Disponibilização de dados de pesquisa em repositórios universitários da América Latina

JAQUELINE INSAURRIAGA SILVEIRA

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

jaqueline.silveira@ufrgs.br

SAMILE ANDRÉA DE SOUZA VANZ

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

samilevanz@terra.com.br

CATERINA MARTA GROPOSO PAVÃO

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

caterina@cpd.ufrgs.br

EIXO TEMÁTICO

Datos abiertos

RESUMO

O trabalho analisa a disponibilidade de dados de pesquisa em repositórios universitários da América Latina. Metodologicamente, o estudo abrange uma pesquisa documental onde identifica no re3data repositórios registrados como sendo para dados de pesquisa, para posteriormente selecionar os que são institucionais de universidades. Descreve os repositórios quanto ao armazenamento e disponibilidade de dados de pesquisa, restrições de acesso e licenças de uso. Constata-se que o re3data não representa uma fonte de pesquisa fidedigna para recuperação de repositórios de dados, uma vez que as informações registradas são autodeclarações das instituições. Dessa forma, a pesquisa sofreu interferência devido a escolha do Diretório como ferramenta de coleta de dados. Considera-se que o objetivo do presente estudo não foi totalmente alcançado, uma vez que dos 20 repositórios analisados, somente seis são exclusivos para dados de pesquisa.

PALAVRAS-CHAVE

Dados de pesquisa; disponibilização de dados; repositório institucional; América Latina.

ABSTRACT

The paper analyzes the availability of research data in university repositories in Latin America. Methodologically, the study covers documentary research where it identifies re3data repositories registered as being for research data, to later select those that are institutional of universities. It describes the repositories in terms of storage and availability of research data, access restrictions and use licenses. It appears that re3data does not represent a reliable research source for retrieving data repositories, since the information recorded are self-declarations by the institutions. Thus, the research suffered interference due to the choice of the Directory as a data collection tool. It is considered that the objective of the present study was not fully achieved, since of the 20 repositories analyzed, only six are exclusive for research data.

KEYWORDS

Research data; data availability; institutional repository; Latin America.

1 Introdução

A Ciência Aberta é entendida como um movimento que envolve práticas e iniciativas de abertura da ciência em diferentes âmbitos. Dentre estas, destaca-se para o presente estudo, a abertura dos dados de pesquisa com suas implicações para a ciência, oportunizando sua reprodutibilidade e desenvolvimento de novos estudos; para os pesquisadores, potencializando a troca de conhecimento, otimizando tempo e recursos de pesquisa; e para a sociedade, por meio de descobertas que impulsionam políticas públicas.

Ao possibilitar o compartilhamento e reuso, a abertura de dados de pesquisa proporciona a reprodutibilidade da ciência, além de fornecer agilidade aos processos científicos e otimização de recursos financeiros para novos estudos. A partir desse entendimento, atores envolvidos nos contextos científico, financeiro e político da ciência (pesquisadores, comunidade científica,

instituições, governos e agências de fomento), empenham esforços para que, pesquisas financiadas com recursos públicos estejam acessíveis ao público, em formato digital sem restrição ou com restrição mínima (Torino, Roa-Martinez & Vidotti, 2020).

No entanto, o compartilhamento de dados de pesquisa requer iniciativas que garantam a sua disponibilização e integridade, de forma permanente ao longo do seu ciclo de vida. Ao encontro dessa exigência, tem-se, os repositórios digitais e especialmente os Repositórios de Dados de Pesquisa (RDP) como sistemas de informação adequados para o depósito, acesso e uso dessa tipologia documental, uma vez que dispõem das necessárias condições técnicas e tecnológicas de armazenamento e preservação.

A região Latino Americana, segundo Amaro (2021), é tida como modelo de produção científica em acesso aberto, tendo em vista a forte atuação das editoras universitárias. Especificamente no Brasil, o lançamento da SciELO em 1998 e a tradução do software livre Open Journal Systems pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) em 2003 foram os principais motivos pelo qual o acesso aberto se difundiu rapidamente (Vanz & Silva Filho, 2019). Diante desse contexto de reconhecimento da grande adesão da América Latina às publicações em acesso aberto, pode-se inferir que, em relação a abertura e disponibilização de dados de pesquisa, a Região repete o mesmo protagonismo.

Partindo dessa inferência, este estudo objetiva conhecer as iniciativas de abertura de dados de pesquisa dos países da América Latina, por meio da disponibilização em repositórios universitários de dados de pesquisa, a existência ou não de restrições de acesso, e a atribuição de licenças de uso. As próximas seções apresentam os procedimentos metodológicos, uma apresentação descritiva dos repositórios encontrados seguida pela discussão destes resultados e conclusão.

2 Procedimentos metodológicos

Para a identificação dos repositórios no âmbito da América Latina, realizou-se uma busca no **Registry of Research Data Repositories (re3data)** no período de 12 a 22 de fevereiro de 2023, utilizando-se do filtro "País". Foram recuperados 59 repositórios sediados em seis países, com o seguinte quantitativo de repositórios: Argentina (oito); Brasil (18); Chile (dois); Colômbia (doze); México (dezessete); Panamá (dois). Dentre estes, realizou-se uma triagem selecionando somente os repositórios institucionais de universidades, resultando neste quantitativo: Argentina (quatro); Brasil (quatro); Chile (um); Colômbia (nove) e México (dois). O repositório argentino Humadoc Repositorio[1] não foi analisado por não atender aos critérios metodológicos estabelecidos para o presente estudo, ou seja, analisar repositórios institucionais de universidades, não fazendo parte do estudo aqueles vinculados a órgãos específicos das instituições.

Cada um dos repositórios registrados no re3data foi visitado para coleta de dados, caracterizando a pesquisa documental. A obtenção das informações sobre acesso, restrições e licenças de uso se deu por meio de análise dos conjuntos de dados (datasets) e seus metadados recuperados por meio de estratégias de busca e mecanismos de visualização disponibilizadas pelos repositórios. Quando não se obteve sucesso na recuperação de datasets, buscou-se informações nos sites e políticas de gestão desses repositórios. Importante ressaltar que o re3data não coleta informações sobre os repositórios, sendo estas registradas por seus gestores. Dessa forma é possível a existência de outras iniciativas que por não estarem registradas no Diretório, não fazem parte deste estudo.

