



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE ARTES E COMUNICAÇÃO
DEPARTAMENTO DE CIÊNCIA DA INFORMAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

EDUARDA FERREIRA FIGUEIRÊDO

REPOSITÓRIOS DIGITAIS DE PUBLICAÇÕES AMPLIADAS: Um estudo de caso

Recife
2018

EDUARDA FERREIRA FIGUEIRÊDO

REPOSITÓRIOS DIGITAIS DE PUBLICAÇÕES AMPLIADAS: Um estudo de caso

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de Mestre em Ciência da Informação.

Área de concentração: Informação, Memória e Tecnologia

Orientadora: Profa. Dra. Sandra de Albuquerque Siebra.

Recife
2018

Catálogo na fonte
Bibliotecária Jéssica Pereira de Oliveira, CRB-4/2223

F475r

Figueirêdo, Eduarda Ferreira

Repositórios Digitais de publicações ampliadas: um estudo de caso /
Eduarda Ferreira Figueirêdo. – Recife, 2018.

68f.: il.

Orientadora: Sandra de Albuquerque Siebra.

Dissertação (Mestrado) – Universidade Federal de Pernambuco.
Centro de Artes e Comunicação. Programa de Pós-Graduação em Ciência
da Informação, 2018.

Inclui referências.

1. Repositórios Digitais. 2. Publicações ampliadas. 3. Repositórios de
dados. 4. Acesso aberto. 5. Dados abertos. I. Siebra, Sandra de
Albuquerque (Orientadora). II. Título.

020 CDD (22. ed.)

UFPE (CAC 2019-39)

EDUARDA FERREIRA FIGUEIRÊDO

REPOSITÓRIOS DIGITAIS DE PUBLICAÇÕES AMPLIADAS: Um estudo de caso

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de mestre em Ciência da Informação.

Aprovada em: 27/02/2018

BANCA EXAMINADORA

Prof^a D^{ra} Sandra de Albuquerque Siebra (Orientadora)

Universidade Federal de Pernambuco

Prof. Dr. Marcos Galindo Lima (Examinador Interno)

Universidade Federal de Pernambuco

Prof. Dr. Henry Poncio Cruz de Oliveira (Examinador Externo)

Universidade Federal da Paraíba

Prof^a D^{ra} Celly de Brito Lima (Examinador Externo)

Universidade Federal de Pernambuco

AGRADECIMENTOS

Primeiramente, agradeço a Deus por essa oportunidade e por me dar forças e permitir que eu chegasse até aqui.

À minha mãe, Rizonete Gonzaga, por sempre me motivar e mover céus e terras para que eu pudesse ter uma boa oportunidade de educação. E aguentar meu humor em todas essas noites mal dormidas.

À minha orientadora, Sandra Siebra, por não ter desistido de mim, mesmo nos momentos mais difíceis que enfrentei para conseguir dar prosseguimento à minha pesquisa. Por sempre acreditar que eu era capaz e pela motivação para nunca desistir.

Aos meus amigos, Marília Santana, Marina de Holanda, Larissa Guimarães, João Marcelo, Andrey Reis, Gerlaine Souza, Rafael Ferreira, Fabíola Queiroz, Tatianna Dias, Fernando Rafael, pela grande força e motivação em todos os momentos que eu pensei em desistir e eles me fizeram ver que eu era capaz e que conseguiria concluir. Por terem entendido um pouco a minha distância na vida social e toda torcida durante esses dois anos.

À todos os meus colegas da turma de 2016 do Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Pernambuco, pela preocupação, companheirismo e motivação que foi dada desde o início dessa caminhada juntos.

Aos docentes do PPGCI por fazerem parte do meu amadurecimento como estudante e pesquisadora, e por todo aprendizado adquirido ao longo dessa jornada.

A todos, meu muito obrigado.

RESUMO

Devido à explosão da geração de dados no contexto da e-Science e visando a disponibilização livre dos dados gerados pelas pesquisas científicas para o avanço da ciência, tornou-se essencial ampliar o conceito da publicação armazenada nos repositórios digitais. A ideia das publicações ampliadas é criar pontes que liguem os conteúdos das publicações dos repositórios digitais a conteúdos de repositórios de dados. Nesse cenário, essa pesquisa objetiva investigar a disponibilização de publicações ampliadas em repositórios digitais institucionais brasileiros. Essa é uma pesquisa descritiva, qualitativa, que teve como objeto de estudo os repositórios CarpeDien do Instituto de Energia Nuclear (IEN) e o RIDI do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT). Os resultados mostram que, apesar da tendência que se fortalece cada vez mais no contexto da ciência aberta de se criarem repositórios de dados abertos de pesquisa e integrá-los a repositórios digitais ou incorporá-los a eles, especialmente no exterior, esses repositórios ainda não existem efetivamente no Brasil. Observa-se algumas iniciativas em prol de uma futura integração, mas ainda de maneira limitada, como também algumas iniciativas embrionárias de repositórios de dados dentro de instituições acadêmicas tais como o protótipo de Repositório de Dados de Pesquisa da UNESP e a Base de Dados Científicos da Universidade Federal do Paraná. Cenário esse que deve se modificar em alguns anos, devido às exigências crescentes das instituições de fomento a pesquisa pela criação e execução de planos de gestão de dados de pesquisa. O que deve motivar a criação e expansão de repositórios de dados, assim como o uso de publicações ampliadas.

Palavras-chave: Repositórios Digitais. Publicações Ampliadas. Repositórios de Dados. Acesso Aberto. Dados Abertos.

ABSTRACT

Due to the explosion of data generation in the context of e-science and to the free availability of data generated by scientific research for the advancement of science, it has become essential to broaden the concept of the publication stored in digital repositories. The idea of expanded publications is to create bridges linking the content of publications in digital repositories to the contents of data repositories. In this scenario, this research aims to investigate the availability of extended publications in Brazilian institutional digital repositories. This is a descriptive, qualitative research that had as object of study the CarpeDien repositories of the Nuclear Energy Institute (IEN) and the RIDI of the Brazilian Institute of Information in Science and Technology (IBICT). The results show that, despite the tendency that is increasingly strengthened in the context of the open science of creating repositories of open research data and integrating them with digital repositories or incorporating them, especially abroad, these repositories do not yet exist in Brazil. There are some initiatives for future integration, but still limited and some embryonic initiatives within academic institutions as the Repositório de Dados de Pesquisa of UNESP and the Base de Dados de Pesquisa da Federal University of Paraná. This scenario is expected to change over the next few years due to the increasing demands of research institutions for the creation and execution of research data management plans. This should motivate the creation and expansion of data repositories as well as extended publications.

Keywords: Digital repositories. Expanded Publications. Data Repositories. Open Access. Open Data.

LISTA DE FIGURAS

Figura 1 - Elementos do modelo de dados de publicações ampliadas	39
Figura 2 - Página inicial do repositório do IBICT	48
Figura 3 - URI de Recurso no Repositório RIDI.....	49
Figura 4 - Imagem dos Arquivos associados do site do Ibict.....	50
Figura 5 – Uso de URI externa no repositório RIDI	51
Figura 6 - Página do Repositório CarpeDien	53
Figura 7 - Busca Facetada no site do repositório CarpeDien	53
Figura 8 - URI de Recurso do Repositório CarpeDien	54
Figura 9 - Associação de arquivos complementares no CarpeDien	55
Figura 10 - Uso de URI externa no repositório CarpeDien.....	56

LISTA DE QUADROS

Quadro 1 – Iniciativas que recomendam a realização de auditorias	36
Quadro 2 - Análise do repositório do RIDI.....	49
Quadro 3 - Análise do repositório do CarpeDien.....	54

SUMÁRIO

1	INTRODUÇÃO.....	10
2	MEMÓRIA E SUPORTES DE MEMÓRIA.....	16
2.1	MEMÓRIA CIENTÍFICA.....	18
2.2	SUPORTES E FORMAS DE REGISTRO DA MEMÓRIA	19
3	OS REPOSITÓRIOS DIGITAIS E A CIÊNCIA ABERTA.....	23
3.1	REPOSITÓRIOS DIGITAIS	24
3.2	CLASSIFICAÇÃO DOS REPOSITÓRIOS	27
3.3	PLATAFORMAS/SOFTWARES PARA REPOSITÓRIOS.....	28
3.4	E-SCIENCE E REPOSITÓRIOS DE DADOS	30
4	PUBLICAÇÕES AMPLIADAS.....	37
4.1	ELEMENTOS DA PUBLICAÇÃO AMPLIADA.....	39
4.2	REQUISITOS E RECOMENDAÇÕES PARA A PUBLICAÇÃO AMPLIADA	41
4.3	SISTEMAS DE INFORMAÇÃO PARA PUBLICAÇÃO AMPLIADA	42
4.3.1	Sistemas de Informação para Interconexão com Dados de Pesquisa	43
5	METODOLOGIA	45
6	ANÁLISE DOS REPOSITÓRIOS DE PUBLICAÇÕES AMPLIADAS	47
6.1	REPOSITÓRIO INSTITUCIONAL DO INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT).....	47
6.2	REPOSITÓRIO CARPEDIEN: DADOS E INFORMAÇÕES EM ENERGIA NUCLEAR – INSTITUTO DE ENGENHARIA NUCLEAR.	52
7	CONSIDERAÇÕES FINAIS.....	57
	REFERÊNCIAS.....	59

1 INTRODUÇÃO

Nas últimas décadas a produção científica cresceu substancialmente e esse crescimento tem contribuído para o desenvolvimento da ciência e possibilitado a disseminação do conhecimento nas mais diversas áreas científicas. De fato, a ciência está sempre em evolução, pois a todo tempo novas pesquisas e descobertas são feitas e publicadas ao redor do mundo.

A modificação da forma de intercâmbio de informações nas comunidades científicas e no processo de comunicação científica em si, em função do uso de computadores, *tablets*, *smartphones* e internet, têm apontado para mudanças paradigmáticas na área, de onde se refletem novas possibilidades de divulgação de informações científicas e de preservação da memória da ciência. De fato, a comunidade científica é influenciada pelos paradigmas, teorias e tecnologias vigentes. Sobre o contexto paradigmático, Nehmy e colaboradores (1996), utilizam Thomas Khun (1994) para definir paradigma como um consenso, que representa um momento histórico, junto com suas teorias e métodos, que lidam com determinados tipos de problemas, ou anomalias. Quando tudo que é estudado dentro daquele âmbito paradigmático não consegue mais suprir as necessidades referentes aos problemas da época, a ciência passa por um momento de crise e, em resposta a essa crise, é preciso encontrar novas teorias e formas de resolver aquele problema, dessa forma nasce outro paradigma para suprir as lacunas existentes.

