



CAPÍTULO 12

REPOSITÓRIOS DIGITAIS ABERTOS

Fernanda Maciel Rufino¹
Bernardo Dionízio Vechi²
Milton Shintaku³



¹ Universidade de Brasília (UnB) e Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict). ORCID: <https://orcid.org/0000-0002-7593-8146>. E-mail: fernandarufino@ibict.br.

² Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict). ORCID: <https://orcid.org/0000-0002-7727-3889>; E-mail: bernardovechi@ibict.br.

³ Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) e Universidade Federal do Paraná (UFPR). ORCID: <https://orcid.org/0000-0002-6476-4953>; E-mail: shintaku@ibict.br.

12.1 INTRODUÇÃO

O surgimento da internet, ainda na década de 1970, possibilitou a criação de redes interligando computadores, que na época ainda eram de grande e médio porte, em sua grande parte institucional ou departamental. Assim, parte dos serviços ofertados pela internet inicial estava centrada no envio e recebimento de mensagens e na troca de arquivos. Com isso, fomentou a criação de documentos em formato digital, mesmo em formatos mais simples.

Com o surgimento dos computadores pessoais, em 1976, com o lançamento do modelo Apple I, da empresa de mesmo nome, uma nova era na criação de documentos em formato digital também se iniciou. Com os computadores pessoais – *personal computer* –, chamados de PC, os usuários puderam produzir conteúdo em formato digital, mesmo que ainda se mantivessem certas características do impresso, como se criassem documentos e imprimissem no formato digital.

Esse formato de documentos nato digitais, mas como se fossem impressos digitalmente, mantiveram-se em muitas iniciativas, como nos casos dos documentos portáteis, do inglês *Portable Document Format* (PDF). Em parte, por não permitir mudanças no conteúdo, como no caso dos impressos, o que garante, em parte, a sua integridade, mas que cria limitações ante as possibilidades do mundo digital.

Na década de 1990, com o surgimento da *World Wide Web* (WWW), ou apenas a Web, com todas as suas possibilidades, descortinou-se grandes oportunidades, principalmente com o uso das linguagens de marcação, como o *HyperText Markup Language* (HTML – Linguagem de Marcação de Hipertexto) ou o *Extensible Markup Language* (XML – Linguagem de Marcação Extensível). Inicialmente com as páginas estáticas, criadas como serviços informacionais, em que o conteúdo era apresentado de forma hipertextual, modificaram a percepção de como um documento digital poderia ser, mesmo que muitas instituições mantenham os documentos impressos digitalmente na disseminação da informação.

Para a atuação com a informação, o surgimento da Web teve um impacto grande, a ponto de estudiosos compararem a criação da prensa de tipos móveis de Johannes Gutenberg (1400-1468). Assim como a prensa de tipos móveis possibilitou uma democratização dos livros, e dessa forma a informação impressa, a Web possibilitou que pessoas pudessem acessar a informação de forma remota, por meio do seu computador pessoal, aumentando de forma significativa a possibilidade de acesso ao conteúdo.

Nesse contexto, a Web está em evolução, de forma que cada vez mais novas funcionalidades sejam agregadas. Tanto que O'Reilly (2005) cunhou o termo Web 2.0, apresentando as diferenças ocorridas neste ambiente desde a sua criação, chamada pelo autor de Web 1.0, que era composta praticamente de páginas estáticas, onde os usuários eram apenas consumidores de conteúdo, prioritariamente textual. Na versão da Web 2.0, o ambiente se tornou mais interativo, em que os usuários passam a ser produtores de conteúdo, com *blogs* e redes sociais.

Posteriormente, com a evolução tecnológica, a próxima versão Web 3.0 nasce junto com o uso cada vez maior dos dispositivos móveis e a força das redes sociais. Outro ponto é o uso da Web Semântica, integração entre dispositivos e Inteligência Artificial. A próxima geração do ambiente Web, a 4.0, está fundamentada no uso maior das nuvens e Inteligência Artificial, com um imenso sistema operacional Web que utiliza todos os recursos disponíveis.

Entretanto, mesmo com toda essa evolução alguns problemas ainda são recorrentes. Entre os problemas está o formato dos documentos, principalmente artigos, teses e dissertações, documentos oficiais, que ainda se mantêm como impressos na Web, formato textual com pouco uso das funcionalidades hipertextuais. Outro ponto está na quantidade de documentação, que forma um acervo humanamente impossível de verificar a sua relevância.

Por isso, alguns sistemas de informação foram desenvolvidos para a gestão documental das instituições ou organizações, de forma a facilitar o processo de seleção, depósito e compartilhamento de documentos digitais. Com isso, atender a necessidade de dar acesso ao documento em texto completo, descrevê-los de forma a facilitar a sua recuperação e armazená-los de forma organizada.