[1] Repositório da Faculdad de Humanidades de la Universidad Nacional de Mar del Plata.

Nos casos em que a coleta de dados e informações diretamente no repositório não atendeu a demanda da pesquisa, foi realizado contato com o setor responsável pelo repositório por e-mail ou Whatsapp.

3 Apresentação e análise dos resultados

A seguir, apresenta-se os repositórios identificados no re3data, por país e suas caraterísticas principais.

3.1 Argentina

Identificaram-se quatro repositórios universitários, apresentados a seguir.

3.1.1 Producción Académica UCC

O Producción Académica UCC é o repositório institucional da Universidad Católica de Córdoba (UCC), fundada em 1956. Gerenciado pelo Sistema de Bibliotecas da Universidade, reúne a produção acadêmica e científica gerada por professores e pesquisadores da Universidade, objetivando sua preservação a longo prazo e divulgação, com a garantia de acesso livre e aberto (UCC, 2021). Disponibiliza em seu website as políticas de: conteúdo; de acesso; de depósito; de metadados; de preservação; de serviços; e de confidencialidade e privacidade.

Embora sua Política de conteúdo contemple os dados de pesquisa nos tipos documentais, ao realizar-se uma busca em sua plataforma, recuperou-se somente um dataset, o qual é disponibilizado em acesso aberto, sem restrições e nem exigência de identificação de usuário, correspondendo a um conjunto de documentos XML que dão respaldo a um caderno de trabalho (workingpaper - WP).

Não foi possível localizar diretrizes específicas para dados de pesquisa, sendo assim, infere-se que estes obedeçam às orientações destinadas aos outros tipos de documentos. O acesso ao Repositório é aberto e todos os

conteúdos depositados estão sob a proteção de licenças Creative Commons (CC). Os autores devem conceder ao Repositório uma licença de distribuição não exclusiva sob a qual será feito o arquivamento, a divulgação e a preservação dos documentos. Para os conteúdos, os autores são encorajados a atribuir uma Licença Atribuição - Não Comercial - Sem Trabalhos Derivados (CC BY-NC-ND 4.0).

Visando confirmar a recuperação de somente um dataset nas buscas realizadas no Repositório, uma vez que sua interface de pesquisa apresentou instabilidade recorrente em diferentes momentos da pesquisa, buscou-se contato via e-mail com o Sistema de Bibliotecas da UCC. Obteve-se, por meio da Diretora do Sistema, a confirmação do armazenamento de somente um dataset e, da não existência de uma política específica para dados de pesquisa, justificado pelo fato da Universidade estar dando os primeiros passos ao encontro da gestão de dados de pesquisa na Instituição.

3.1.2 Repositorio de Datos Académicos UNR (RDA-UNR)

O RDA-UNR, lançado no ano de 2022, além de ser o repositório de dados de pesquisa da Universidad Nacional de Rosario (UNR), é o primeiro repositório aberto de dados de pesquisa de uma universidade argentina. Também é um repositório de autopublicação, onde os usuários carregam, gerenciam e publicam seus dados, contando com a ajuda de curadores e guias de utilização. Com o objetivo de dar visibilidade, promover a reutilização e garantir a preservação dos dados, o Repositório além de receber dados de pesquisa gerenciados por professores e alunos pesquisadores da Universidade, permite também que pesquisadores de outras instituições que publicam artigos em periódicos do Portal de Revistas UNR depositem os dados correspondentes às suas investigações (Universidad Nacional de Rosario, 2022).

Disponível no website do Repositório, sua política de gestão abrange as diretrizes de: Acordo de uso geral do RDA-UNR; Contrato de liberação de dados de pesquisa; Contrato de uso de dados; Licenças de uso de dados;

Compromisso de citação de dados; Dados aceitos no repositório; Quantidade e tamanho dos dados; Dados sensíveis e confidenciais; e de Preservação.

A gestão e uso dos dados de pesquisa no Repositório incorrem na aceitação das condições do Contrato de liberação de dados de pesquisa, pelos autores de dados e, do Contrato de uso de dados, pelos usuários que acessam, baixam e utilizam os dados de pesquisa armazenados no RDA-UNR. Aos dados de pesquisa é concedida uma Licença Atribuição 4.0 Internacional (CC By BY4.0), permitindo compartilhamento e adaptação dos dados, com sua devida citação de autoria.O RDA-UNR organiza-se em 17Dataverses vinculados a faculdades e órgãos da UNR.

3.1.3 Repositorio Institucional de la Universidad Católica de Salta (UCASAL)

O repositório institucional da Universidad Católica de Salta (UCASAL) tem como objetivo a promoção do acesso aberto à produção científico-tecnológica gerada pela Instituição. O acesso ao Repositório se dá pela página do Sistema de Bibliotecas da Universidade, e armazena em formato digital a documentação acadêmica e científica gerada e publicada pela instituição, disponibilizando seu conteúdo em acesso aberto.

O Repositório está organizado em quatro grandes coleções: Colección Tesis y Trabajos Finales; Colección Investigación; Colección Histórica Institucional e Colección Archivo Fotográfico. De acordo com a Resolução Reitoral 1017/17, os dados de pesquisa devem ser armazenados na Colección Investigación, porém estes, não estão contemplados na descrição de conteúdo dessa coleção. Em contato com o Sistema de Bibliotecas da Universidade obteve-se a confirmação de que no presente momento não há conjunto de dados processados no Repositório.

Para o licenciamento de conteúdos, a Resolução Reitoral 1017/17 promove o uso das Creative Commons como forma de manter a autoria da produção científica da Universidade, facilitando seu uso para fins acadêmicos e de investigação (Universidad Católica de Salta [UCASAL], 2017).

3.1.4 Servicio de Difusión de la Creación Intelectual (SEDICI)

O Servicio de Difusión de la Creación Intelectual (SEDICI) é o repositório institucional da Universidad Nacional de La Plata (UNLP), tendo por objetivo indexar, preservar e disponibilizar em acesso aberto toda a produção intelectual produzida na Universidade, dentre estes, datasets. O Repositório divide-se em quatro grandes coleções: Publicações acadêmicas e científicas; Coleções patrimoniais; Dados de pesquisa; e Documentos institucionais.