A cerca de mil anos havia apenas a ciência empírica, cuja forma de execução estava pautada na descrição de fenômenos naturais. Posteriormente, a ciência teórica formulou leis e equações de grande abrangência, por meio do uso de modelos e generalizações. Nos anos subsequentes, a dificuldade de resolução dos modelos teóricos e empíricos cresceu e, como consequência, nas últimas décadas, o modo de fazer ciência passou a ser o da simulação via computadores. Este terceiro modelo promoveu um crescimento exponencial dos dados científicos, trazendo à tona, no meio da ciência, um quarto paradigma, chamado de e-Science (HEY; TANSLEY; TOLLE, 2009). Esse quarto paradigma objetiva a exploração e a reutilização de dados e a unificação das teorias, experimentos e simulações utilizando metodologias e ferramentas voltadas para o gerenciamento e reuso desses dados de forma interdisciplinar e inovadora, alterando a forma como a ciência evolui (SALES, 2014). Pois, os conhecimentos científicos registrados se convertem em informações passíveis de consulta para geração de novos conhecimentos. De acordo com Roland (2006), o princípio basilar da ciência acadêmica é que os resultados de pesquisa devem ser públicos e os conhecimentos, dados, descobertas e experimentos não podem ser considerados como pertencentes ao

conhecimento científico, se não forem relatados e armazenados de forma permanente. Como consequência, tornou-se ainda mais necessário formas de organização, armazenamento, gestão e disponibilização dos dados e informações, de forma a criar uma memória para disponibilização para as gerações atuais e futuras, possibilitando o acesso a partir de qualquer lugar e a qualquer momento.

Nesse novo cenário, se tornou ainda mais mandatório a divulgação ampla dos avanços e descobertas das Ciências, a necessidade de compartilhar ideias, resultados e questionamentos. Como uma primeira solução, foram criadas as revistas científicas (primeiro as analógicas e, posteriormente, as eletrônicas), uma plataforma para divulgar as pesquisas e seus resultados para o restante da sociedade acadêmica ou não, tornando-se o principal marco da estrutura da comunicação científica. (WEITZEL, 2006). A utilização dessas revistas, trouxe um outro patamar para a comunicação científica, pois elas conseguiam atingir um nível maior de leitores geograficamente dispersos. Porém, mesmo com a utilização dessas revistas, alguns problemas foram identificados com o passar do tempo. Primeiramente, o tempo para a publicação de pesquisas nessas revistas era muito longo, pois era preciso passar por muitas etapas e revisões até que as pesquisas pudessem ser publicadas. Isso ocasionava que algumas publicações acabavam por estar com informações obsoletas no momento da publicação, já existindo dados/informações mais novos sendo pesquisados (WEITZEL, 2006). Esse foi um problema não resolvido ainda nos dias atuais, visto que as publicações ainda podem demorar para serem realizadas, pois dependem de avaliações por pares, que nem sempre são obtidas rapidamente. Outra questão era a exclusividade requerida pelas editoras das revistas, tornando a produção do pesquisador restrita aos assinantes da revista ou a compras individuais de seu artigo. Assim, o acesso democrático ao conhecimento produzido pelas instituições científicas se tornou complexo, em função do alto custo de produção e aquisição dos periódicos onde, em geral, o conhecimento produzido era e ainda é divulgado. Pois as editoras cobravam/cobram altos valores para vender suas assinaturas às bibliotecas e instituições que desejam acesso aos materiais por elas disponibilizados (BAPTISTA et al., 2007), mesmo os materiais tendo sido produzido pelos pesquisadores dessas instituições com o apoio de recursos públicos. Nesse sentido, as editoras universitárias são as que possuem maior abertura para o acesso aberto, pois recebem incentivo de algumas unidades de fomento que financiam e estimulam projetos de pesquisa. Some-se a isso que, com o desenvolvimento da ciência e o e-Science, a questão da publicação das pesquisas, ultrapassou as possibilidades de publicação que os periódicos podiam oferecer.

Considerando as dificuldades anteriormente citadas e as facilidades trazidas pelas TICs para o acesso à comunicação científica, surgiu o movimento de acesso aberto (MUELLER, 2006), com a proposta de acesso livre (*open access*), como também a proposta de arquivos abertos¹ (*open archives*), que buscam a disseminação do conhecimento científico, representados pelas *Open Archives Initiative* – OIA - e a *Budapest Open Access Initiative* – BOIA (LIMA, 2009). O movimento de acesso aberto e livre tem como objetivo desenvolver e promover a implantação e a disseminação dos conteúdos dos arquivos abertos (CAFÉ; LAGE, 2002). Muitas instituições e pesquisadores passaram a aderir a esse movimento preocupados com a disseminação dos resultados das pesquisas realizadas (KURAMOTO, 2014). Concomitante com isso, a comunidade científica/acadêmica, inserida na dinâmica do movimento do acesso livre à informação, começou a utilizar os repositórios digitais como uma ferramenta de auxílio para a publicação e disseminação de pesquisas científicas. Repositórios digitais são uma forma de armazenamento de objetos digitais que tem a capacidade de manter e gerenciar material por longos períodos de tempo e prover o acesso apropriado (VIANA; ARELLANO, 2006). Eles incentivam e gerenciam a publicação pelo pesquisador, utilizam tecnologia aberta e podem ser acessados por diversos provedores de serviços nacionais e internacionais. Por isso, constroem as condições necessárias para permitir o acesso livre à produção científica de forma legítima, como também são uma alternativa para guarda e visualização da memória de uma instituição, alterando o processo de produção, disseminação e uso da informação (BLATTMAN; BOMFÁ, 2006).

Em 2003, a Declaração de Berlin sobre o Acesso Aberto ao Conhecimento em Ciências e Humanidades disponibilizada pela Universidade do Minho, explicitou que se compreende por contribuições de acesso livre os “resultados de pesquisas científicas originais, dados brutos [dados não processados] e metadados, fontes originais, representações digitais de materiais pictóricos e gráficos além de material acadêmico multimídia”. Com isso, ampliou-se as fronteiras do movimento de acesso livre e o conceito deixou de estar relacionado apenas às publicações acadêmicas tradicionais, como artigos de periódicos, avançando para a disponibilização aberta e de forma inteligível de outras formas de conteúdo, como os dados de pesquisa (SAYÃO; SALES, 2014). Essa ampliação tem implicações, visto que o volume de dados de pesquisas disponibilizados digitalmente vem crescendo a uma velocidade exponencial, requerendo concepções novas de documentos e redesenhando o ciclo tradicional

¹ Arquivos *on-line* de acesso público, também definidos como “diretórios existentes em um computador que estão abertos para o acesso via *ftp* ou *http*, armazenando uma coleção de séries de artigos ou uma coleção de dados sobre artigos armazenados em outro local (SENA, 2000).

de comunicação científica, levando em conta também a característica da colaboração e compartilhamento de dados e informações entre pesquisadores, sem limitações de tempo e espaço (ABBOT, 2008; SAYÃO; SALES, 2014). Porém, a diversidade, heterogeneidade e a complexidade dos registros de resultados de pesquisa não conseguem ser expressos por documentos convencionais únicos, impressos ou digitais. De fato, dados e informações gerados pelas atividades de pesquisa necessitam de novos modelos de representação, de custódia e de gestão de conteúdos científicos digitais.

Assim, surgiram propostas sobre como representar esses novos dados, tais como: pacotes de publicações científicas (chamadas publicações ampliadas) que encapsulam e relacionam dados brutos e seus subprodutos, publicações e metadados em objetos compostos (HUNTER, 2007); documentos sobrepostos (*overlay documents*) que interligam dados, anotações, artigos de periódicos revisados por pares e comentários (HEY, 2009) e objetos digitais compostos que combinam recursos distribuídos com dados e tipos múltiplos de mídia (texto, imagens e vídeos) (LAGOZE; SOMPEL, 2008). Os dados de pesquisa nessas representações vem sendo armazenados nos chamados repositórios de dados confiáveis² (SAYÃO; SALES, 2012), onde são gerenciados e preservados sob os princípios da curadoria digital, mantendo sua capacidade de reuso.

Apesar dessas novas representações, na atual infraestrutura de comunicação científica, esses objetos compostos não são linkados às publicações onde eles são analisados e discutidos. O que é justamente a ideia por trás das publicações ampliadas, onde as publicações científicas podem ser ligadas aos conteúdos de repositórios de dados confiáveis relacionados (VERHAAR, 2008; SALES; SAYÃO; SOUZA, 2013). Por exemplo, uma tese de doutorado poderia ser linkada com os dados científicos que foram utilizados para produzi-la. Em outras palavras, publicações ampliadas são “objetos digitais compostos que combinam e-Prints com um ou mais recursos de dados, um ou mais registros de metadados, ou qualquer combinação deles” (VERHAAR, 2008, p.11). Essa combinação poderia melhorar a produtividade científica, por meio do aumento da velocidade com que os dados e as informações poderiam ser tratados e produzidos (LYNCH, 2011), além de evitar que sejam despendidos esforços e recursos em coletas de dados duplicadas, reduzindo os custos e investimentos aplicados pelas agências de fomento (MONTEIRO, SANT’ANA, 2016). Isso se faz relevante porque os resultados de pesquisas alcançados por alguns pesquisadores vem sendo “freqüentemente

² Repositórios Digitais confiáveis são repositórios que têm como “missão oferecer à sua comunidade-alvo acesso confiável e de longo prazo aos recursos digitais por ela gerenciados, agora e no futuro” (RLG/OCLC, 2002, p.5).

retomados por outros cientistas, teóricos ou aplicadores, que dão continuidade ao estudo, fazendo avançar a ciência ou produzindo tecnologias ou produtos neles baseados” (MUELLER, 2000, p. 25).