12.2 MOVIMENTO DOS ARQUIVOS ABERTOS PRECURSOR DO REPOSITÓRIO

Em 1991, o pesquisador Paul Ginsparg, professor de física e computação da Cornell University, no estado de Nova Iorque, Estados Unidos da América, começou o desenvolvimento do ArXiv, voltado para disseminação dos artigos de *preprints* – pré-publicações ou pré-impressos – na universidade de Los Alamos. Essa iniciativa é muito importante por tratar dos chamados *preprints*, prévia de publicação, utilizado para se referir a artigos na versão do autor, ou seja, antes de passar por processo editorial das revistas.

O sistema de informação ArXiv revolucionou parte do cenário de comunicação científica fornecendo acesso amplo a documentação científica que até então era restrita. Possibilitou que os autores de artigos ainda não avaliados pelos pares pudessem apresentá-lo à comunidade, sem o intermédio de canais de publicação, no que depois foi conhecido como autoarquivamento, sendo a ação do autor de depositar o documento digital em sistema de informação.

O ArXiv ainda está em atividade sendo um canal importante de disseminação da informação científica. Tanto que Ginsparg (2011), diante da evolução desta importante iniciativa, denominou-o de ArXiv 2.0, apresentou os resultados de mais de duas décadas, ressaltando o papel de democratizador da informação do ArXiv, dando visibilidade a produção científica, mesmo antes da sua publicação. Para o autor, o ArXiv ainda é útil na comunicação científica, para uma geração nativa digital, com menor barreira no acesso à informação científica.

O ArXiv foi um precursor de um grande movimento nas ciências chamado *Open Archives* (Arquivos Abertos), lançado oficialmente na *Convenção de Santa Fé*, no Novo México, em 2000¹. Para os idealizadores do evento, na época, chamou-se de *archives* (arquivos) os documentos científicos que estavam sendo publicados digitalmente. Dessa forma, os arquivos poderiam ser disseminados de forma informal, distribuídos pela internet, estabelecendo uma nova modalidade para a comunicação científica.

O movimento *Open Archives* (Arquivos Abertos) foi crucial para a modernização de processos da comunicação científica diante da evolução tecnológica, tanto que muitas das suas premissas foram absorvidas pelos movimentos posteriores, como o movimento *Open Access* (Acesso Aberto), por exemplo. De certa forma, o movimento dos Arquivos Abertos foi o primeiro grande movimento científico voltado para a modernização da comunicação científica no mundo digital.

Nas suas premissas, uma das que foram adotadas de forma ampla, foi o que era chamado de autoarquivamento, ou seja, um processo pelo qual os autores submetem os seus manuscritos aos sistemas de informação. Esse processo, inicialmente utilizado no ArXiv, foi amplamente aceito em vários outros sistemas, como nas revistas eletrônica, repositórios e bibliotecas digitais. Vale lembrar que antes do autoarquivamento, muitas instituições recebiam os originais por e-mail ou correio, em dispositivos como disquetes.

Sistemas de informação com um módulo de submissão pode parecer simples, mas não era comum na chamada Web 1.0, sendo umas das mudanças que mar-

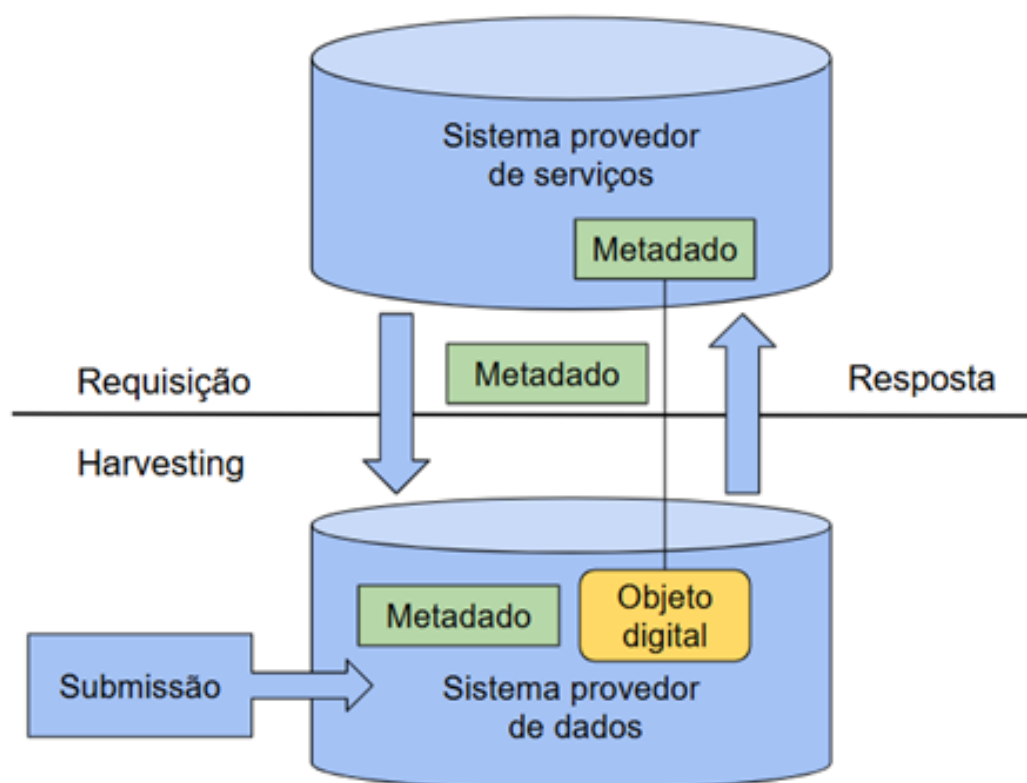
¹ Disponível em: https://www.openarchives.org/sfc/sfc_entry.htm. Acesso em: 24 abr. 2025.