A coleção Dados de pesquisa contém 85 datasets, e verificou-se que os dados estão lincados com as publicações relacionadas. A disponibilização dos dados se dá em acesso aberto sem exigência de identificação e registro de usuário. Em relação às licenças de uso, salvo indicação contrária, são atribuídas Creative Commons, permitindo compartilhar, copiar e redistribuir o material em qualquer meio ou formato para qualquer finalidade, mesmo comercialmente, desde que seja dado o crédito apropriado e este não seja modificado (CC BY-ND).

3.2 Brasil

A pesquisa no re3data recuperou 18 repositórios, sendo quatro destes universitários, a saber:

3.2.1 Base de Dados Científicos da Universidade Federal do Paraná – BDC/UFPR

A Base de Dados Científicos da Universidade Federal do Paraná (BDC/UFPR) tem por objetivo reunir os dados científicos utilizados em pesquisas publicadas pela comunidade da Universidade Federal do Paraná (UFPR) em teses, dissertações, artigos de periódicos e outros materiais bibliográficos. Lançada em setembro de 2017, a BDC/UFPR é gerenciada pelo Sistema de Bibliotecas da Universidade em parceria com o Centro de Computação Científica e Software Livre (C3SL) da Instituição, e representa a primeira base de dados de pesquisa de uma universidade pública brasileira (Estevão, 2018).

A coleção de dados conta com 65 documentos publicados entre os anos de 2017 e 2022. Realizou-se o acesso aleatório à coleção de dados de pesquisa, constatando-se sua disponibilização em acesso aberto com a atribuição de licenças Creative Commons CC0 1.0 Universal (CC0 1.0), ou, uma Atribuição-NãoComercial-SemDerivados 3.0 Brasil (CC BY-NC-ND 3.0 BR).

3.2.2 Open Research Data@PUC-Rio

O Open Research Data@PUC-Rio é um agregador para facilitar o acesso à dados de pesquisa, entre outros conteúdos digitais no Repositório Maxwell. O Maxwell é o repositório institucional da Pontifícia Universidade Católica do Rio de Janeiro (PUC-RIO), desenvolvido e mantido pelo Laboratório de Automação de Museus, Bibliotecas Digitais e Arquivos (LAMBDA), laboratório da Vice-reitoria Acadêmica (VRAc) da Universidade (Pontifícia Universidade Católica do Rio de Janeiro [PUC-Rio], 2022). Com o objetivo de compartilhar dados que foram desenvolvidos e/ou modificados no âmbito da Universidade para que estes sejam utilizados em estudos na Instituição, o Open Research Data@PUC disponibilizou em 2015 seu primeiro conjunto de dados de pesquisa. Para submissão de conteúdo é necessário que, pelo menos, um dos autores pertença à comunidade PUC-Rio e que os dados sejam disponibilizados em acesso aberto. Todos os datasetes submetidos devem ser licenciados com uma licença Creative Commons, que segundo informações disponibilizadas no site, é uma atribuição (CC BY-NC-SA 4.0).

3.2.3 Repositório Institucional UNESP

O Repositório Institucional UNESP pertence à Universidade Estadual Paulista, que sob a responsabilidade do Grupo Gestor da Política do Repositório Institucional UNESP (GRI-UNESP) tem por objetivo armazenar, preservar, divulgar e disponibilizar em acesso aberto à produção intelectual dos pesquisadores e alunos da Universidade (Universidade Estadual Paulista [UNESP], 2022). Dentre os documentos armazenados disponibiliza também dados de pesquisa e data paper. Em pesquisa realizada utilizando-se o filtro

'Tipo de documento', recuperou-se 83 datasets, publicados entre os anos de 2008 e 2023. Dentre estes, há dados em acesso aberto e dados com restrição de acesso. Em relação à atribuição de licenças de uso, não foi possível localizar nenhuma informação no site do Repositório e nem nos metadados dos datasets.

3.2.4 Repositório de Dados de Pesquisa da Unicamp (REDU)

O Repositório de Dados de Pesquisa da Unicamp (REDU) é o repositório institucional de dados abertos de pesquisa da Universidade Estadual de Campinas (Unicamp). Criado em 2020 e coordenado por um comitê científico e técnico composto por bibliotecários, profissionais de TI e pesquisadores representantes dos usuários, o REDU armazena, preserva, e disponibiliza os dados de pesquisa produzidos na Universidade, permitindo seu compartilhamento e reutilização. De acordo com a política de Ciência Aberta da Unicamp, o Repositório apresenta cinco dataverses (Artes e Humanas; Exatas; Ciências Aplicadas; Biomédicas; Tecnológica), onde encontram-se depositados 138 datasets e 933 arquivos publicados entre os anos de 2018 e 2023. A disponibilização dos dados de pesquisa se dá por prerrogativa em acesso aberto com uma atribuição de Domínio Público (CC0), contudo pode ser necessário o embargo de dados, mas com disponibilização dos seus metadados em acesso aberto.

3.3 Chile

No re3data recuperaram-se dois repositórios chilenos, contudo somente um deles é universitário, o **Repositorio de Datos de Investigación de la Universidad de Chile**. Este, é um repositório institucional que objetiva a preservação, divulgação e disponibilização de dados de pesquisa gerados por sua comunidade de pesquisadores e acadêmicos. É composto por 19 dataverses, 50 datasets e 334 arquivos de dados publicados entre os anos de 2019 e 2023. A disponibilização dos arquivos se dá em acesso aberto com indicação de

atribuição de Domínio Público (CC0) nos Termos de uso dos dados, e também em acesso restrito, exigindo a identificação de usuário.

3.4 Colômbia

A pesquisa no re3data retornou 12 repositórios, sendo somente nove pertencentes à universidades, os quais são apresentados na sequência.