Nesse cenário, se tem a necessidade de repositórios digitais desenvolvidos para armazenar, além das publicações científicas, os dados de pesquisa relacionados a essas publicações. Repositórios que consigam integrar dados de toda natureza e publicações tradicionais em formatos digitais, criando um novo gênero de publicação Web, chamado publicações ampliadas (SALES, 2014). Nessa conjuntura esta pesquisa tem como objetivo geral investigar o uso de publicações ampliadas em repositórios digitais institucionais brasileiros. E como objetivos específicos:

- Realizar levantamento bibliográfico sobre repositórios, publicações ampliadas e repositórios de dados confiáveis;
- Investigar os repositórios institucionais de instituições, institutos e/ou centro de pesquisas cadastrados no site do Opendoar para verificar a utilização feita de publicações ampliadas, ou se há interligação de publicações com dados de repositórios de dados.
- Buscar no referencial teórico repositórios institucionais referenciados como sendo de publicações ampliadas;
- Analisar os repositórios institucionais que indicam fazer uso de publicações ampliadas.

Essa pesquisa se enquadra no chamado quarto paradigma científico³, que tem como foco a exploração de dados e a unificação das teorias, experimentos e simulações utilizando metodologias e ferramentas voltadas para o gerenciamento e reuso desses dados (SALES, 2014). E, devido a quantidade cada vez maior de dados gerados pelos pesquisadores nas instituições e a possibilidade de reutilizá-los de forma interdisciplinar e inovadora, fica evidente que a adição de outros recursos ao texto, como dados científicos, imagens, sons e interatividade, agora se torna fundamental. De fato, os princípios da ciência aberta estão embasados no argumento que os dados podem ter vários usos e aplicações que podem ser, inclusive, diferentes daqueles para os quais foram inicialmente criados (EVANS, REIMER, 2009; VISION, 2010), ampliando as chances de serem obtidos novos resultados e conhecimentos científicos, decorrentes dos dados previamente disponibilizados (WALLIS;

³ Apesar de enquadrar a pesquisa no quarto paradigma (considerando o contexto internacional), tem-se ciência de que este paradigma ainda não está consolidado no Brasil.

ROLANDO; BORGMAN, 2013). Além disso, considera-se a premissa de que os resultados de pesquisas realizadas com financiamento público são bens públicos tangíveis e, por isso, devem estar disponíveis sem custo para reuso (DALRYMPLE, 2003).

Adicionalmente, ressalta-se que essa pesquisa considera que, promover a ligação semântica entre a produção científica e dados de pesquisa, facilita a visualização da pesquisa de forma global. E que integrar os dados brutos e/ou processados e os documentos correlatos aos resultados publicados é uma forma de registrar a memória da pesquisa (memória científica) desenvolvida pela instituição, dando visibilidade a mesma e abrindo possibilidades de colaboração científica. Dessa forma, justifica-se a realização da pesquisa, uma vez que a Ciência da Informação tem na memória e na comunicação científica (e no fluxo desta), objetos de estudo constante (SOUZA, 2012), trazendo contribuições nas formas de comunicação da informação e de registro e acesso à memória científica, colaborando assim com os anseios por informação da sociedade científica

A partir desse ponto, o trabalho se subdivide da seguinte forma:

No capítulo 2 será tratada a temática Memória, abordando conceitos, seus suportes e formas de registro.

No capítulo 3 serão apresentados os Repositórios Digitais, abordando suas classificações, formas de armazenamento, as plataformas mais utilizadas para a construção dos repositórios, além do campo da ciência aberta e seus componentes.

No capítulo 4 serão discutidas as publicações ampliadas, trazendo seus elementos, os requisitos mínimos para um repositório ser considerado ampliado, e os sistemas de informação que são utilizados nesse tipo de repositório.

No capítulo 5 será abordada a metodologia utilizada para a produção dessa pesquisa.

No capítulo 6 serão analisados dois repositórios brasileiros: o do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e o Repositório do Instituto de Engenharia Nuclear (CarpeDien), a partir de requisitos previamente apresentados e os resultados serão comparados.

E, por fim, serão feitas as considerações finais baseadas na análise de todos os pontos abordados anteriormente.

2 MEMÓRIA E SUPORTES DE MEMÓRIA

Memória “[...] envolve vários campos do saber, uma vez que a memória, como propriedade de conservar certas informações, nos remete a um conjunto de ações psíquicas, com as quais atualizamos impressões passadas” (DODEBEI, 1997, p. 44). O conceito de memória é uma maneira de pensar o passado em função do futuro que se deseja. Dessa forma, a pretensão em manter viva a memória surge do anseio de se manter vivas as lembranças, ideias, conhecimentos e ideias para que estejam ao alcance no instante em que for preciso resgatá-los (GONDAR, 2009). O que é endossado por Monteiro, Carelli e Pickler (2008, p. 14) quando frisam que “[...] a própria noção de memória consiste em uma analogia da memória humana, responsável por reter informações na mente e recuperá-las quando necessário”. Pois, “a memória não só não se apresenta como um mero mecanismo que copia informação e a armazena na nossa cabeça, mas também como a experiência de recuperar essa informação e de combinar de maneira a formar pensamentos novos (FENTRESS; WICKMAN, 1994, p. 29). Dessa maneira, pode-se observar que os autores abordam a memória como uma forma de trazer acontecimentos do passado, vistos e analisados no contexto do presente, para suprir necessidades do presente e, em alguns casos, produzir novas informações, já que é por meio da memória de um conhecimento ou um raciocínio antigo que pode nascer uma nova reflexão (DODEBEI, 1997).

Na tradição filosófica, e também no modo de pensar comum, a memória parece referir-se a uma persistência, a uma realidade de alguma forma intacta e contínua; a reminiscência (ou anamnese ou reevocação), pelo contrário, remete à capacidade de recuperar algo que se possuía antes e que foi esquecido (ROSSI, 2010, p. 15, grifo nosso).

Nora (1993) ressalta que a memória está sempre em evolução e está aberta a dialética do esquecimento e da lembrança. Destaca-se que, a princípio, o ser humano é orientado ao esquecimento, por ser este instintivo e espontâneo. Porém, a necessidade de pertencimento a um grupo sobrepõe o instinto primitivo e acarreta no estabelecimento de relações sociais, na necessidade de controlar o esquecimento para comprometimento com o futuro e, por conseguinte, do registro da memória para interesses coletivos (SOUZA, 2012). Assim, a lembrança e o esquecimento são componentes da memória, um não existe sem o outro, no processo de atualização do passado, quando evocado.

Jo Gondar (2016) apresenta a memória como polissêmica e um fenômeno em constante mutação e, dessa maneira, não é possível formular um conceito único para ela.

Um conceito costuma nos dizer o que alguma coisa é, no presente, no passado e no futuro, a despeito de qualquer mudança. A memória, contudo, nunca é: na variedade de seus processos de conservação e transformação, ela não se deixa aprisionar numa forma fixa ou estável. A memória é, simultaneamente, acúmulo e perda, arquivo e restos, lembrança e esquecimento (GONDAR, 2016, p.19).

Em alguns contextos os conceitos de memória e história são confundidos, a partir do momento da sua transcrição. Le Goff (2003, p. 477) relaciona memória e história, quando afirma “A memória onde cresce a história, que por sua vez a alimenta, procura salvar o passado para servir o presente e futuro”. Porém, Nora (1993) afirma que a memória é viva, é tratada como um fenômeno em constante evolução e vulnerável a usos e manipulações, nem sempre construída intencionalmente. Já a história é a representação do passado, “a reconstrução sempre problemática e incompleta do que não existe mais” (p. 9), não necessariamente relacionada com o presente. Assim, é possível afirmar que a pretensão de se manter viva a memória, surgiu da vontade de se manterem vivas as recordações, ideias, conhecimentos, culturas e tradições, de forma que fosse possível resgatá-las, quando fossem necessárias.

Segundo Le Goff (1990), a memória pode ser tratada de duas maneiras, memória individual e coletiva. A primeira, a individual, repleta de afetividade e manipulável pelo seu transmissor por meio de emoções e intenções, e a segunda, a coletiva, relacionada a junção da busca coletiva por um mesmo objetivo. O que diverge da visão de Halbwachs (1993) para o qual a memória é sempre construída de forma coletiva. Porém, o mesmo autor afirma que a memória individual é definida pela junção de diferentes influências sociais, e a coletiva, é um trabalho social feito a partir de um grupo. De fato, apesar da memória possuir uma dimensão individual, a maioria de seus referenciais são sociais e são eles que permitem que, além da memória individual, que é por definição única, tenha-se também uma memória intersubjetiva, compartilhada e coletiva. Assim, Halbwachs afirma que a memória coletiva envolve as memórias individuais, mas não se confunde com elas e evolui de acordo com suas próprias leis. Logo, "se algumas lembranças individuais penetram algumas vezes nela, mudam de figura assim que sejam recolocadas num conjunto que não é mais [aquele da] consciência pessoal" (HALBWACHS, 1990, p. 53-54). O que é endossado por Le Goff (1990), quando afirma que a memória é algo que os próprios homens estabelecem a partir de suas relações sociais. Dessa forma, o homem é produtor e produto de seu quadro social e cultural, ao mesmo tempo, pois ele não é um ser isolado e sim interativo com a sociedade, então é

formador da memória como construção social. Logo, o conhecimento que constrói a memória coletiva se torna inteligível quando o homem passa a ter consciência que a memória é algo que eles mesmos estabelecem a partir das relações sociais (LE GOFF, 1990).

O conceito de memória social, para Le Goff (1990), é associado a documentos escritos, uma vez que, para o autor, esse conceito é aplicado às sociedades onde a escrita já tenha se instalado, enquanto a memória coletiva diz respeito aos povos sem escrita.

No contexto dessa pesquisa a memória científica é social, uma vez que é originária de um grupo (comunidade científica) e, para que não pereça, ela precisa estar acessível em suportes de memória, porque enquanto produto da ciência legitimada, é vital que esteja disponível a longo prazo para os pesquisadores.