caram o Web 2.0. O processo de autoarquivamento inicialmente era um processo simples no qual o usuário depositaria o seu objeto digital sem a intervenção de outras equipes. Por causa do autoarquivamento, evidenciaram-se questões sobre qualidade dos metadados, na medida em que esse processo consiste no preenchimento de formulário de metadados, carregamento de objetos digitais e marcação sobre direitos autorais e licenciamento.

Como dito anteriormente, no movimento dos Arquivos Abertos, nomeava-se os sistemas de informação como *arquivos*, possivelmente influenciados pelo projeto ArXiv. Tanto que, nos documentos da *Convenção de Santa Fé*, relatam sobre a unidade de informação *e-prints* (impressão eletrônica), para definir os documentos em formato digital, e os arquivos no qual essa documentação estaria disponibilizada para acesso ao texto integral dos *e-prints*. Posteriormente, muitas dessas nomenclaturas foram alteradas, com a evolução conceitual e tecnológica.

Possivelmente, uma das contribuições dos Arquivos Abertos, que implementaram uma funcionalidade tecnológica distinta, tenha sido a interoperabilidade, termo até então pouco conhecido na comunicação científica. Para implementação dessa nova funcionalidade (Figura 12.1) nos sistemas de informação, foi preciso definir uma série de elementos constituintes da interoperabilidade.

Figura 12.1 – Modelo de interoperabilidade



Fonte: Os autores (2025), conforme os preceitos dos Arquivos Abertos.

Neste modelo da Figura 12.1, o Sistema Provedor de Dados implementa o processo de submissão (autoarquivamento), no qual o autor submete um objeto digital acompanhado pelo seus metadados, compondo o item. Esses metadados devem atender a um princípio básico de conjunto mínimo, muitas vezes utilizando o padrão Dublin Core. O Sistema Provedor de Dados deve também atender a utilização de um identificador único e persistente (*handle*) para o conjunto de metadados e objeto digital (item).

O Sistema Provedor de Serviços, por sua vez, faz uso de um protocolo de comunicação, que possibilita executar o processo de *Harvesting* (coleta), emitindo uma requisição ao provedor de dados, tendo como resposta os metadados do item. Assim, o provedor de serviços passa a ter uma cópia dos metadados, no qual consta um link para o objeto digital. Dessa forma, um usuário pode acessar o objeto digital pelos Sistemas Provedor de Dados e Provedor de Serviços.

Um dos sistemas de informação que implementaram os preceitos dos Arquivos Abertos foram as chamadas Bibliotecas Digitais. Suleman (2002), um dos idealizadores da interoperabilidade, define a biblioteca digital como um sistema de armazenamento eletrônico de informação focado em atender às necessidades de busca de informação de seus usuários. O autor foi um dos responsáveis pelo desenvolvimento das primeiras versões do protocolo de interoperabilidade *Open Archives Initiative - Protocol for Metadata Harvesting*² (OAI-PMH), que possibilitou a criação dos provedores de dados e serviços.

Assim, um dos maiores expoentes no movimento de Arquivos Abertos, com o uso de interoperabilidade e bibliotecas digital, é a *Networked Digital Library of Theses and Dissertations* (NDLTD), um provedor de dados que coleta metadados de várias iniciativas no mundo. Para Suleman e Fox (2002), seus idealizadores, a NDLTD funciona como uma federação de entidades independentes que publicam teses e dissertações de forma digital. Com isso, revela o objetivo primeiro das bibliotecas digitais no mundo para publicação da literatura cinzenta, como as teses e dissertações.

No Brasil, como reflexo da NDLTD, o Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) instituiu o projeto Biblioteca Digital Brasileira (BDB) para a criação de um consórcio de bases de teses e dissertações em formato eletrônico (Marcondes; Sayão, 2003). Posteriormente, a BDB evoluiu para a Biblioteca Digital Brasileira de Teses e Dissertações (BDTD) com a agregação de quase todas as universidades que possuem programas de pós-graduação *stricto sensu* (Macêdo; Shintaku; Mathias; Novaes, 2013).

² Disponível em: <https://www.openarchives.org/pmh/>. Acesso em: 28 abr. 2025.