3.4.1 Repositório Institucional Intellectum (RII)

O Intellectum é o repositório institucional da Universidad de La Sabana, tendo sua criação aprovada pela Comissão de Assuntos Gerais do Conselho Superior na Lei no 1370, de 6 de dezembro de 2012. Criado para gerenciar, preservar e divulgar a produção intelectual, científica, cultural e histórica de sua comunidade universitária, tem sua administração a cargo da Direção da Biblioteca Octavio Arizmendi Posada. O RII é composto por seis comunidades, sendo elas: Trabajos de Grado; Recursos Educativos Digitales Abiertos (REDA); Memoria Institucional; Publicaciones Académicas; Archivo Histórico Cipriano Rodriguez Santa María; Artículos de Investigación Indexados; e, Otros Documentos.

A disponibilização dos documentos se dá em acesso aberto e acesso restrito e para licenciamento de uso são atribuídas licenças Creative Commons adaptadas para o país. Por padrão, a indicação é de atribuição (CC BY-NC-SA 4.0), salvo quando orientação diferente. Não foi possível verificar a disponibilização de dados de pesquisas no RII, uma vez que não se obteve sucesso na recuperação destes no Repositório. Da mesma forma, em suas políticas não existem diretrizes para essa tipologia documental. Ainda buscando confirmar a existência ou não de dados de pesquisa no RII, entrou-se em contato via WhatsApp com a Biblioteca gestora do Repositório. Obteve-se a informação de que não há dados de pesquisa armazenados no RII, visto estarem ainda em processo de criação da coleção Datos de Investigación.

3.4.2 Repositorio Universidad Autónoma de Bucaramanga (UNAB)

O Repositório UNAB é o repositório institucional da Universidad Autónoma de Bucaramanga (UNAB), para fins de armazenamento, preservação e divulgação da produção intelectual, científica, cultural e histórica da sua comunidade universitária. É composto por seis comunidades de documentos, sendo elas: Investigación; Normatividad institucional; Publicaciones académicas; Recursos multimedia; Tesis y disertaciones; e, Trabajos de grado. Não se obteve resultado ao se efetuar uma busca por dados de pesquisa na plataforma, utilizando-se o filtro 'Tipologia documentária'. Da mesma forma, não foram localizadas as políticas de gestão do Repositório.

3.4.3 Repositorio Institucional Universidad Distrital Francisco José de Caldas (RIUD)

O RIUD é o repositório institucional da Universidad Distrital Francisco José de Caldas destinado a armazenar sua produção intelectual. Conta com cinco comunidades de documentos: Documentos RIUD; Producción Editorial Universidad Distrital Francisco José de Caldas; Recursos de Ensenãnza; Tesis y Disertaciones Académicas; Trabajos de grado. Não foi possível localizar dados de pesquisa por meio de busca aleatória em suas comunidades documentais.

Para as licenças de uso de conteúdo, sua política explicita que os titulares de direitos autorais cedem à Universidade uma licença não exclusiva, limitada e gratuita. A licença é aplicada por 10 anos prorrogáveis indefinidamente, de acordo com os direitos de autor, o qual pode rescindi-la mediante aviso prévio antes de findar o prazo inicial ou de renovação (Universidad Distrital Francisco José de Caldas, 2022). A disponibilização dos documentos no RIUD se dá em acesso aberto e acesso restrito e, para licenciamento de uso desses, segundo sua política, atribui-se licenças Creative Commons, não ficando claro contudo, se a atribuição das licenças fica a cargo do depositante ou da Instituição.

3.4.4 Repositorio Institucional de la Universidad Santo Tomás

O Repositorio Institucional de la Universidad Santo Tomás é destinado ao armazenamento, preservação e divulgação de toda produção acadêmica e administrativa da Universidade. Gerencia 10 comunidades de documentos digitais: Banco de objetos virtuales de aprendizaje; Centro de Recursos para el Aprendizaje, la Investigación y la Docencia; Colegios de la Orden de Predicadores; Documentos CRAI-USTA; Gestión Institucional; Grupos de Investigación; Libros; Publicaciones seriadas; Responsabilidad Social Universitaria; Tesis, Trabajos de Grado y Disertaciones Académicas. Realizou-se uma busca aleatória em suas comunidades para identificação de datasets, contudo não se obteve êxito, e dessa forma não se pode verificar como ocorre a disponibilização destes. O que se pode observar em relação aos documentos acessados aleatoriamente dentro das Comunidades é que estes são disponibilizados em acesso aberto com atribuição das licenças Creative Commons.

3.4.5 Repositorio UNAL

O Repositorio UNAL é o repositorio institucional da Universidad Nacional de Colombia (UNAL), gerenciado por sua **Biblioteca Digital**, com a finalidade de conservar e difundir a produção acadêmica e científica da Universidade. O Repositório disponibiliza acesso aberto aos documentos armazenados e disposto nas seguintes comunidades: Tesis y Disertaciones; Libros y capítulos de libro; Documentos e tutoriales para publicar; Proyectos temáticos; Trabajos de grado; General; Revistas UNAL. Dentro da tipologia documentária armazenada nas comunidades, não foi possível identificar datasets. Em relação ao licenciamento dos conteúdos depositados, os autores concedem ao Repositório uma licença não exclusiva, limitada e gratuita. A disponibilização dos documentos se dá em acesso aberto com atribuição de licenças Creative Commons.

3.4.6 Repositorio Institucional Universidad de Antioquia

O Repositorio Institucional Universidad de Antioquia é coordenado pelo Sistema de Bibliotecas da Universidade. Sua política de gestão expressa a importância do acesso aberto aos conteúdos digitais como um meio de democratização do conhecimento. Armazena artigos de revistas científicas e acadêmicas; livros; capítulos de livros; trabalhos de graduação, mestrado e doutorado; trabalhos de eventos; workingpapers; relatórios de pesquisa; material audiovisual; e documentos históricos e patrimoniais. Estes conteúdos são aceitos somente após passarem por uma revisão por pares.

O Repositório conta com nove comunidades, sendo elas: Biblioteca Virtual de Antioquia; Editoriales; Información institucional; Investigaciones; Patrimonio bibliográfico y documental; Producción docente; Producción estudiantil; Trabajos de posgrados; e Trabajos de pregrado. Analisou-se a descrição de conteúdo das Comunidades e em nenhuma delas consta datasets, contudo, a Universidade possui uma Política Institucional de Acesso Aberto, onde discorre que o Plano de Desenvolvimento Institucional (2017-2026) se compromete com a política e estratégias da Ciência Aberta e que a Universidade procurará incorporar gradualmente em seu repositório, dados de pesquisas, recursos educativos, software e outros.