2.1 MEMÓRIA CIENTÍFICA

Monteiro, Carelli e Pickler (2006 p. 115, grifo nosso) afirmam que o termo memória,

foi associado, especialmente na arquivologia, biblioteconomia e museologia, ao conjunto das informações registradas, isto é, aos documentos e representações que podem ser consultados, servindo de memória social ou memória de longo prazo. Com efeito, essas três áreas valem-se da memória no sentido de armazenagem e preservação dos saberes (conservação), para a posterior recordação por parte da sociedade. (MONTEIRO, CARELLI E PICKLER, 2006, p.115)

Nesse contexto, pode-se considerar memória científica a documentação produzida e acumulada no decorrer das atividades científicas e daquelas que as viabilizam e concorrem para seu desenvolvimento, difusão e acesso. Também os arquivos gerados no desenvolvimento de atividades que possibilitam e viabilizam o fazer científico e naqueles decorrentes de atividades que são fundamentais à atividade científica, como a difusão da ciência realizada em revistas especializadas e em grandes bancos de dados que referenciam a produção científica nacional e internacional, além de patentes portadoras das inovações tecnológicas (BRITO, 2003). Estão aí também incluídos, os museus científicos e as grandes redes de comunicação, impressa, televisiva e eletrônica. Pois, o caráter cumulativo da ciência focada nas contribuições dos pesquisadores gera um corpo de conhecimento baseado no consenso, que é representado pela literatura técnico-científica, produto sujeito à mensuração da atividade científica (SAYÃO, 1996). Sendo assim, “a memória científica encontra-se nas instituições onde a ciência é feita, ensinada, planejada, administrada, financiada, exposta, difundida e acessada” (BRITO, 2003, p. 2).

De fato, pode-se afirmar que tanto a produção de pesquisas, quanto a informação científica⁴ compõem o acervo da memória social da comunidade científica, então se verte em uma memória própria a esse grupo: a memória científica (SOUZA, 2012). Para que não pereça, seu conteúdo precisa estar acessível em suportes e formatos e registro da memória, porque enquanto produto da ciência, é vital que esteja disponível a longo prazo para os pesquisadores. Pois,

Quando um pesquisador, diante de um microcomputador ligado a um banco de dados que pode estar em qualquer parte do mundo, vasculha suas estantes eletrônicas à procura de informações que definam, completem ou estabeleçam as fronteiras do seu trabalho de pesquisa, ele repete o mesmo gesto de quem mergulha na memória de seu grupo para reconstruir as lembranças comuns e dessa forma manter íntegra a sua comunidade (SAYÃO, 1996, p. 314)

Ressalta-se que não apenas a guarda e a preservação documental são importantes, pois preservar é mais do que guardar, é tornar acessível à comunidade acadêmica e científica e à sociedade em geral informações que tragam conhecimento e novos saberes, de tal forma que a comunidade possa usufruir desses benefícios, gerando novas pesquisas (ABRAHÃO, 2010).

Dessa forma, o acesso aos documentos que registram o desenvolvimento da ciência tornou-se uma ferramenta importante no intercâmbio de informações entre as Instituições congêneres. Logo, é fundamental a sensibilização da comunidade de pesquisadores em relação ao estabelecimento de políticas para a preservação da memória científica e a consciência da necessidade de não só guardar os documentos, mas, também entender porque foram gerados e a sua importância na realização de futuras pesquisas (RIBEIRO, 2010). Pois, só assim, os pesquisadores poderão tirar proveito do que outros colegas já fizeram no passado para, a partir de ideias, dados e informações passadas, gerarem novas possibilidades e perspectivas. Porém, segundo Abrahão (2010), para que isso ocorra, é de suma importância que as informações científicas estejam disponíveis em suportes de memória, senão, pode haver perda do conhecimento.

2.2 SUPORTES E FORMAS DE REGISTRO DA MEMÓRIA

Ao longo da história humana, estoques informacionais foram formados e materializados em diferentes suportes e formas de registro, que vão da utilização da pedra,

⁴ “A informação científica é o insumo básico para o desenvolvimento científico e tecnológico de um país” (KURAMOTO, 2006, p. 91). Esse tipo de informação, resultado das pesquisas científicas, é divulgado à comunidade por meio do sistema de comunicação científica.

passando pelo uso do papiro, pergaminho, papel, aos mais modernos instrumentos tecnológicos. Pois, como nos mostra a História, construir formas de preservar a memória, o conhecimento sempre foi uma das grandes preocupações do homem. Assim, foram sendo produzidos acervos dos mais variados, por meio da escrita, sons, imagens, vídeos, entre outros modos de expressão e comunicação.

Em um breve retrospecto, primeiramente a memória era armazenada apenas por meio de pessoas, onde homens importantes, que vivenciavam acontecimentos, eram chamados de “homens-memórias” e reuniam todos as ocorrências relevantes da época ou as informações e episódios cotidianos considerados úteis eram passados de um indivíduo para os seus descendentes. Por conta disso, a transmissão dessas ocorrências era feita apenas de forma oral, sendo assim de fácil manipulação, pois o controle total estava sob os guardiões da memória (LE GOFF, 1990).

A escrita veio como uma forma de transformação da memória, primeiramente na forma de desenhos e símbolos, de maneira a auxiliar na disseminação e guarda das informações, onde assim podia-se alcançar mais gerações, pois antes eram apenas compartilhadas de forma oral e com a visão dos fatos do transmissor. Com o advento da escrita, a memória se exterioriza e passa a permitir em um suporte externo à mente e a voz o registro da informação produzida pelo ser humano, possibilitando, assim, a multiplicação da informação, sua perdurabilidade e seu uso repetido e dilatado no tempo (LE COADIC, 2004; RIBEIRO, 2005). De fato, os documentos escritos dão um suporte material à memória, de forma a ampliá-la e transformá-la, além de estabelecer a fronteira onde a memória coletiva torna-se social.

Além disso, a transformação da oralidade para a escrita foi um marco crucial na evolução da memória, pois as informações transmitidas passaram a ser mais precisas e sua transmissão atingiu uma forma mais fiel ao momento referenciado. Assim, pode-se dizer que a escrita surgiu para auxiliar a sociedade na comunicação e armazenamento da memória. Os registros do conhecimento em suportes concederam à sociedade a capacidade de revelar a memória social com fins de preservação por tempo indeterminado.

Os meios de memória relacionados à oralidade e a escrita são tratados por Vera Dodebei (2009) como memória individual e memória auxiliar.

a memória individual que ao invés de acumular processa para produzir novas informações, e à memória auxiliar criada pela sociedade da escrita representada por arquivos, bibliotecas e museus, juntam-se a memória

informática que transforma o visual e o oral em números (DODEBEI, 2009, p.130).

Logo, no decorrer do tempo, os suportes de memória evoluíram acompanhando a transição humana da oralidade para a escrita. E, com o progresso da memória escrita, os suportes continuaram a se renovar até se converterem em acessíveis por meios eletrônicos, impulsionados pela tecnologia da informação e comunicação (TICs) e as redes como a internet. Os suportes de memória eletrônicos provocaram mudanças na comunicação científica uma vez que possibilitaram oferecer informação sem limitações temporais ou espaciais. Pois, na verdade, “A memória eletrônica radicaliza um traço intrínseco a todo suporte de memória: a capacidade de promover a desterritorialização dos eventos no próprio ato de sua recuperação” (FERREIRA; AMARAL, 2004, p. 141).

Em resumo, no contexto da evolução da memória há 5 momentos referentes ao salto que memória percorreu (LE GOFF, 1990).

- 1) Transmissão Oral
- 2) Transmissão escrita com tábuas ou índices
- 3) Fichas Simples
- 4) Mecanografia
- 5) Seriação Eletrônica (Banco de Dados)

Desde a oralidade até a memória eletrônica, o objetivo foi sempre o armazenamento e a transmissão de informações registradas no passado, para isso sendo necessário criar suportes e formas de registro da memória, além de ferramentas para seu acesso.

Os suportes e formas de armazenamentos foram se aperfeiçoando, sempre para alcançar a menor perda de informações possível, para permitir uma maior quantidade de informações armazenadas a um custo mais baixo, para aprimorar a forma de promover acesso e para se adequar às tecnologias existentes. Assim, a memória eletrônica passou a ser utilizada para suprir a necessidade de armazenamento de informações que a memória do ser humano não era mais capaz de armazenar de forma satisfatória. Como consequência, os computadores passaram a ser usados como dispositivos de memória auxiliar, colaborando no acesso e manutenção das informações produzidas pela sociedade (DODEBEI, 2016).

Paralelamente à evolução da memória, seus suportes e métodos de armazenamento, foram sendo criadas ferramentas para auxiliar na organização, guarda e acesso às informações armazenadas a longo prazo. Como exemplo, pode-se citar os repositórios digitais que são, de

acordo com Arellano (2008, p. 124), “um serviço de armazenamento de objetos digitais que tem a capacidade de manter e gerenciar materiais por longos períodos de tempo e prover o seu acesso apropriado”. Os repositórios e sua relação com a memória científica serão abordados no próximo capítulo.

3 OS REPOSITÓRIOS DIGITAIS E A CIÊNCIA ABERTA

Destaca-se que o princípio basilar da ciência acadêmica é que os resultados de pesquisas devem ser públicos. Assim, “as descobertas dos cientistas não podem ser consideradas como pertencentes ao conhecimento científico se não forem relatadas e gravadas de forma permanente. A instituição fundamental da ciência é, então, o sistema de comunicação” (ROLAND, 2006, p. 87). Uma das principais formas de divulgação/comunicação do conhecimento científico é viabilizada pela publicação científica, que é uma colaboração em larga escala, com a possibilidade de acesso à pesquisa e possível reprodução de resultados das pesquisas e experimentos (LYNCH, 2011).

A comunicação científica consiste em assegurar o intercâmbio de dados e informações sobre os trabalhos em andamento ou já finalizados e possibilita o contato entre pesquisadores e suas pesquisas sendo possível o trabalho colaborativo (LE COADIC, 2004). Ela está situada no coração da ciência (MEADOWS, 1999).