Outro ponto que merece destaque foi a criação da Biblioteca Digital Jurídica (BD-Jur), em 2004, pelo Superior Tribunal de Justiça (STJ), para disseminar a memória técnica do tribunal (Basevi, 2005). Com isso, as bibliotecas digitais se estabelecem no processo de disseminação de documentação que não passa por processo editorial tradicional, como a literatura cinzenta e a memória técnica. Mesmo que o movimento dos Arquivos Abertos tenha nascido no âmbito da comunicação científica, nota-se que atua em documentação mais ampla que os considerados como científicos na visão mais purista e restrita.

O movimento dos Arquivos Abertos foi suplantado pelo movimento do Acesso Aberto, logo no início dos anos 2000. Entretanto, grande parte das premissas dos Arquivos Abertos foram adotadas pelo novo movimento, tanto nas questões técnicas, quanto nas conceituais. Assim, o movimento de abertura das ciências se manteve em evolução, na qual o movimento dos Arquivos Abertos deu visibilidade à literatura cinzenta, técnica e aos *preprints*, que até então possuíam disseminação reduzida.

12.3 MOVIMENTO DE ACESSO ABERTO E O NASCIMENTO DOS REPOSITÓRIOS

O Movimento do Acesso Aberto (*Open Access*) inicialmente se mesclou ao seu antecessor, o Movimento de Arquivos Abertos (*Open Archives*), muito por possuírem a mesma sigla. Entretanto, os seus objetivos são diferentes, assim como a sua motivação. Enquanto o Arquivos Abertos tinha um foco mais na literatura cinzenta, técnica e *preprints*, o Acesso Aberto atua principalmente na documentação científica, inicialmente, nos artigos publicados em periódicos.

A necessidade da abertura – tornar livre o acesso – dos resultados de pesquisas publicadas em artigos de revistas possui antecedentes que remontam a década de 1970, como relata Mueller (2006), com a chamada crise dos periódicos. A situação foi sendo agravada conforme o negócio de editoração, dominado pelas grandes editoras, tornaram impeditivo a aquisição das assinaturas pelas bibliotecas, devido ao alto custo.

Assim, com o impulso tecnológico ofertado pelas bibliotecas digitais e seus sistemas de oferta de acesso ao texto completo surgem os repositórios. Originalmente, repositório, como a formação da palavra, indica o *locus* digital onde se deposita uma cópia de um documento publicado em outro canal. Tanto que Weitzel

(2006) descreve os repositórios como um sistema de informação que atua como segunda fonte. Por isso a formação morfológica do termo é *re* (novamente) + *por* (colocar) + *tório* (local), ou seja, local onde se recoloca algo. Posteriormente, esse termo ganha novas acepções.

Harnad *et al.* (2004), denominaram os repositórios como via verde, ou seja, um sistema de informação que dá sinal verde ao acesso gratuito a um documento publicado em outro lugar, geralmente restrito. O repositório (via verde) era complementado com os periódicos de acesso aberto (via dourada), que davam acesso livre aos seus artigos, como díade perfeita do Movimento de Acesso Aberto. Seguindo o entendimento, este movimento fundamentava-se nas revistas de acesso aberto, primeira fonte, e os repositórios, segunda fonte, para dar acesso livre aos resultados de pesquisa.

O Movimento de Acesso Aberto teve um impacto muito grande na comunicação científica, em vários sentidos (Costa, 2006), em parte por democratizar o acesso, mas também por dar visibilidade à produção científica de países fora do eixo Europa e Estados Unidos da América. Muito deste impacto deve-se aos repositórios, na medida em que se tornaram presentes em quase todas as universidades, como um sistema que dissemina a produção acadêmica da instituição.

Para Björk (2007), em seu modelo de comunicação científica, os repositórios são facilitadores de acesso à documentação produzida nas pesquisas. Com isso, os documentos que publicam os resultados de pesquisa continuam a ser disseminados pelos canais tradicionais como os artigos de periódicos e eventos e os livros, mesmo os de acesso livre, mas cópias são dispostas em repositórios aumentando a possibilidade de serem acessados e usados.

No Brasil, os repositórios foram catalisados por meio de um projeto de pesquisa desenvolvido pelo Ibict, financiado pela Financiadora de Estudos e Projetos (Finep), denominado BDB. Assim, *kits* tecnológicos foram distribuídos por meio de editais às bibliotecas de universidades que queriam implementar repositórios, juntamente com capacitação a profissionais de informação e informática, em especial às universidades públicas.

Como resultado, parte das universidades brasileiras possuem repositórios, em grande parte, com o uso do *software* livre DSpace, muito pela ação do Ibict. De modo que parte da produção acadêmica depositada nos repositórios é coletada pelo Portal Brasileiro de Publicações e Dados Científicos em Acesso Aberto³

³ Disponível em: <https://oasisbr.ibict.br/vufind/>. Acesso em: 28 abr. 2025.

(Oasisbr), tornando o Brasil como expoente na disseminação de acesso livre no mundo.

Se inicialmente os repositórios mantinham o acervo composto quase que exclusivamente por cópias de documentação científica publicada, aos poucos adotaram políticas mais flexíveis, disseminando outras documentações. Shintaku e Vidotti (2016) relatam como esse processo tornou os repositórios como publicadores de documentação de primeira fonte de documentação acadêmica, disseminando teses, dissertações e outros.