3.4.7 Repositorio Institucional Universidad Autónoma de Manizales (REPOUAM)

O REPOUAM, repositório institucional da Universidad Autónoma de Manizales (UAM), tem o compromisso com o acesso aberto a todos os tipos de conteúdo digital, incluindo dados de pesquisa. Estruturado em comunidades e coleções, o REPOUAM contempla as comunidades: Institucional (composta por documentos institucionais); Investigación; Recursos educativos; Sello editorial (composta por produções editorais da Universidade; Tesis y disertaciones; e Trabajos de grado. Em uma busca aleatória no Repositório, foi possível localizar na Comunidade Investigación, um dataset relacionado a um estudo

sobre contratura muscular. A disponibilização deste dataset se dá em acesso aberto licenciado por uma atribuição (CC BY-ND 4.0).

3.4.8 Repositorio Digital UTB

O Repositorio Digital UTB é o repositório institucional da Universidad Tecnológica de Bolívar (UTB). Disponibiliza todos os documentos produzidos no âmbito da Universidade por meio das comunidades: Fototecas; Investigaciones; Producción académica; e Sello editorial. Não se obteve sucesso na recuperação de dados de pesquisa em sua plataforma e na tentativa de localizar suas políticas para esclarecer sobre as diretrizes e orientações sobre a tipologia de conteúdo aceita para depósito. Em relação à disponibilização de acesso aos documentos em geral armazenados, pode-se constatar que essa se dá em acesso aberto e acesso restrito e, os documentos são licenciados com atribuições Creative Commons.

3.4.9 Repositorio de datos de investigación de la Universidad del Rosario

Lançado em 2019, o Repositório tem por objetivo preservar e disponibilizar acesso aos dados de pesquisa resultante de projetos da Universidade, visando seu compartilhamento e reuso. Suas normativas seguem a Política Institucional de Gestão de Dados de Pesquisa, orientada pela Portaria Reitoral nº 1.576, de março de 2019. A Portaria orienta sobre a tipologia de dados reconhecidos para depósito no Repositório, permitindo a inclusão de novos tipos conforme necessidade futura. Da mesma forma, discorre sobre dados anônimos, anonimizados, pessoais, públicos, semiprivados, privados, sensíveis, e dados pessoais de crianças e adolescentes. No momento da coleta de dados para este estudo, o Repositório contava com 118 registros de dados de pesquisa dispostos em diversas coleções. Muitos datasets estão licenciados com Creative Commons Domínio Público (CC0) e Atribuição-NãoComercial 4.0 Internacional (CC BY-NC 4.0), outros têm acesso restrito para consulta, sendo disponibilizados somente seus metadados.

3.5 México

Recuperaram-se no re3data 17 repositórios mexicanos, contudo somente dois universitários.

3.5.1 Portal de Datos Abiertos da Universidad Nacional Autónoma de México (UNAM)

O Portal de Datos Abiertos da UNAM funciona como um ponto de acesso online para busca de acervos universitários autorizados para uso, reutilização e redistribuição gratuita desde que respeitados os Termos de Uso Gratuito de Dados Abertos da UNAM. Contém dados, objetos digitais e camadas geoespaciais de coleções biológicas, trabalhos artísticos, música, medicina veterinária, projetos universitários, entre outros (Universidad Nacional Autónoma de México [UNAM], 2022). O Portal tem uma estrutura onde, inicialmente, apresenta os números estatísticos das coleções, contando com 2.106.655 registros nas diversas coleções. A disponibilização de acesso aos datasets se dá em acesso aberto, mas alguns conjuntos de dados contam com algumas restrições que demandam encaminhamento de solicitação de acesso ao coordenador do projeto. Todos os conteúdos são licenciados com Creative Commons Atribuição 4.0 Internacional (CC-BY 4.0), exceto quando se indique o contrário.

3.5.2 Repositorio de Acceso Abierto (RIAA)

O RIAA é o repositório institucional de acesso aberto da Benemérita Universidad Autónoma de Puebla (BUAP), que permite consultas às informações científicas, tecnológicas e de inovação produzidas na Universidade, assegurando a disponibilidade e acessibilidade a esse conteúdo, sem restrições, desde que respeitada a observância das atribuições das licenças Creative Commons (Benemérita Universidad Autónoma de Puebla [BUAP], 2018b).

De acordo com a descrição de conteúdo do RIAA-BUAP, este abriga principalmente artigos, livros, teses, documentos de eventos científicos e

outros materiais voltados para a produção de conhecimento; edições especiais, livros incunábulos, materiais antigos, manuscritos e coleções fotográficas (BUAP, 2018a). Na descrição de conteúdo das comunidades e coleções não se identificou a indicação de dados de pesquisa. Contudo, a comunidade EcoBUAP (Ecosistema de Aprendizaje Abierto BUAP) dá acesso ao **Repositorio EcoBUAP**, o qual disponibiliza a comunidade Conjunto de Datos. No entanto ao realizar-se busca por assunto nesta comunidade, não se obteve resultados.

4 Considerações sobre os repositórios

Surgidos no contexto do Open Access, no início dos anos 2000, os repositórios institucionais (RI) são uma extensão do conceito de repositórios digitais desenvolvidos na década de 1980, originalmente para o armazenamento de softwares de computadores (Macgregor, 2023). Suas características específicas estão diretamente relacionadas ao conteúdo documental a ser armazenado. De forma geral Bicknese (2003-2004), na década de 2000, classificou os repositórios em dois modelos, sendo o primeiro modelo destinado à coleta de documentos de uma única disciplina, e o outro modelo destinado à coleta de todo material produzido por uma faculdade ou universidade, referindo-se a este como sendo os repositórios institucionais. Posteriormente, à medida que os repositórios digitais foram evoluindo tecnicamente e assim diversificando seus propósitos e funcionalidades, desencadeou-se o surgimento de novas e distintas tipologias (Macgregor, 2023). Na tipologia proposta por Macgregor (2023) têm-se alguns exemplos em funcionamento, são os repositórios: de preprints (ex., EmeRI, 2020, cooperação entre a ABEC e o IBICT, com apoio da Unesco); de objetos de aprendizagem (MIT OpenCourseWare, 2001); de dados governamentais (Portal Brasileiro de Dados Abertos); agregadores (LA Referencia, 2012, Acordo de Cooperação entre países da América Latina); mega-repositórios (repositórios de escopo generalista, a exemplo do Zenodo, 2013); e os de dados de pesquisa. Macgregor (2023) ainda inclui em sua proposta os “repositórios confiáveis” que

seriam aqueles que satisfazem a requisitos rigorosos de confiança, lhes sendo atribuídos um selo como o “CoreTrustSeal”.