Como formas de comunicação científica, por muitos anos, destacaram-se as revistas científicas impressas, até que a crise dos periódicos científicos causada pelo alto valor de aquisição desses periódicos, no final da década de 90, fizeram surgir diversas manifestações em favor do acesso aberto ou acesso livre à informação científica⁵. Assim, foram procuradas alternativas no sentido de manter o acesso a essas revistas por meio da formação de consórcios, criando-se portais de acesso às revistas eletrônicas, o que ainda trazia dificuldade com os editores dos periódicos (KURAMOTO, 2006). Posteriormente, com o estabelecimento do modelo *Open Archives* e o desenvolvimento de diversas ferramentas de software para a construção de repositórios digitais e publicações periódicas eletrônicas em conformidade com esse modelo, o movimento do acesso livre ganhou consistência. Ressalta-se que o movimento do acesso livre teve como principal objetivo defender o acesso aberto às pesquisas científicas e à informação em meio digital, sem restrições ou cobranças, a fim de evitar o monopólio de editoras, que cobravam/cobram altos valores para vender suas assinaturas às bibliotecas e instituições que desejem acesso aos materiais por elas disponibilizados por deterem os direitos autorais de periódicos científicos e dos materiais neles publicados (BAPTISTA *et al.*, 2007).

Assim, os avanços das Tecnologias da Informação e Comunicação (TIC), em especial das redes de comunicação eletrônica, revolucionaram os fluxos de informação, as formas de acesso e a troca de informações, além de possibilitar o uso de novas formas de publicação,

⁵ O movimento se baseia no princípio de que todos os resultados de pesquisas financiadas com recursos públicos devem ser de livre acesso.

disseminação e acesso à informação (MARINHO, 2012). E esse contexto foi um dos pontos de motivação para o movimento de acesso aberto (*open access*), como também para a proposta de arquivos abertos (*open archives*), que buscam a disseminação do conhecimento científico e do patrimônio cultural, representados pelas *Open Archives Initiative – OIA* - e a *Budapest Open Access Initiative – BOIA* - (LIMA, 2009).

Em um contexto de acesso aberto, de um paradigma pós-custodial, informacional, tecnológico e científico⁶ (SILVA, 2017), voltado para a reflexão acerca da disponibilização da informação, o acesso à comunicação acadêmica e às pesquisas científicas ganhou nova dimensão, possibilitando que os pesquisadores ganhassem a oportunidade de ter acesso às informações que necessitassem, quebrando barreiras geográficas e físicas, em um curto espaço de tempo. Nesse âmbito, algumas ferramentas passaram a ser utilizadas para auxiliar na guarda e disseminação de informações, entre elas os repositórios digitais, que serão enfocados nessa pesquisa.

3.1 REPOSITÓRIOS DIGITAIS

Dobratz e colaboradores (2006, p. 2) definem Repositórios Digitais (RDi) como

uma organização (consistindo em sistemas de pessoas e técnicas) que assume a responsabilidade da preservação e acessibilidade a longo prazo dos objetos digitais, assegurando a sua usabilidade por parte de um grupo-alvo específico, ou comunidade designada.

Shintaku e Suaiden (2015, p. 33), os apresentam como “ferramentas de compartilhamento de conteúdo, em sua forma completa, seja texto integral ou outro formato, desde que apresente o conteúdo e não apenas os metadados”. Outra definição é dada por Abreu e Vidotti (2016) que os definem como ambientes digitais que permitem reunir dados e informações de cunho científico, administrativo, técnico, artístico, cultural, entre outros, cuja principal função é promover a visibilidade de seus objetos digitais, preservando-os por meio do gerenciamento de informação. Corujo (2014, p. 46) detalha o conceito, quando afirma que os RDis, em sua essência

são sistemas de informação que têm o intuito de gerir e armazenar colecções de objectos digitais. Esses sistemas de informação são da responsabilidade de um organismo ou instituição, que vai definir a finalidade, os objectivos, os tipos e características dos objectos que armazena, tendo em conta as

⁶ Segundo Miranda et. al (2012), o paradigma pós-custodial fundamenta-se na garantia de localização do conteúdo produzido e na promoção da visão social.

necessidades e expectativas da comunidade de interesse, potenciais clientes do repositório digital.

Esse mesmo autor também caracteriza o repositório afirmando que a arquitetura do repositório dever gerir tanto o conteúdo, quanto a metainformação a ele associada; que o conteúdo do repositório pode ser depositado pelo produtor ou proprietário do objeto digital ou ainda por terceiros; que o RDi deve ser sustentável, confiável, ter bom suporte e administração; além de ser capaz de oferecer um conjunto mínimo de serviços básicos, como por exemplo: admissão, disseminação, recuperação de informação e controle de acesso.

Monteiro e Brasher (2007) consideram o RDi como uma plataforma aberta que veio como uma saída para os obstáculos na publicação de pesquisas em periódicos pagos, tais como: o custo da assinatura do periódico, o dispendioso processo de submissão e publicação e a inacessibilidade aos documentos que estão fora dos circuitos editoriais. Eles foram idealizados como

um modo de fornecer acesso imediato e amplo a trabalhos de investigação científica, mas têm vindo a assumir o papel de curadores dessa produção científica, obrigando à adopção de políticas e ferramentas específicas para a sua preservação e curadoria (CORUJO, 2014, p. 32)

Os repositórios proporcionam informações estruturadas, classificação e indexação, o compartilhamento de informações e a preservação do conhecimento explícito (HAKOPOV, 2016). O que é endossado por Santos e Flores (2015, p. 206) quando falam que “com o auxílio de repositórios digitais é possível manter os formatos de arquivos sempre atualizados, desta forma, as estratégias de migração serão mais eficazes”, prevenindo, assim, que os documentos e as tecnologias utilizadas se tornem obsoletas e ocorra perda de informações. Adicionalmente, a utilização de repositórios aumenta a visibilidade de pesquisas, publicações e trabalhos diversos (ex: vídeos, áudios, etc) nele depositados (WEITZEL, 2006) e promove a interoperabilidade entre sistemas de informação (WEITZEL; LEITE; ARELLANO, 2008).

um repositório digital deve estar em conformidade com as normas e padrões estabelecidos, e trabalhar de forma colaborativa com outros serviços de preservação digital de forma a possibilitar níveis de interoperabilidade com outros repositórios digitais e sistemas informatizados para documentos arquivísticos digitais. (SANTOS; FLORES, 2015, p. 206)

O que também é enfatizado por Lampert (2013) quando diz que a utilização de RDi acaba valorizando os pesquisadores, pois com essa ferramenta, os mesmos podem

disponibilizar o andamento de suas pesquisas, sem precisar esperar por edições posteriores das revistas científicas, além de poderem disponibilizar, se não houver impedimento legal, artigos previamente publicados.

É possível depositar artigos já publicados ou quaisquer outras publicações online em repositórios digitais, a fim de possibilitar o acesso aos textos completos dos trabalhos já publicados bem como os seus dados descritivos (metadados) de autoria, título, palavras-chave etc. (WEITZEL, 2006, p. 61)

Para assegurar o alcance de seus objetivos, preservação e acesso a longo prazo, Rodrigues (2003) relata a importância de os repositórios: a) existirem no âmbito de um sistema organizacional, assegurando o acesso a longo prazo; b) fazer a manutenção dos recursos digitais; c) ter responsabilidade fiscal e base de sustentação; d) basear o sistema de acordo com convenções e normas internacionais; e) possuir metodologias para a avaliação da qualidade; e f) possuir políticas e práticas que possam passar por auditorias. Os repositórios se preocupando com esses quesitos citados anteriormente, têm um grande potencial para fornecer seus serviços por um longo prazo e promover o acesso à informação.

Sob a ótica de preservação de documentos, Lampert e Flores (2013) apontam que os repositórios digitais representam uma forma de garantir o acesso ao longo do tempo e de preservar os documentos nele armazenados. O que é endossado por Corujo (2014, p. 42) quando afirma que os repositórios

são vistos como a garantia do armazenamento e autenticidade dos conteúdos digitais. Os repositórios digitais vão integrar as problemáticas e soluções técnicas referentes à preservação e autenticidade da informação digital, que só faz sentido guardar tendo em vista o seu acesso aos públicos.

Para Ferreira (2011, p. 72), a preservação a partir de utilização de Repositórios Digitais significa que

(i) os dados são mantidos no repositório sem serem danificados, perdidos ou adulterados; (ii) os dados podem ser encontrados, recuperados e disponibilizados ao utilizador; (iii) os dados podem ser interpretados e entendidos pelo utilizador (não basta preservar a *bitstream* que representa a informação preservada, o desafio reside em assegurar que os utilizadores podem aceder ao conteúdo tal como este foi depositado no repositório); (iv) todas as condições referidas podem ser mantidas ao longo do tempo.

A autora reforça ainda que é preciso

assegurar que os documentos sejam preservados tal qual foram depositados, ou seja, que se mantêm a sua autenticidade e integridade e que todas as operações de preservação levadas a cabo estão devidamente justificadas e documentadas.

Apenas assim será possível garantir acessibilidade, autenticidade e confiabilidade das informações armazenadas.

3.2 CLASSIFICAÇÃO DOS REPOSITÓRIOS

Os repositórios podem ser classificados a partir de uma ou mais características como o tipo, formato ou origem do objeto que o compõem. Shintaku e Suaiden (2015) exemplificam alguns classificações como, pelo:

- a) Tipo de objetos: ex: repositório de distribuição de software; repositórios educacionais.
- b) Formato dos objetos: ex: repositórios fotográficos; repositórios textuais, repositórios multimídia.
- c) Origem do acervo: ex: Repositório Institucional

Traçando uma tipologia para os repositórios, Leite (2009) os divide em:

- a) Repositórios temáticos ou disciplinares: agregam trabalhos de uma área específica do conhecimento, ou seja, particulariza a produção intelectual em função das áreas do conhecimento.
- b) Repositório de teses e dissertações: lidam exclusivamente com teses e dissertações defendidas de programas de pós-graduação das diversas áreas do conhecimento;
- c) Repositório institucional (RI): voltados à produção intelectual de uma instituição, especialmente universidades e institutos de pesquisa. Eles podem englobar repositórios temáticos, pois são a junção de documentos de várias áreas diferentes, mas realizadas a partir de uma mesma instituição. Em consequência disso, possuem um conteúdo bastante heterogêneo e podem ter uma variedade de tipos de documentos (CAFÉ et al., 2003).

Esse trabalho foca em repositórios institucionais, por isso, esse tipo de repositório será melhor detalhado.

A partir do final de 2002, iniciou-se o processo de criação dos Repositórios Institucionais (RIs), motivados pelo movimento do acesso aberto. Esses repositórios surgiram

como uma estratégia, permitindo que as universidades realizassem mudanças no desenvolvimento científico e na forma de comunicar suas produções⁷ (LYNCH, 2003).