Nesse contexto, os repositórios se firmam como sistemas de informação que atuam na gestão da informação da instituição, tanto para os usuários internos quanto externos. Os repositórios possibilitam o acesso a grande parte da produção acadêmica de uma instituição, destarte as disciplinas que ainda publicam em periódicos fechados, possibilitando a democratização do conhecimento científico.

Mesmo que os periódicos de acesso aberto, na forma ouro ou diamante, sejam os sistemas de informação principal do Movimento de Acesso Aberto, por publicarem os resultados de pesquisa, os repositórios possuem um papel importante neste cenário. Tanto que, Aguillo *et al.* (2010), criaram um ranking webométrico de repositórios, em contrapartida as métricas de revistas, pela importância desses sistemas de informação para a comunicação científica. Com isso, promove a visibilidade do sistema e da instituição mantenedora.

Com o surgimento do Movimento da Ciência Aberta os repositórios se tornaram mais flexíveis, na medida em que expandiu as suas políticas para que o acervo fosse mais abrangente. Esta evolução se torna necessária, visto as mudanças ocorridas no cenário da comunicação científica, em muito pelo uso cada vez mais das facilidades oferecidas pela internet e informática, além da pressão da comunidade pela transparência no processo científico.

12.4 CIÊNCIA ABERTA E OS REPOSITÓRIO DIGITAIS ABERTOS

De uma forma superficial, pode-se apresentar a abertura das ciências pelos seus movimentos, iniciando com o Movimento dos Arquivos Abertos (*Open Archives*), dando acesso aos *preprints* e literatura cinzenta, e com o Movimento de Acesso Aberto (*Open Access*) que possibilitou o acesso ao texto integral dos resultados de

pesquisa. Entretanto, os primeiros movimentos de abertura estavam restritos aos resultados das pesquisas, às publicações com os resultados dos estudos, como artigos de eventos e revistas.

O Movimento da Ciência Aberta, por sua vez, abre outros elementos da pesquisa. Pinheiro (2014), por exemplo, ressalta a abertura dos dados de pesquisa, como se a Ciência Aberta fosse um alargamento teórico e prático do Movimento de Acesso Aberto, possibilitando novas análises, requerendo curadoria e gestão em sistemas de informação apropriados aos dados. Nesse ponto, surge um novo tipo de repositório, o de dados de pesquisa. Albagli, Clinio e Raychtock (2014) relata que a Ciência Aberta torna-se um termo guarda-chuva que transcende ao Acesso Aberto, incluindo dados científicos abertos, ferramentas científicas abertas, *hardware* científico aberto, cadernos científicos abertos e wikipesquisa, ciência cidadã, educação aberta e outros.

No que tange especificamente aos repositórios no âmbito do Movimento da Ciência Aberta, Silveira *et al.* (2023) apresentam uma taxonomia revisitada sobre o movimento, no qual tratam dos seguintes tipos de repositórios:

- Acesso Aberto
 - ▶ Repositórios (Via Verde)
 - ◇ Repositórios institucionais
 - ◇ Repositórios temáticos
 - ◇ Repositório de consórcio
 - ◇ Repositórios de *preprints*
- Dados Abertos
 - ▶ Repositórios de dados
- Infraestrutura e ferramentas científicas abertas
 - ▶ Repositórios de códigos abertos

Essa taxonomia proposta por Silveira *et al.* (2023) apresenta os temas adotados pela Ciência Aberta e mostra que este novo movimento incorporou o Movimento

do Acesso Aberto, assim como o Acesso Aberto incorporou os Arquivos Abertos. Tanto que trás os principais tipos e repositórios, desde o repositório de *preprint* (Movimento dos Arquivos Abertos), aos de consórcio que atende a mais de uma instituição, institucionais e temáticos de acesso aberto.

No que tange aos repositórios de dados, ainda não há um total consenso, visto que a própria ideia de dados ainda não está clara diante das diferenças nas pesquisas nas diversas disciplinas. Assim, um repositório de dados atende ao depósito dos dados coletados em pesquisa, independente do tipo, para compartilhamento e reuso, de forma organizada e com descrição que atenda as necessidades de recuperação. Assim, utiliza os preceitos de repositório do Movimento de Acesso Aberto, ajustando-os para os dados de pesquisa.