Para os repositórios digitais de dados de pesquisa, segundo Sayão e Sales (2016) há um certo consenso na literatura quanto a sua tipologia, caracterizando-os como: repositórios institucionais de dados de pesquisa, vinculados a instituições acadêmicas ou institutos de pesquisas, com a finalidade de armazenar toda produção científica gerada por seus pesquisadores; repositórios disciplinares de dados de pesquisa, destinados ao arquivamento de material de áreas específicas do conhecimento; repositórios multidisciplinares de dados de pesquisa, reúnem datasets de diversas áreas do conhecimento; e os repositórios de dados de pesquisa orientados por projetos, os quais recebem coleções de dados oriundas de projetos de pesquisa ou resultantes de resolução de problemas específicos.

A partir das tipologias apresentadas e das análises realizadas, têm-se que dos 59 repositórios recuperados no re3data, somente 20 são repositórios digitais institucionais vinculados a universidades, e dentre esses 20, somente seis são repositórios institucionais de dados de pesquisa, sendo os outros 14, repositórios institucionais acadêmicos, como mostra o Quadro 1.

Países	Total de repositórios no re3data	Repositórios Universitários	Repositório de Dados de Pesquisa
Argentina	8	4	1
Brasil	18	4	3
Chile	2	1	1
Colômbia	12	9	1
México	17	2	0
Panamá	2	0	0
Total	**59**	**20**	**6**

QUADRO 1.Quantitativo de repositórios universitários e de dados recuperados no re3dataFonte

Fonte: Dados da pesquisa

Pode-se observar no Quadro 1, que o Brasil apresenta o maior número de repositórios institucionais registrados no Diretório, 18 repositórios, seguido pelo México com 17, Colômbia com 12, Argentina com oito, e Chile com somente dois repositórios. Quando se seleciona somente os repositórios institucionais de universidades, a Colômbia aparece com o maior número de repositórios registrados no re3data, seguida pelo Brasil e Argentina, depois México e, por último o Chile. Ao passo que dentre os repositórios universitários, identificando-se somente os que são de dados de pesquisa, o Brasil aparece novamente, com o maior quantitativo registrado no Diretório, seguido pela Argentina, Chile e Colômbia com somente um repositório cada país. Por meio de divulgação de grupos de pesquisas e instituições de pesquisa, sabe-se que no Brasil além dos três repositórios registrados no re3data, estão em fase de planejamento distintas iniciativas de implementação de repositórios de dados de pesquisa (PAVÃO et al., 2022).

No Quadro 2, a seguir, encontra-se uma síntese sobre os 20 repositórios recuperados no re3data, podendo-se identificar: o país de origem; o ano de sua criação; sua tipologia; software utilizado para o gerenciamento das coleções; resultado da busca por dados de pesquisa na plataforma; e como se dá a disponibilização de acesso a esses dados.

REPOSITORIO	PAÍS	ANO [2] CRIAÇÃO	TIPO DE REPOSITÓRIO	*SOFTWARE*	RECUPERAÇÃO DADOS DE PESQUISA	TIPO DE ACESSO AOS DADOS	LICENÇA ATRIBUÍDA
Producción Académica UCC		2013	Institucional acadêmico	EPrints	sim	Acesso aberto	(CC BY-NC-ND 4.0)
Repositorio de Datos Académicos UNR		2022	Dados de pesquisa	Dataverse	sim	Acesso aberto	(CC BY4.0)

[2] Informações coletadas nos sites dos repositórios e em pesquisas na *web*.

Repositorio	País	Ano [2] criação	Tipo de repositório	*Software*	Recuperação dados de pesquisa	Tipo de acesso aos dados	Licença atribuída
(RDA-UNR)							
Repositorio Institucional UCASAL		2017	Institucional acadêmico	PhpMyBibli (PMB)	não	Não identificado	Não especificado
SEDICI (Servicio de Difusión de la Creación Intelectual)		2003	Institucional acadêmico	DSpace	sim	Acesso aberto	(CC BY-ND)
Base de Dados Científicos da Universidade Federal do Paraná		2017	Dados de pesquisa	DSpace	sim	Acesso aberto	(CC0 1.0), ou (CC BY-NC-ND 3.0 BR)
Open Research Data@PUC-Rio		2015	Dados de pesquisa	Maxwell	sim	Acesso aberto	(CC BY-NC-SA 4.0)
Repositorio Institucional UNESP		2016	Institucional acadêmico	DSpace	sim	Acesso aberto e fechado	Não especificado
Repositorio de Dados de Pesquisa da Unicamp (REDU)		2020	Dados de pesquisa	Dataverse	sim	Acesso aberto e fechado	(CC0)
Repositorio de Datos de Investigación de la Universidad de Chile		2019	Dados de pesquisa	Dataverse	sim	Acesso aberto	(CC0)
Repositorio Institucional Intellectum (RII)		2012	Institucional acadêmico	DSpace	não	Não identificado	(CC BY-NC-SA 4.0)
Repositorio Universidad Autónoma de Bucaramanga		2021	Institucional acadêmico	DSpace	não	Não identificado	Não especificado
Repositorio Institucional Universidad Distrital - RIUD		2016?	Institucional acadêmico	DSpace	não	Não identificado	Não especificado
Repositorio Institucional de la		?	Institucional	DSpace	não	Não	Não

