disponível

Software	Informação não disponível
Interoperabilidade	Informação não disponível

Fonte: <http://www.inmet.gov.br/projetos/rede/pesquisa/>, acesso em: 11 de jul de 2018

■ **Instituto Brasileiro de Geografia e Estatística (IBGE)**

O banco de dados do IBGE objetiva retratar o Brasil com informações necessárias ao conhecimento de sua realidade e ao exercício da cidadania. Constitui-se no principal provedor de dados e informações do País, que atendem às necessidades dos mais diversos segmentos da sociedade civil, bem como dos órgãos das esferas governamentais federal, estadual e municipal. Oferece uma visão completa e atual do País, através do desempenho de suas principais funções: produção e análise de informações estatísticas, coordenação e consolidação das informações estatísticas, produção e análise de informações geográficas, coordenação e consolidação das informações geográficas, estruturação e implantação de um sistema de informações ambientais, documentação e disseminação de informações e coordenação dos sistemas estatístico e cartográfico nacionais.

URL	https://ww2.ibge.gov.br/
Âmbito	Nacional
Instituição participante	Instituto Brasileiro de Geografia e Estatística (IBGE), Brasil.
Temática	Estatística e Geociências
Formatos dos dados	Plano de Dados Abertos (https://www.ibge.gov.br/np_download/novoportal/documentos_institucionais/Plano_de_Dados_Abertos_IBGE_2018_2019.pdf) Formatos abertos (JSON, ODS e XML) e recursos, como links (documentação da pesquisa e página para a tabela agregada) em formato HTML.
Políticas	Informação não disponível
Metadados	https://metadados.ibge.gov.br/consulta
Identificador persistente	Informação não disponível
Software	Informação não disponível
Interoperabilidade	Os parâmetros estabelecidos na e-PING - arquitetura de interoperabilidade do governo eletrônico - e os vocabulários e ontologias de Governo Eletrônico e-VoG e e-MAG - Modelo de Acessibilidade de Governo Eletrônico (instituído pela Portaria nº 03, de 07 de maio de 2007)

Fonte: <https://www.ibge.gov.br/index.php>, acesso em: 11 de jul de 2018

■ **Sistema Maxwell**

O Sistema Maxwell, da PUC-Rio, é um repositório institucional que abriga e disponibiliza parte da produção acadêmica da universidade. Este sistema é um produto do Laboratório de Automação de Museus, Bibliotecas Digitais e Arquivos do Departamento de Engenharia Elétrica (LAMBDA).

URL	https://www.maxwell.vrac.puc-rio.br/research_data.php?tipBusca=inicio
Âmbito	Nacional
Instituição participante	Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Brasil.
Temática	Engenharia e Física
Formatos dos dados	Informação não disponível
Políticas	Política de direito autoral, Lei 9610/98
Metadados	Dublin Core
Identificador persistente	DOI
Software	Desenvolvimento local em PHP
Interoperabilidade	OAI-PMH

■ **Repositório de dados PELD**

O repositório de dados PELD, do Programa de Pesquisa Ecológica de Longa Duração, foi criado para atender à necessidade dos sítios PELD de armazenamento qualificado dos dados produzidos pelo Programa. A estrutura do repositório foi desenvolvida no âmbito do Sistema de Informação sobre a Biodiversidade e os Ecossistemas Brasileiros (SiBBr).

URL	http://peld.inpa.gov.br (em manutenção, 05 jul. 2018)
Âmbito	Nacional
Instituições participantes	Ministério de Ciência, Tecnologia, Inovações e Comunicações (MCTIC), Programa das Nações Unidas para o Meio Ambiente (PNUMA), Rede Nacional de Pesquisa (RNP) e Laboratório Nacional de Computação Científica (LNCC), Brasil; Fundo Global para o Meio Ambiente (GEF), internacional.
Temática	Ecologia de Ecossistemas
Formatos dos dados	Informação não disponível
Políticas	Informação não disponível
Metadados	Informação não disponível
Identificador persistente	Informação não disponível

Software	Metacat https://knb.ecoinformatics.org/knb/docs/
Interoperabilidade	Informação não disponível

Fonte: <http://memoria.cnpq.br/repositorio-peld>. Acesso em: 11 de jul.2018

■ **Projeto speciesLink**

O Projeto speciesLink é um sistema distribuído de informação que integra dados primários de coleções científicas. A rede speciesLink tem por princípio promover o acesso livre e aberto aos dados, informações e ferramentas disponíveis a qualquer indivíduo ou grupo.