O objetivo dos repositórios institucionais, segundo Martins (2013 s/p), é “favorecer a difusão dos conteúdos acadêmicos da instituição, dar visibilidade à investigação realizada na instituição pelos seus membros, facilitar a conservação e preservação dos objectos documentais criados, armazenados e organizados”. Sayão e Marcondes (2009, p. 10) defendem que:

Repositórios institucionais são entendidos hoje como elementos de uma rede ou infraestrutura informacional de um país ou de um domínio institucional destinados a garantir a guarda, preservação a longo prazo e, fundamentalmente, o livre acesso à produção científica de uma dada instituição.

Os RIs reúnem toda a produção científica e/ou acadêmica de uma instituição produzida em formato digital ou digitalizada, em um único ambiente padronizado e interoperável. Eles são uma iniciativa que constrói as condições necessárias para permitir o acesso livre à produção científica de forma legítima, como também são uma alternativa para guarda e visualização da memória intelectual, cultural e científica de uma instituição, alterando o processo de produção, disseminação e uso da informação (BLATTMAN; BOMFÁ, 2006).

Um repositório institucional é uma base de dados digital e virtual (web-based database), de caráter coletivo e cumulativo (memória da instituição), de acesso aberto e interoperável que coleta, armazena, dissemina e preserva digitalmente a produção intelectual da instituição (DODEBEI, 2009, p. 91).

Para além disso, resguarda a memória digital através do tempo, materializado na produção realizada de artigos científicos, relatórios técnicos, livros, atas, entre outros.

3.3 PLATAFORMAS/SOFTWARES PARA REPOSITÓRIOS

Existem diversas plataformas e softwares que podem ser utilizadas para a criação de repositórios digitais. Uma das mais conhecidas e amplamente utilizadas é a chamada *Dspace*⁸. Essa plataforma foi desenvolvida pelo MIT e HP, no ano de 2002, baseada na comunicação científica, de forma colaborativa. Segundo Ferreira (2011), o *DSpace* “tem o objetivo de fornecer um sistema de repositório para documentos digitais resultantes de investigação ou

⁷ A comunicação científica consiste em proporcionar o intercâmbio de dados e informações sobre os trabalhos em andamento ou já finalizados e possibilitar o contato entre pesquisadores e suas pesquisas, viabilizando o trabalho colaborativo (LE COADIC, 2004).

⁸ <http://www.dspace.org>

destinados à educação e distribuído sob uma licença de código aberto”. O *DSpace* tem a função de:

gerenciar os diferentes e complexos recursos no formato digital (padrões de documentos como TIFF, AIFF, XML, ou publicados como especificações PDF, RIFF) para coletar, preservar, indexar e distribuir os itens digitais de ambientes acadêmicos (departamentos, laboratórios, centros, escolas, programas) (BLATTMANN; WEBER, 2008, p. 467).

Ela é uma plataforma gratuita, disponível online e completa quando se refere às características necessárias para um RD, já mencionadas anteriormente. Além do *DSpace*, outras plataformas também vem sendo utilizadas, mesmo em menor escala, são elas:

- Fedora ⁹(*Flexible Extensible Digital Object and Digital Repository Architecture*): é um software livre *open source*, desenvolvido em conjunto por equipes da *Cornell University* (Ciência da Informação) e pela Biblioteca da *University of Virginia*, ambas nos Estados Unidos. Surgiu em 1997, e se difere do *DSpace* e do *Eprints*, pois não é uma ferramenta específica para repositórios. Ferreira (2011) traz alguns recursos relacionados ao quesito de preservação que foram implantados no Fedora ao longo de sua trajetória: a) suporta versões de fluxos de dados, preservando a ligação entre as versões dos dados. b) possui um módulo que permite validar os formatos dos ficheiros e extrair os metadados a partir de seus fluxos de dados. A plataforma Fedora também visa fornecer um software de repositório de livre acesso e serviços relacionados para servir de base para muitos tipos de sistemas de gestão da informação.
- *EPrints*¹⁰ - Software *open source*, desenvolvido pela *School of Eletronics and Computer Science of Southampton*, no Reino Unido É descrito como a maneira mais fácil e rápida de criar repositórios de acesso livre para base de dados científica, teses, relatórios e multimídia. Foi traduzido para o Português e customizado pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT). A plataforma *EPrints* possui muitos recursos semelhantes ao *DSpace*, mas tem o diferencial de fornecer acesso ao autor/publicador, enquanto o segundo aborda estratégias que lidam com a preservação dos documentos a longo prazo. (TANSLEY, et al., 2003). O *EPrints* possui algumas características além da preservação, como a possibilidade de rastrear possíveis alterações efetuadas nos registros, no intuito de controlar esses registros. Os autores/depositantes podem ou

⁹ <http://fedorarepository.org>

¹⁰ <http://www.eprints.org/uk/>

não conceder permissões referente aos arquivos depositados no repositório, dando maior autoridade sob o uso desses arquivos. (FERREIRA, 2011)

- Archivematica¹¹ - Software *open source*, desenvolvido pela *Artefactual Systems Inc* com apoio do Programa Memória do Mundo – UNESCO, Biblioteca da *University British Columbia* e Arquivo da cidade de Vancouver. É um sistema de preservação digital para processamento e armazenamento de objetos digitais a longo prazo fundamentado nas estratégias de preservação digital de emulação, migração e normalização.
- RODA¹² (Repositório de Objetos Digitais Autênticos) - Software *open source*, desenvolvido pela Direção Geral de Arquivos (DGARQ), de Portugal, na plataforma Fedora, é um sistema de repositório digital criado para recolher, armazenar, preservar e dar acesso ao património arquivístico digital. Visa a preservação a longo prazo de objetos digitais, bem como pesquisa e acesso ao património arquivístico digital. Objetiva garantir o acesso continuado à informação custodiada, sem que haja prejuízo de performance e/ou da segurança dos dados.

3.4 E-SCIENCE E REPOSITÓRIOS DE DADOS

Em 2003, a Declaração de Berlim sobre o Acesso Aberto ao Conhecimento em Ciências e Humanidade (BERLIM, 2003) ampliou o escopo do que se compreendia por acesso livre ao afirmar que ele deveria incluir resultados de pesquisas científicas originais, dados não processados e metadados, fontes originais, representações digitais de materiais pictóricos e gráficos e materiais acadêmicos multimídia.

Nesse contexto, surgiu o termo e-Science, introduzido por John Taylor no ano de 2001, no Conselho Nacional de Pesquisa do Reino unido, para caracterizar um novo método utilizado na produção de informações científicas, que é baseada no compartilhamento de informações, metodologias e ferramentas. (BERTIM; VISOLI; DRUCKER, 2017). Segundo Taylor (2002) a e-Science é a ciência desenvolvida, de forma crescente, por meio de colaborações distribuídas ao redor do mundo, viabilizadas pela Internet e que faz uso de grande coleções de dados, recursos computacionais com grande capacidade de armazenamento e visualização de alto desempenho. Assim, a e-Science prioriza o dado de

¹¹ <https://www.archivematica.org/en/>

¹² <https://demo.roda-community.org>

pesquisa¹³ como o seu elemento central e faz uso de tecnologias atualizadas para auxiliar no desenvolvimento de pesquisas, abrindo possibilidades de pesquisas similares serem feitas de forma simultânea, onde pesquisadores podem reutilizar dados e metodologias de outros pesquisadores, proporcionando um melhor aproveitamento de recursos e de pesquisas já realizadas.

De fato, essa nova maneira científica, traz como prioridade a colaboração, compartilhamento e gerenciamento dos dados de pesquisa, tratando-se de uma ciência colaborativa. Ela visa a exploração de dados e a unificação das teorias, experimentos e simulações utilizando metodologias e ferramentas voltadas para o gerenciamento e reuso desses dados, alterando a forma como a ciência evolui. Por isso, a trajetória da evolução da ciência é acompanhada pelo progresso das formas de comunicação científica, bem como dos suportes de memória (SOUZA, 2012). Em especial porque, como afirma Ferreira (2011), muitos dos resultados de investigação científica tem valor duradouro e precisam ser preservados de forma a garantir a maximização dos investimentos, de forma que o conhecimento possa ser reutilizado, a fiabilidade dos dados possa ser garantida e a memória organizacional conservada. O que é endossado por Mueller (2000, p. 25), quando fala que os resultados das pesquisas alcançados por determinados pesquisadores “são frequentemente retomados por outros cientistas, teóricos ou aplicadores, que dão continuidade ao estudo, fazendo avançar a ciência ou produzindo tecnologias ou produtos neles baseados”. Também Targino (2000) reforça a importância que, no contexto da ciência, a informação deve ser tanto produzida, quanto comunicada/disseminada.

A ciência recorre, inevitavelmente, à informação e à comunicação. No entanto, é insensato restringir a comunicação à mera troca de informações entre cientistas, pois a ciência como sistema social integra elementos que vão desde a figura do pesquisador/cientista/acadêmico ao fluxo de ideias, fatos, teorias, métodos, literatura científica e instrumentos que permitem a operacionalização das investigações. Em linha similar de pensamento, é fundamental não se pensar apenas no processo do cientista buscando e usando a informação. Ao mesmo tempo que o pesquisador está envolvido no seu próprio trabalho e na aquisição de informações, está também produzindo e disseminando novas informações para os demais (TARGINO, 2000, p. 24).

Nesse cenário, as ações de e-Science costumam privilegiar a disponibilização e manutenção de bases de dados abertos, de acesso público, que subsidiem o trabalho de

¹³ Registros de fatos usados como fontes primárias na investigação científica e que geralmente são aceitos na comunidade científica como necessários para a validação dos resultados da pesquisa (OCDE, 2007).

pesquisa tanto no âmbito individual como no colaborativo, dentro do espírito da ciência aberta. Essa técnica de produção coletiva veio para agregar valores positivos na evolução da ciência. Desde então o crescimento, tanto da quantidade como na qualidade, de pesquisas científicas vem aumentando, trazendo assim uma nova realidade para a comunidade acadêmica poder preencher mais lacunas informacionais.

Essa nova realidade que a ciência enfrenta de um grande crescimento exponencial na produção de pesquisas científicas, abre portas para a utilização de novas metodologias, entre as quais alguns autores trazem a e-Science como uma alternativa.