Repositorio	País	Ano [2] criação	Tipo de repositório	*Software*	Recuperação dados de pesquisa	Tipo de acesso aos dados	Licença atribuída
Universidad Santo Tomás			acadêmico			identificado	especificado
Repositorio UNAL		?	Institucional acadêmico	DSpace	não	Não identificado	Não especificado
Repositorio Institucional Universidad de Antioquia		2008	Institucional acadêmico	DSpace	não	Não identificado	Não especificado
Repositorio Institucional Universidad Autónoma de Manizales (REPOUAM)		?	Institucional acadêmico	DSpace	sim	Acesso aberto	(CC BY-NC-SA 4.0)
Repositorio Digital UTB		2017?	Institucional acadêmico	DSpace	não	Não identificado	Não especificado
Repositorio de datos de investigación de la Universidad del Rosario		2019?	Dados de pesquisa	Dataverse	sim	Acesso aberto e fechado	(CC0) e (CC BY-NC 4.0),
Portal de Datos Abiertos da Universidad Nacional Autónoma de México (UNAM)		2016	Institucional acadêmico	Não identificado	sim	Acesso aberto e fechado	(CC-BY 4.0)
Repositorio de Acceso Abierto (RIAA)		2022?	Institucional acadêmico	DSpace	não	Acesso aberto	Não especificado

Quadro 2. Repositórios, país, ano de criação, tipo, software, recuperação e tipo de acesso

Fonte: Elaborado pelas autoras

O mais antigo dos repositórios analisados é o da Universidad de Antioquia, criado em 2008, enquanto os outros repositórios foram implantados somente entre os anos de 2016 e 2022. Destes repositórios, a grande maioria

corresponde a repositórios institucionais acadêmicos que disponibilizam acesso a diferentes tipologias documentárias provenientes da produção acadêmica e de pesquisa das universidades. Entre esses documentos podem conter dados de pesquisas, como é o caso dos repositórios: Producción Académica UCC; SEDICI; Repositório Institucional UNESP; REPOUAM; e o Portal de Datos Abiertos da Universidad Nacional Autónoma de México (UNAM). Apenas seis repositórios identificados no re3data são específicos para dados de pesquisa.

Comparativamente, destaca-se o estudo realizado por Silveira (2021), onde foram analisadas as políticas de 18 repositórios universitários, em sua maioria dos Estados Unidos, e que disponibilizam as políticas em seus websites. Em comparação ao presente estudo, os repositórios da referida pesquisa foram criados na primeira década dos anos 2000, correspondendo a um período maior de consolidação de práticas de armazenamento, preservação e disponibilização de conteúdo digital. Isto posto, é possível inferir que a maturidade dos repositórios se impõe na apresentação de informações e normativas relacionadas aos mesmos e publicizadas em seus sites. Tal constatação pode explicar em parte a dificuldade em localizar informações no site ou repositórios latino-americanos.

No que se refere ao software utilizado na implementação dos repositórios, verifica-se que o DSpace está presente em 12 dos 20 repositórios analisados, possivelmente pelo fato de o DSpace ser um software livre, originalmente desenvolvido para a criação de repositórios e ser facilmente adaptado pelas instituições (IBICT, 2012). Em segundo lugar aparece o Dataverse, presente em quatro repositórios, sendo estes específicos para dados de pesquisa, podendo sua escolha ser justificada por ser um software desenvolvido para atender as demandas desse tipo de repositório (Pavão et al., 2018, p. 42). Em contrapartida, a Base de Dados Científicos da Universidade Federal do Paraná, utiliza-se do DSpace para a gestão de suas coleções de dados de pesquisa. Essa escolha provavelmente está atrelada ao uso do DSpace no repositório

institucional da Universidade. De acordo com Treloar, Groenewegen e Harboe-ree (2007), embora a abordagem de um único repositório para toda produção científica de uma instituição possa parecer atraente, existe uma série de desafios relacionados a curadoria da ampla variedade de documentos. Em estudo comparativo entre os sistemas DSpace e Dataverse, realizado por Rocha et al (2021), os autores relatam que um diferencial do Dataverse em relação ao DSpace, é que este apresenta recursos de configurações que possibilitam a criação de vários tipos de ambientes de repositório, enquanto o DSpace exige um esforço maior em suas configurações para se alcançar o mesmo resultado.

Quanto a verificação da disponibilização dos dados de pesquisa, pôde-se constatar que em sua maioria os repositórios têm por prerrogativa o acesso aberto, contudo permitem o acesso restrito aos conteúdos se assim for necessário. Em relação às licenças de uso dos dados de pesquisa, verifica-se que, na grande maioria dos repositórios a indicação é de licenciamento com as licenças Creative Commos, atribuição CC0 ou uma Atribuição-Não-Comercial sem obra derivada (CC BY-NC-ND 4.0).

5 Considerações finais

Uma das limitações desta pesquisa refere-se à falta de informações nos repositórios e sites das bibliotecas e instituições para realizar uma pesquisa documental consistente. As políticas dos repositórios muitas vezes não fazem qualquer menção a dados de pesquisa, mesmo nos repositórios registrados na ferramenta re3data. Salienta-se a importância da publicização das políticas de repositórios, uma vez que estas são instrumentos oficiais norteadores das condutas a serem adotadas no acesso, depósito e uso do conteúdo armazenado nesses sistemas. Outro fato interessante e inesperado é a falta de dados de pesquisa nesses repositórios, constatando-se que o re3data não é totalmente confiável, possivelmente pela iniciativa de registro partir da própria

instituição, em uma autodeclaração. Sendo assim, infere-se que as informações registradas pelas universidades estejam em consonância com suas políticas institucionais, mesmo que não se tenha obtido sucesso nas buscas realizadas em alguns repositórios. Da mesma forma, como a maioria dos repositórios identificados são institucionais e não específicos de dados de pesquisa, infere-se que as buscas infrutíferas possam ser em decorrência de um processo de implementação de armazenamento desse conteúdo que ainda esteja em desenvolvimento nessas universidades.

Em relação a identificar a disponibilização de dados de pesquisa, como a maioria dos repositórios, 14 dos 20 investigados, não são exclusivos de dados de pesquisa, essa análise não foi satisfatoriamente possível. Dessa forma, considera-se que o objetivo do presente estudo não foi totalmente alcançado. A escolha por analisar repositórios universitários sem delimitar para repositórios específicos de dados de pesquisa acarretou prejuízo ao estudo. Da mesma forma, a utilização do re3data como uma única ferramenta de identificação dos repositórios parece não abranger o real cenário de iniciativas de implantação de repositórios de dados de pesquisa, a exemplo do Brasil.