Sayão e Sales (2016), tratando dos repositórios de dados, relatam sobre os benefícios dos repositórios dos dados, sendo:

- **Visibilidade dos dados:** ao disponibilizar os dados de forma ampla, proporciona visibilidade aos dados, dos seus produtores e da instituição que os gerou;
- **Compartilhamento dos dados:** repositórios possibilitam a democratização do acesso aos dados de pesquisa;
- **Crédito ao autor dos dados:** possibilita que se identifiquem os autores dos dados, possibilitando inclusive a citação;
- **Preservação digital:** possibilita que as bases de dados possam estar disponíveis por longo espaço de tempo;
- **Memória científica e transparência:** promove a memória da pesquisa, por meio da oferta de acesso aos dados por longo tempo, ofertando transparência na pesquisa;
- **Segurança dos dados:** repositórios, como um sistema de informação, possuem processos de preservação e anti-intrusão, o que garante certa segurança;
- **Disponibilidade:** os dados em repositórios estão disponíveis online para acesso por pessoas ou outros sistemas;

- **Curadoria digital:** repositórios promovem fluxos informacionais de avaliação, de adição de valor, reformatação, agregação e recriação de dados, entre outros;
- **Serviços inovadores:** a evolução técnica e tecnológica possibilita a criação de novos serviços ofertados pelos repositórios de dados;
- **Reuso dos dados:** promove o uso e reuso dos dados de forma a diminuir esforços repetidos;
- **Rede de repositórios:** Permite que os repositórios possam se agregar em redes, com serviços de busca ou de atendendo a preceitos do Linked data;
- **Indicador de qualidade e produtividade da instituição:** os dados nos repositórios são evidências da pesquisa, representando sua produtividade, possibilitando avaliar a qualidade das pesquisas.

Para os mesmos pesquisadores, seguindo, em parte dos tipos de repositórios institucionais de resultados de pesquisa, e estudos de outros estudiosos, chegaram a uma tipologia básica de tipos de repositórios de dados. Assim, podem-se estabelecer os seguintes tipos de repositórios de dados: repositórios institucionais de dados de pesquisa, repositórios disciplinares de dados de pesquisa, repositórios multidisciplinares de dados de pesquisa e repositórios de dados de pesquisa orientados por projetos.

Quanto aos repositórios de códigos abertos, não é uma iniciativa criada pela Ciência Aberta, mas incorporada preceitos originários do movimento dos *softwares* livres, nascido ainda na década de 1980, no qual mudava o modelo de negócio sobre ferramentas informatizadas. Os *softwares* livres mudaram a forma pelo qual produtores e adquirentes de ferramentas informatizadas negociavam, tratando-os como bens para serviços. Assim, um *software* deixa de ser um bem comercializado e passa a ofertar serviços pagos de manutenção, apoio, atualização e outros.

Da mesma forma, por atuar na democratização do uso dessas ferramentas sem custos, surgiram os repositórios de códigos fontes, para a sua distribuição. Com isso, outro conceito foi criado, a de comunidade, que representa os usuários da ferramenta e seus desenvolvedores. Nesses repositórios, conhecidos pela comunidade como Git⁴, em referência a um projeto desenvolvido por Linus Torvalds

⁴ Git é um sistema de controle de versão distribuída, é um repositório de controle de versão completa, onde os desenvolvedores confirmam o trabalho localmente e depois sincronizam a cópia do repositório com a cópia no servidor (Microsoft, 2023).

criador do sistema operacional livre Linux, as ferramentas ficam disponíveis para serem baixadas e utilizadas sem nenhum custo.

Nas ciências, o uso de *software* livre é bastante comum. No contexto do acesso aberto, Costa (2006) advoga pela chamada “filosofia aberta”, consistindo do uso de *softwares* livre para disponibilizar documentação sem barreiras de acesso. Assim, repositórios de códigos abertos pode ser um termo polissêmico e abrangente, podendo se tratar da iniciativa, dos repositórios de *softwares* abertos, de licenciamento livre e códigos abertos, mas também dos *softwares* abertos para a criação dos diversos tipos de repositórios.

Na taxonomia proposta por Silveira *et al.* (2023), repositórios de códigos abertos estão contextualizados na Infraestrutura e Ferramentas Científicas Abertas. Assim, pode-se olhar os repositórios de códigos tanto como infraestrutura, quanto como ferramenta. Dessa forma, ter os repositórios e as ferramentas para criação dos repositórios, todos utilizando preceitos da abertura, de democratização, colaboração, entre outros.

Assim, repositórios na Ciência Aberta assumem três grandes papéis: (i) para disseminar resultados de pesquisa (mantendo as premissas do Acesso Aberto); (ii) para os dados de pesquisa (voltada ao movimento dos dados abertos), e; (iii) para disponibilizar *softwares* abertos (alinhados ao movimento do *software* livre). Dessa forma, pode-se inferir que os repositórios nesse movimento foram adotados de outras iniciativas da ciência, alinhados aos preceitos desse movimento.

Nesse contexto, reforça a premissa de que o Movimento da Ciência Aberta é um grande guarda-chuva, que cobre uma série de temas, muitos dos quais são oriundos de outros movimentos. Esse ponto reforça a máxima da ciência em que a transformação e reuso como sua base. Não se cria novos conhecimentos sem usar o compêndio existente. Ao quebrar paradigmas, utiliza-se o conhecimento existente, muitas vezes, reunindo os de outras áreas.

Nesse caminho, reforça também como a tecnologia serviu de catalisador aos processos de abertura, principalmente no que diz respeito aos repositórios, que precisam de ferramentas informatizadas para a sua implementação. Assim, repositórios digitais abertos na Ciência Aberta tem uma infraestrutura informatizada, mas são estabelecidos com base teórica que garantam a sua abertura.