Trata-se de metodologia de pesquisa emergente, a qual influencia preponderantemente os rumos tomados no desenvolvimento da pesquisa científica contemporânea, levando-se em consideração a combinação e ampla utilização dos elementos característicos da *e-Science* assinalados e determinantes para a implementação de alguns experimentos científicos realizados em ambientes colaborativos digitais. (FERREIRA; VILLALOBOS; MOURA, 2016, não paginado)

Para proporcionar o desenvolvimento de pesquisas colaborativas, autores como Sarita Albagli, André Luiz Appel e Maria Lucia Maciel (2013, p. 3) abordam que a e-Science requer a construção de uma infraestrutura informática computacional de uso distribuído, que possibilitar a colaboração à distância de equipes de pesquisa, envolvendo o uso intensivo e o compartilhamento de dados e recursos computacionais. Junto com novas perspectivas e novos métodos, para acompanhar essa evolução da ciência, também é preciso pensar no contexto geral, que se refere à políticas, tecnologias que surgem em paralelo para auxiliar nessa nova realidade.

Esse novo fazer científico pressupõe que novas regras e práticas sejam ensejadas pela comunidade científica, instituições de fomento, empresas e governo. Esse fato demanda que se pense amplamente nas dimensões epistemológica, política, ético-legal-cultural, morfológica e tecnológica que sustentam a efetivação da ciência aberta. (OLIVEIRA; SILVA, 2016, p.11)

Esses autores reforçam a ideia de que o e-Science é um modelo de ciência voltado para disponibilizar informações científicas de forma livre e aberta, para que outros pesquisadores possam utilizar esses dados para referenciar pesquisas futuras. Junto à isso, colocam a necessidade da criação de políticas e práticas próprias e utilizar tecnologias específicas para auxiliar ainda mais na evolução e desenvolvimento da ciência.

Essa concepção de grandeza na exploração dos dados disponibilizados pelas novas pesquisas, traz com ela uma preocupação: conseguir gerir e aproveitar da melhor maneira

possível toda essa informação que está sendo concedida a comunidade científica. Assim, pode-se dizer que a e-Science engloba o uso de uma nova geração de infraestrutura científica, seu caráter colaborativo e multi/interdisciplinar e de um grande volume de dados, além do importante papel desempenhado pela Computação na pesquisa científica. O que acaba por fazer com que o e-Science se relacione com outro conceito relevante o Big Data.

Big data compreende a geração, tratamento e análise de grandes volumes de dados que excedem as capacidades convencionais de processamento, e vem sendo explorado também por empresas, governos e outros segmentos interessados em extrair informações a partir de grandes quantidades de dados não estruturados. Já e-Science incorpora, além do uso intensivo de dados, a pesquisa colaborativa e o uso de recursos compartilhados para sua exploração. (ABAGLI; APPEL; MACIEL, 2013, não paginado)

A e-Science se refere à ciência desenvolvida com o mínimo possível de limitações de recursos colaborativos e computacionais, que englobam desde hardware e software, até a curadoria de dados e uma estrutura organizacional que estimule a cooperação entre os pesquisadores e potencialize os benefícios alcançados com a infraestrutura disponível (VAZ, 2011).

No contexto do e-Science, houve um crescimento exponencial dos dados científicos e, com isso, a gestão de dados científicos tornou-se um aspecto primordial no processo de investigação científica (HEY; TANSLEY; TOLLE, 2009; SAYÃO; SALES, 2014). Pois, de acordo com Bell (2011, p. 11) os dados científicos devem “permanecer para sempre num estado submetido à curadoria e acessível para o público para análise contínua”, o que converge mais uma vez para a questão do acesso aberto à informação científica.

Isso contribuiu para um repensar da divulgação e disponibilização de dados nos periódicos científicos e em repositórios digitais (TOLLE; TANSLEY; HEY, 2011; LYNCH, 2011). Essa valorização dos dados de pesquisas como recursos relevantes para uma ciência aberta tem reflexo na implantação de infraestruturas gerenciais e tecnológicas para o arquivamento desses dados. E, apesar que a e-Science no Brasil ainda está em um estágio embrionário, tanto os pesquisadores, quanto as instituições acadêmicas e as agências de fomento à pesquisa começam a compreender que esses dados, se devidamente tratados, preservados e gerenciados, podem constituir uma fonte inestimável de recursos informacionais (SAYÃO; SALES, 2014). Isso porque o trabalho colaborativo pode ser potencializado com a disponibilização de dados primários científicos em periódicos ou repositórios de dados científicos, já que esses dados podem habilitar pesquisadores a formular novas hipóteses de pesquisa, podem servir de entrada para novos estudos e pesquisas, para a

replicação dessas, além de para novas descobertas ou refutação de descobertas previamente realizadas (MAYERNIK et al., 2012). A formação desses conjuntos de dados é, em geral, decorrente de resultados de projetos de pesquisa financiados com recursos públicos, sendo disponibilizados para livre acesso após um período de carência.

De fato, “os dados que coletamos hoje podem ser usados no futuro de forma que ainda não conseguimos imaginar. (...) Quando você coleta os seus dados, reúne informações que, no futuro, poderão ser analisadas de formas muito diferentes”¹⁴. Além disso, os dados e conjuntos de dados de pesquisas providenciam as evidências necessárias para conferir veracidade, autenticidade e capacidade de reprodutibilidade ao corpo de conhecimento publicado nos periódicos e repositórios digitais. No entanto, mesmo projetos acadêmicos de e-Science, cuja intenção é trabalhar com dados abertos, falham em um ou mais aspectos, especialmente no que diz respeito às normas de divulgação de informações (DAVID; BESTEN; SCHROEDER, 2006). Além disso, nem sempre é considerada a importância de uma boa descrição dos dados compartilhados de forma que possam ser compreendidos e reutilizados no futuro. Pois, realmente, para que os dados de pesquisa possam manter a capacidade de serem interpretados em domínios distintos, poderem ser reusados, é necessário que eles estejam suficientemente organizados e documentados. Dessa forma, torna-se imprescindível que informações contextuais, estruturais e semânticas os acompanhem, de forma a torná-los auto descritos. Isso pode ser realizado, na prática, por esquemas de metadados. Esses esquemas devem registrar tudo o que deve ser de interesse do usuário, incluindo modelos de dados, equipamentos especiais, especificação da instrumentação, linhagem dos dados, entre outros.

Nesse cenário, os repositórios de dados se incorporaram à infraestrutura mundial de informação científica, proporcionando que os acervos de dados possam ser usados, reusados e compartilhados. Repositório de Dados (RD) são sistemas digitais de informação que armazenam dados provenientes de pesquisas científicas, a fim de colaborar com o seu acesso, uso, reuso e preservação. De acordo com Formeton (2015), eles têm sua importância como recurso informacional e vem se tornando um dispositivo de intercâmbio de experiências e compartilhamento de dados científicos, além de parte fundamental das infraestruturas

¹⁴ POLIAKOFF, M. [Depoimento]. In: JONES F. Editor-chefe da Nature fala sobre a abertura da ciência. Agência FAPESP, São Paulo, 06 mar. 2013. Disponível em: <<http://agencia.fapesp.br/16919>>. Acesso em: 30 set. 2017.

mundiais de pesquisa, tornando visível e aberta para toda a sociedade uma parcela importante da atividade de pesquisa, caracterizando a chamada Open Science ou Ciência Aberta¹⁵.

Entre as plataformas existentes para criação de um RD que vem sendo utilizadas, especialmente no contexto nacional, encontra-se o Dataverse. Ele é um RD de código aberto para compartilhar, preservar, citar, explorar e analisar dados de pesquisa. Cada conjunto de dados contém metadados descritivos e arquivos de dados (incluindo documentação e código que acompanham os dados) (DATAVERSE PROJECT, 2018).

Sayão e Sales (2012; 2014) afirmam que é um fato promissor observar que os dados de pesquisas estão sendo cada vez mais armazenados em repositórios de dados confiáveis onde, gerenciados sob os princípios da curadoria digital são preservados e mantêm a sua capacidade de reuso.

Repositórios de Dados Confiáveis, segundo o relatório do RLG/OCLC (2002, p. 1, tradução nossa), é conceituado como “aquele cuja missão é fornecer acesso confiável e de longo prazo e gerenciar os recursos digitais para sua comunidade designada, agora e no futuro”. O relatório também identifica algumas características chaves desse tipo de repositório, tais como: sustentabilidade, segurança no sistema, adaptação tecnológica e estrutura baseada no modelo de acesso aberto. Uma das características principais desse tipo de repositório é a autenticidade.

A autenticidade de um documento está diretamente ligada ao modo, à forma e ao status de transmissão desse documento, bem como às condições de sua preservação e custódia. Isso quer dizer que o conceito de autenticidade refere-se à adoção de métodos que garantam que o documento não foi adulterado após a sua criação e que, portanto, continua sendo tão fidedigno quanto era no momento em que foi criado (RONDINELLI, 2005, s/p).

A autenticidade é muito debatida nesse quesito pois os documentos e dados digitais possuem essa fragilidade e facilidade no seu manuseio. Por isso mesmo, um repositório digital confiável deverá atender aos procedimentos arquivísticos e aos requisitos de confiabilidade (CONARQ, 2014). Assim, também, como a confiabilidade¹⁶, que deve ser levada em conta nas medidas de segurança, desde a concepção dos repositórios digitais, a fim de garantir que os conteúdos armazenados permanecerão autênticos a longo prazo (MÁRDERO ARELLANO, 2008).

¹⁵ Movimento para tornar a pesquisa científica, os dados e a divulgação destes acessíveis a todos os níveis da sociedade (WIKIPÉDIA, 2018).

¹⁶ A confiabilidade está relacionada ao momento em que o documento é produzido e à veracidade do seu conteúdo. Para tanto, há que ser dotado de completeza e ter seus procedimentos de produção bem controlados. (...) A confiabilidade é uma questão de grau, ou seja, um documento pode ser mais ou menos confiável (CONARQ, 2011, p. 21).