Metodologicamente decidiu-se por analisar somente repositórios universitários, contudo no re3data recuperou-se um total de 39 outros repositórios que não fizeram parte do corpus desta pesquisa. Sendo assim, sugere-se para estudos futuros a análise desses repositórios para conhecer como se dá as iniciativas de abertura de dados de pesquisa nesses sistemas de informação.

Bibliografia

AMARO, B. (2021). Ciência aberta: desafios e perspectivas. <https://www.youtube.com/watch?v=7r_n_nrMwts&t=6558s>

BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA. (2018a). Contenido. <https://repositorioinstitucional.buap.mx/page/content>

BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA. (2018b). What are we? <https://repositorioinstitucional.buap.mx/page/buapAbout>

BICKNESE, D. (2003-2004). Institutional repositories and the institution's repository: what is the role of university archives with an institution's on-line digital repository? *Archival Issues, 28*(2), 81-94. <https://minds.wisconsin.edu/bitstream/handle/1793/45952/MA28_2_2.pdf?sequence=3&isAllowed=y>

ESTEVÃO, J. S. B. (2018). Apresentação da Base de Dados Científicos da Universidade Federal do Paraná. Curitiba. 50 slides. <https://acervodigital.ufpr.br/bitstream/handle/1884/56095/Apresenta%c3%a7%c3%a3o_BDC%20%20UFPR.pdf?sequence=1&isAllowed=y>

IBICT. (2012). Sistema para Construção de Repositórios Institucionais Digitais (DSpace): sobre o DSpace. <http://sitehistorico.ibict.br/pesquisa-desenvolvimento-tecnologico-e-inovacao/Sistema-para-Construcao-de-Repositorios-Institucionais-Digitais>

MACGREGOR, G. (2023). Digital repositories and discoverability: definitions and typology. In: Discoverability in digital repositories: systems, perspectives, and user studies. <https://strathprints.strath.ac.uk/78405/7/Macgregor_2022_Digital_repositories_and_discoverability_definitions_and_typology.pdf>

PAVÃO, C. et al. (2023). Planejamento e incubação de repositórios de dados em ICTs brasileiras. In: Actas de la XI Conferencia Internacional de Bibliotecas y Repositorios Digitales, (p. 233-249). La Plata: Universidad Nacional de La Plata: SEDICI. <https://www.lume.ufrgs.br/bitstream/handle/10183/255620/001162849.pdf?sequence=1>

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO. (2022). Sobre [laboratório LAMBDA]. <https://www.maxwell.vrac.puc-rio.br/sobre.php>

ROCHA, R. P. DA ET AL. (2018). Acesso aberto a dados de pesquisa no Brasil: soluções tecnológicas: relatório 2018. <http://hdl.handle.net/10183/185126>

ROCHA, R. P. DA ET AL. (2021). Análise dos sistemas DSpace e Dataverse para repositórios de dados de pesquisa com acesso aberto. *Revista Brasileira de Biblioteconomia e Documentação, 17*, 1-25. <https://brapci.inf.br/index.php/res/download/160968>

SALES, L. F. E SAYÃO, L. F. (2019). Uma proposta de taxonomia para dados de pesquisa. *Conhecimento em Ação, Rio de Janeiro, 4*(1), 31-48. <https://revistas.ufrj.br/index.php/rca/article/download/26337/14573>

SAYÃO, L. F. E SALES, L. F. (2016). Algumas considerações sobre os repositórios digitais de dados de pesquisa. *Informação & Informação,, 21*(2), 90-115.

<https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/27939/20122>

SILVEIRA, J. I. E VANZ, S. A. de S. (no prelo). Diretrizes para políticas de depósito, acesso e uso de dados de pesquisa: proposta a partir da análise de repositórios de dados universitários internacionais. AtoZ Novas Práticas em Informação e Conhecimento.

TORINO, E., ROA-MARTÍNEZ, S. M. E VIDOTTI, S. A. B. G. (2020). Dados de pesquisa: disponibilização ou publicação? In: Shintaku, M.; Sales, L. F e Costa, M. (Orgs). Tópicos sobre dados abertos para editores científicos p. 183-201. Botucatu, SP: ABEC, <doi:10.21452/ 978-85-93910-04-3.cap15>

TRELOAR, A., GROENEWEGGEN, D.,E HARBOE-REE, C. (2007). The Data Curation Continuum: Managing Data Objects in Institutional Repositories. *D-Lib Magazine,* *13*(9-10). <http://www.dlib.org/dlib/september07/treloar/09treloar.html>

UNIVERSIDAD CATÓLICA DE CÓRDOBA. (2021). Política del Repositorio Institucional Producción Académica. <http://pa.bibdigital.ucc.edu.ar/policies.html>

UNIVERSIDAD CATÓLICA DE SALTA. (2017). Resolución Rectoral nº 1017/17. <http://bibliotecas.ucasal.edu.ar/opac_css/doc_num_data.php?explnum_id=1084>

UNIVERSIDADE ESTADUAL PAULISTA. (2022). Apresentação. <https://repositorio.unesp.br/>

UNIVERSIDAD DISTRITAL FRANCISCO JOSÉ DE CALDAS. (2022). Formato Licencia de uso y publicación. <https://repository.udistrital.edu.co/handle/11349/3841>

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO. (2022). Sobre el Portal de Datos Abiertos UNAM. <https://datosabiertos.unam.mx/informacion/sobreportal.html>

UNIVERSIDAD NACIONAL DE LA PLATA. (2022). Políticas del repositorio. <http://sedici.unlp.edu.ar/pages/politicas>

UNIVERSIDAD NACIONAL DE ROSARIO. (2022). Se presentó el Repositorio de Datos Académicos. <https://unr.edu.ar/la-unr-presento-su-repositorio-de-datos-academicos/>

VANZ, S. A. DE S. E SILVA FILHO, R. C. (2019). O protagonismo das revistas na comunicação científica: histórico e evolução. In: Carneiro, F. F. B., Ferreira Neto, A. e Santos, W. dos. (Org.). *A comunicação científica em periódicos*. (1ed.), p. 19-44. Curitiba: Appris.