12.5 CONSIDERAÇÕES FINAIS

Na história dos repositórios, iniciada ainda no final do século XX, destaca-se que a sua finalidade de compartilhar objetos digitais, inicialmente textos, depois passando para dados e *softwares*. Assim, tornam-se um instrumento de compartilhamento de recursos, voltados à democratização do acesso aos mais variados tipos de itens em formato digital, tudo de forma aberta, ou seja, sem barreiras.

Não há como negar que a evolução da tecnologia teve papel importante para o surgimento dos repositórios, principalmente a criação da Web, que deu aos primeiros movimentos de abertura das ciências as ferramentas necessárias para a sua implementação. Há uma infinidade de *softwares* abertos para a criação de repositório, dependendo do seu tipo, de forma alinhada ao movimento do *software* livre e da Ciência Aberta. Assim, ferramentas como o DSpace, Dataverse, Ckan, Omeka, Tainacan, Gitlab, entre tantos outros são utilizados para a criação de repositórios abertos para compartilhar resultados de estudos, dados de pesquisas e *softwares* livres.

Por fim, torna-se quase impossível atuar nas ciências ignorando os seus movimentos de abertura, mesmo naquelas disciplinas que adotam a informática mais tardiamente. Repositórios abertos se tornaram comuns na execução das pesquisas, como fonte de informação e para depósito dos resultados, ou mesmo na necessidade de compartilhar dados. Está se tornando comum que revistas peçam que os dados de pesquisa estejam disponíveis para aceitar artigos, da mesma forma que pesquisas que desenvolvem *software*, os disponibilizem de forma livre em repositórios.

REFERÊNCIAS

AGUILLO, Isidro F.; ORTEGA, José L.; FERNÁNDEZ, Mario; UTRILLA, Ana M. Indicators for a webometric ranking of open access repositories. **Scientometrics**, [s. l.], v. 82, n. 3, p. 477-486, Mar. 2010. DOI: <https://doi.org/10.1007/s11192-010-0183-y>. Disponível em: <http://link.springer.com/10.1007/s11192-010-0183-y>. Acesso em: 30 nov. 2023.

ALBAGLI, Sarita; CLINIO, Anne; RAYCHTOCK, Sabryna. Ciência Aberta: correntes interpretativas e tipos de ação. **Liinc em Revista**, Rio de Janeiro, v. 10, n. 2, p. 434-450, nov. 2014. DOI: <https://doi.org/10.18617/liinc.v10i2.749>. Disponível em: <http://revista.ibict.br/liinc/article/view/3593>. Acesso em: 30 nov. 2023.

BASEVI, Teresa. BDJur Consortium: juridical digital library: implementing DSpace in the Brazilian Judiciary. In: INTERNATIONAL CONFERENCE ON ELECTRONIC PUBLISHING (ELPUB), 9., 2005. **Proceedings** [...]. Leuven-Heverlee, Bélgica: Peeters Publishing Leuven, 2005. p. 127-132. Disponível em: <https://elpub.architexturez.net/system/files/pdf/150elpub2005.content.pdf>. Acesso em: 30 nov. 2023.

BJÖRK, Bo-Christer. A model of scientific communication as a global distributed information system. **Information Research**, [s. l.], v. 12, n. 2, p. 1-48, Jan. 2007. Disponível em: <https://informationr.net/ir/12-2/paper307.html>. Acesso em: 28 abr. 2025.

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, n. 2, p. 39-50, maio/ago. 2006. DOI: <https://doi.org/10.18225/ci.inf.v35i2.1139>. Disponível em: <https://revista.ibict.br/ciinf/article/view/1139>. Acesso em: 30 nov. 2023.

GINSPARG, Paul. ArXiv at 20. **Nature**, [s. l.], v. 476, n. 7359, p. 145-147, 11 Aug. 2011. DOI: <https://doi.org/10.1038/476145a>. Disponível em: <http://www.nature.com/articles/476145a>. Acesso em: 31 mar. 2020.

HARNAD, Stevan; BRODY, Tim; VALLIÈRES, François; CARR, Les; HITCHCOCK, Steve; GINGRAS, Yves; OPPENHEIM, Charles; STAMERJOHANN, Heinrich; HILF, Eberhard R. The access/impact problem and the Green and Gold roads to Open Access. **Serials Review**, [s. l.], v. 30, n. 4, p. 310-314, Jan. 2004. DOI: <https://doi.org/10.1080/00987913.2004.10764930>. Disponível em: <http://www.tandfonline.com/doi/abs/10.1080/00987913.2004.10764930>. Acesso em: 3 nov. 2020.

MACÊDO, Diego José; SHINTAKU, Milton; MATHIAS, Gabriel Franklin; NOVAES, Roberto. Biblioteca Digital Brasileira de Teses e Dissertações: dez anos de interoperabilidade. In: CONFERÊNCIA LUSO-BRASILEIRA SOBRE ACESSO ABERTO, 4.,

26 jul. 2013. **Anais** [...]. São Paulo: USP, 26 jul. 2013. Disponível em: <https://conferencias.rcaap.pt/confoa2013/2013/paper/view/332>. Acesso em: 30 nov. 2023.