URL	http://www.splink.org.br/index?setlang=pt
Âmbito	Nacional
Instituições participantes	Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CAPES), Financiadora de Estudos e Projetos (FINEP), Rede Nacional de Ensino e Pesquisa (RNP), Centro de Referência em Informação Ambiental (CRIA), Brasil; Global Biodiversity Information Facility (GBIF), Dinamarca; JRS Foundation, Estados Unidos.
Temática	Ciências biológicas
Formatos dos dados	Informação não disponível
Políticas	Os participantes da rede manifestaram o seu compromisso de tornarem seus dados disponíveis para promover a pesquisa e o desenvolvimento científico no país e no exterior e para estimular o uso público da informação científica. Para a utilização dos dados disponíveis através dessa rede o usuário indica estar ciente que: a qualidade e completude dos dados não podem ser garantidas, o uso e aplicação dos dados correrão por conta e risco do usuário; os usuários deverão observar e respeitar as restrições indicadas por cada provedor; os usuários deverão reconhecer publicamente o uso dos dados e imagens, sempre citando a fonte, rede speciesLink e provedores dos dados.
Metadados	DarwinCore
Identificador persistente	Informação não disponível
Software	SpLinker http://splink.cria.org.br/splinker
Interoperabilidade	Informação não disponível

■ **IPAgriDados**

O IPAgriDados foi implementado para facilitar a identificação das bases de dados existentes na unidade e a catalogação de um conjunto mínimo de metadados e de suas restrições de uso, ampliar o uso, reuso e integração das bases de dados mantidas na unidade, possibilitando o aumento do impacto positivo desse tipo de resultado, além permitir a evidenciação e valorização dos esforços e investimentos para a obtenção e manutenção desses ativos da empresa.

URL	https://www.dados.cnptia.embrapa.br
Âmbito	Nacional
Instituição participante	Empresa Brasileira de Pesquisa Agropecuária (Embrapa), Brasil.
Temática	Ciências Agrárias
Formatos dos dados	O catálogo permite o depósito de dados na forma de qualquer tipo de arquivo digital, incluindo: textos, planilhas, imagens, arquivos de dados geográficos (vetoriais e matriciais). Alternativamente, é possível inserir uma URL1 para acesso à base de dados.
Políticas	Informação não disponível
Metadados	http://wiki.dados.gov.br/Padroes-de-metadados.ashx
Identificador persistente	Informação não disponível
Software	DSpace
Interoperabilidade	Informação não disponível

Fonte: <https://www.dados.cnptia.embrapa.br/jspui/static/doc/manual-ipagridadoscat-v1-1.pdf>.

Acesso em: 11 de jul. de 2018

4 Conclusão

O levantamento realizado permitiu identificar 15 repositórios brasileiros e, mediante informações registradas no Re3data ou na interface do repositório, buscou identificar características relativas a seu âmbito; a suas áreas temáticas de abrangência; a seu número de instituições envolvidas na coleta e na disponibilização de dados; a suas políticas relativas ao depósito dos dados, dos direitos autorais, das licenças de uso, entre outros; a seu formato para depósito dos dados; a sua interoperabilidade com outros sistemas e seus padrões de metadados para descrição dos dados; a seu software utilizado e, por fim, as suas orientações para elaboração do plano de gestão de dados, quando existentes.

Dos 15 repositórios identificados, quatro são de abrangência internacional, ou seja, envolvem várias instituições, com pelo menos uma do Brasil, e 11 são de abrangência nacional, sendo que dois desses são do tipo multi-institucional. Predominam as temáticas voltadas às Ciências Sociais Aplicadas, às Ciências Agrárias, às Ciências Exatas e da Terra, às Ciências Biológicas e à Engenharia.

No item que se refere a metadados, verificou-se que a maioria dos repositórios não especifica o padrão utilizado. Dentre aqueles que informam o padrão, encontramos o Dublin Core, o Data Cite Metadata Schema, a ISO 19115 e o DarwinCore. Da mesma forma, o identificador persistente é citado em apenas quatro repositórios, que utilizam o DOI, e o padrão de interoperabilidade OAI-PMH é citado em cinco repositórios. Nos demais repositórios, não há menção ao padrão utilizado.

O Dspace é utilizado em dois repositórios, assim como o Metacat. O Dataverse é utilizado em um repositório. Em três repositórios, são utilizados softwares desenvolvidos localmente. Já em outros sete, esta informação não é oferecida. Os formatos mais utilizados nos dados depositados são os seguintes: texto, dados brutos, bases de dados, imagens, planilhas e gráficos. Dois repositórios não possuem informação a respeito. Em relação às políticas, oito repositórios dão acesso às suas políticas, nos demais, não foi possível localizar esta informação.

O levantamento revela a existência de poucas iniciativas nacionais relacionadas ao compartilhamento dos dados de pesquisa brasileiros. As universidades – responsáveis pela maior parte da pesquisa desenvolvida no Brasil – não têm protagonismo no que diz respeito aos repositórios de dados de pesquisa: apenas UFPR e PUC Rio foram identificadas no levantamento. Outro detalhe interessante é a existência de alguns repositórios identificados nos websites institucionais, mas não cadastrados no Re3data, o que revela a incipiência das iniciativas.

Dada a atualidade do tema e a solicitação da gestão de dados por parte de agências de financiamento e de revistas científicas, é urgente tratar esse assunto de forma mais sistemática no Brasil.