MARCONDES, Carlos Henrique; SAYÃO, Luis Fernando. Acesso unificado às teses eletrônicas brasileiras. **Informação & Sociedade: Estudos**, João Pessoa, v. 13, n. 1, p. 159-193, jan./jun. 2003. Disponível em: <https://periodicos.ufpb.br/ojs2/index.php/ies/article/view/125>. Acesso em: 30 nov. 2023.

MICROSOFT. Learn. DevOps. **O que é o Git?**. Publicado em: 5 out. 2023. Disponível em: <https://learn.microsoft.com/pt-br/devops/develop/git/what-is-git>. Acesso em: 29 abr. 2025.

MUELLER, Suzana Pinheiro Machado. A comunicação científica e o movimento de acesso livre ao conhecimento. **Ciência da Informação**, Brasília, v. 35, n. 2, p. 27-38, maio/ago. 2006. DOI: <https://doi.org/10.18225/ci.inf.v35i2.1138>. Disponível em: <https://revista.ibict.br/ciinf/article/view/1138>. Acesso em: 30 nov. 2023.

O'REILLY, Tim. What is Web 2.0: design patterns and business models for the next generation of software. **O'Reilly Media**. Publicado em: 30 set. 2005. Disponível em: <https://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html?page=1>. Acesso em: 23 abr. 2025.

PINHEIRO, Lena Vania Ribeiro. Do acesso livre à ciência aberta: conceitos e implicações na comunicação científica. **Revista Eletrônica de Comunicação, Informação & Inovação em Saúde**, [s. l.], v. 8, n. 2, p. 153-165, jun. 2014. DOI: <https://doi.org/10.3395/reciis.v8i2.629>. Disponível em: <https://www.reciis.icict.fiocruz.br/index.php/reciis/article/view/629>. Acesso em: 30 nov. 2023.

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: <https://doi.org/10.5433/1981-8920.2016v21n2p90>. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/27939>. Acesso em: 30 nov. 2023.

SHINTAKU, Milton; VIDOTTI, Silvana Aparecida Borsetti Gregorio. Bibliotecas e repositórios no processo de publicação digital. **Biblos**, [s. l.], v. 30, n. 1, p. 60-79, 2016. Disponível em: <https://periodicos.furg.br/biblos/article/view/5762>. Acesso em: 30 nov. 2023.

SILVEIRA, Lúcia da; RIBEIRO, Nivaldo Calixto; MELERO, Remedios; MORA-CAMPOS, Andrea; PIRAQUIVE-PIRAQUIVE, Daniel Fernando; URIBE-TIRADO, Alejandro; SENA, Priscila Machado Borges; POLANCO CORTÉS, Jorge; FACHIN, Juliana; SANTILLÁN-ALDANA, Julio; SILVA, Fabiano Couto Corrêa da; ARAÚJO, Ronaldo Ferreira; ENCISO-BETANCOURT, Andrés Mauricio; FACHIN, Juliana. Taxonomia da Ciência Aberta: revisada e ampliada. **Encontros Bibli**, Florianópolis, v. 28, artigo e91712, 30 jun. 2023. DOI: <https://doi.org/10.5007/1518-2924.2023.e91712>. Dis-

ponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/91712>. Acesso em: 30 nov. 2023.

SULEMAN, Hussein; FOX, Edward A. Towards universal accessibility of ETDs: building the NDLTD union archive. *In*: INTERNATIONAL SYMPOSIUM ON ELECTRONIC THESES AND DISSERTATIONS, 5., 2002. **Proceedings** [...]. Utah, EUA: ETD, 2002. Disponível em: http://www.husseinsspace.com/research/publications/etd_2002_paper_union.pdf. Acesso em: 30 nov. 2023.

SULEMAN, Hussein. **Open digital libraries**. 2002. Tese (Doutorado em Computer Science and Applications) – Virginia Polytechnic Institute and State University, Blacksburg, EUA, 2002. Disponível em: <http://hdl.handle.net/10919/29712>. Acesso em: 30 nov. 2023.

WEITZEL, Simone da Rocha. O papel dos repositórios institucionais e temáticos na estrutura da produção científica. **Em questão**, Porto Alegre, v. 12, n. 1, p. 51-71, jan./jun. 2006. Disponível em: <https://seer.ufrgs.br/EmQuestao/article/view/19>. Acesso em: 28 abr. 2025.

Como citar este capítulo:

RUFINO, Fernanda Maciel; VECHI, Bernardo Dionízio; SHINTAKU, Milton. Repositórios digitais abertos. *In*: ARAÚJO, Paula Carina de; LIMA, Karolayne Costa Rodrigues de (org.). **Práticas de ciência aberta**. Brasília, DF: Editora Ibict, 2025. Cap. 12, p. 252-268. DOI: 10.22477/9788570131966.cap12.