O caminho do documento até chegar no depósito em um repositório confiável, passa por vários estágios, como por exemplo, publicação, análise, disponibilização, deixando o documento vulnerável a manipulação, pois várias pessoas diferentes conseguem ter acesso à ele e utilizá-lo com finalidades diferentes, nem sempre dando o controle necessário para seus autores/editores. Por isso, para promover a confiabilidade, é bastante importante nesse tipo de repositório a questão da certificação e da auditoria. Algumas iniciativas que recomendam a realização de auditorias podem ser vistas no Quadro 1 (SANTOS; FLORES, 2015).

Quadro 1 – Iniciativas que recomendam a realização de auditorias

Iniciativas	Objetivo
<i>Trustworthy Repository Audit & Certification: Criteria and Checklist (TRAC)</i>	Identifica os repositórios que conseguem realizar, o armazenamento, a migração e promover o acesso.
<i>Audit And Certification of Trustworthy Digital Repositories (ACTDR)</i>	Avalia o repositório digital com relação a infraestrutura organizacional, sustentabilidade financeira e a gestão dos documentos digitais e seus riscos.
<i>Digital Repository Audit Method Based on Risk Assessment (DRAMBORA)</i>	Avalia as qualidades, identifica os pontos fracos, e reconhece os pontos fortes do repositório.

Fonte: SANTOS;FLORES, 2015.

Porém, ressalta-se que, mesmo considerando o uso de repositórios de dados confiáveis, na atual infraestrutura de comunicação científica, os conjuntos de dados não estão conectados às publicações onde são discutidos e analisados. De acordo com Mueller (2007, p. 128), "para a ciência avançar não basta que o conhecimento seja publicado (...) o que exige entendimento de todo o processo de produção e comunicação da informação". Em consonância com a autora, Gray (2007) argumentou que todos os dados científicos precisam estar online, sejam eles oriundos da literatura ou também os dados brutos utilizados no contexto das pesquisas. Assim, a ideia é criar pontes que liguem as publicações existentes nos repositórios institucionais ou temáticos com os conteúdos dos repositórios de dados confiáveis (VERHAAR, 2008), ampliando assim o contexto da publicação. Logo, o documento abstrato que combina publicações e dados de pesquisas é chamado de “publicação ampliada” (*enhanced publication*) e emerge da compreensão de que as publicações tradicionais são limitadas na sua capacidade de incorporar os resultados de todo o ciclo de geração de conhecimentos da ciência contemporânea. Esse tipo de publicação será discutido no próximo capítulo.

4 PUBLICAÇÕES AMPLIADAS

Os pesquisadores interessam-se em obter recompensa científica por seus esforços e em aprender sobre pesquisas, bem como em reutilizar o resultado da pesquisa de outros para reproduzir experimentos semelhantes e evitar erros que já foram cometidos. Conforme Orrico (2012, p. 120), “as descobertas científicas provocam impactos sociais e culturais, envolvendo questões éticas relacionadas ao papel das instituições de pesquisa, dos órgãos de fomento e dos meios de comunicação, até a responsabilidade do cientista nesse processo de divulgar o que faz”. Orrico ainda afirma essa responsabilidade de divulgação é de todos os que compõem a comunidade científica. Por isso, a ciência contemporânea está sendo cada vez mais embasada nas infraestruturas de *e-Science* e vem adotando a abordagem colaborativa da *e-Research*, que promove fortemente o compartilhamento e reutilização de todos os resultados da pesquisa, incluindo publicações tradicionais, conjuntos de dados (GRAY 2007; MOONEY; NEWTON, 2012) e experimentos de pesquisa (DE ROURE; GOBLE; STEVENS, 2009). Esse embasamento é reforçado pelo fato que as agências de fomento e organizações, que são partes interessadas cruciais na cadeia de pesquisa, estão defendendo e tornando cada vez mais obrigatória a publicação e citação de qualquer resultado de pesquisa, a fim de medir o retorno do investimento, melhorar suas estratégias de financiamento, ou ganhar visibilidade e recompensas científicas (BARDI; MAHGHI, 2014). Adicionalmente, muitas vezes apenas o resultado final de um estudo, que aqui se entende como a publicação formal, torna-se uma redução de todo o processo pelo qual o experimento teve que percorrer (WEITZEL, 2006).

Nesse cenário, a disseminação dos resultados da pesquisa por meio de publicações tradicionais, não conseguia mais lidar sozinho com essas novas necessidades dos cientistas, agências e organizações. Dessa forma, para ultrapassar esses limites, a noção de publicação ampliada emergiu. Para Bardi (2014), ela é uma publicação digital que consiste em uma parte narrativa obrigatória (a descrição da pesquisa realizada) mais partes "relacionadas", como conjuntos de dados, outras publicações, imagens, tabelas, quadros, fluxos de trabalho, dispositivos, entre outros.

Para Bardi e Manghi (2014), publicações ampliadas podem ser definidas como objetos compostos dinâmicos, editáveis e identificáveis, combinando uma publicação eletrônica com dados de pesquisa embutidos ou remotos, materiais extras, dados de publicação, registros de banco de dados e metadados. O que é reforçado por Woutersen-Windhouver e Brandsma (2009, p.136) quando afirmam que ela pode conter “dados de pesquisa, materiais extras,

dados de pós-publicação, registros de base de dados (ex. *protein data bank*) e tem uma estrutura baseada em objeto com links explícitos entre objetos”. Assim, um objeto digital composto pode ter como parte um artigo, um conjunto de dados (*data set*), uma imagem, um filme, um comentário, um módulo ou um link para informação em uma base de dados (ARAYA, 2014; ELBAEK; PEDERSEN; SIERMAN, 2009).

Assim, a publicação ampliada é o conjunto formado pela narrativa textual dos resultados da pesquisa e pelos elementos informacionais inerentes ao próprio processo de pesquisa. Essas partes podem, ou não, ter um identificador e os metadados respectivos e se enlaçam por relações semânticas (BARDI; MANGHI, 2014). O que é endossado por Rodrigues e Sant’Ana (2016) quando afirmam que a publicação ampliada é uma composição de objetos digitais: dados, metadados, documentos e outros elementos descritivos, relacionados entre si, e que apresentem em seu todo, uma estrutura de corpo que lhe garante uma identidade e características próprias que possam ser vinculadas ao conjunto (RODRIGUES; SANT’ANA, 2016).

Sales, Sayão e Souza (2013) simplificam a definição alegando que a publicação ampliada é um documento abstrato que combina e-prints e dados de pesquisa. Para os autores, de forma geral, um manuscrito acadêmico é arquivado em um repositório digital institucional ou temático (repositório de e-prints) e outros componentes, arquivados no mesmo ou em outros repositórios, são agregados ao manuscrito como parte do fluxo de trabalho do desenvolvimento da atividade de pesquisa científica.

Analizando as definições encontradas, pode-se perceber que a publicação ampliada é uma ferramenta útil para a abertura e disseminação dos dados de pesquisas de forma integrada, garantindo aos dados sua significação original e a identificação de sua autoria. Além disso, ao agregar os dados de uma pesquisa ao seu resultado final publicado em um e-print, a publicação ampliada permite a preservação da memória da pesquisa científica realizada. Isso pode possibilitar a replicação da pesquisa para fins de validação ou ainda para colaborar com uma nova pesquisa. Sendo assim, a publicação ampliada pode ser considerada um veículo de comunicação científica de grande importância para a comunidade de pesquisadores (SALES; SAYÃO; SOUZA, 2013; SALES, 2014).

Ressalta-se que a ideia de publicações ampliadas pode não ser considerada nova, visto que os hiperlinks que começaram a ser utilizados no contexto da Web já traziam e trazem a ideia de relação entre objetos de vários tipos (imagem, som, vídeo, etc), como bem descrito na obra de Neves, Galindo e Cunha (1997). Pode-se dizer que a ideia dos hiperlinks está sendo

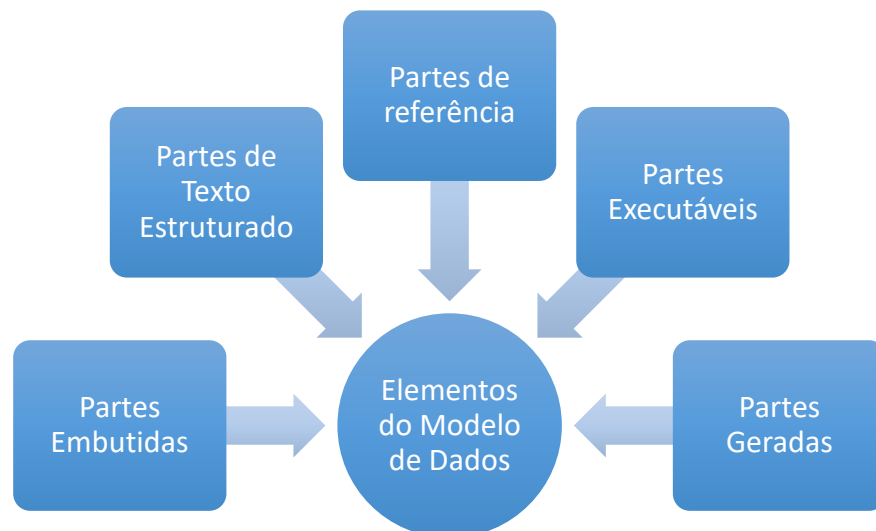
trabalhada agora, dentro do contexto da comunicação científica, em sistemas de informação, tais como os repositórios e periódicos de publicações ampliadas.

4.1 ELEMENTOS DA PUBLICAÇÃO AMPLIADA

Bardi (2014), após analisar diversos modelos de dados existentes para publicações ampliadas, observou que elas são frequentemente composta pelos elementos especificados na Figura 1 e descritos a seguir.

- **Partes Embutidas** (*embedded parts*) ampliam uma publicação com arquivos de materiais suplementares. Esses materiais podem melhorar a compreensão dos resultados apresentados na parte narrativa textual e podem ser slides, imagens de alta resolução, tabelas ou quadros que não couberam na íntegra na publicação original, apêndices, anexos, etc. Essas partes agregam um valor contextual à publicação.

Figura 1 - Elementos do modelo de dados de publicações ampliadas



Fonte: Adaptado de BARDI, 2014.

- **Partes de texto estruturado** (*structured-text parts*) – ampliam a publicação fornecendo uma estrutura editorial aos seus subcomponentes textuais. É composta por uma parte narrativa estruturada em várias partes interligadas (resumo, seções, figuras, tabelas, quadros, bibliografia, etc.). Esse modelo de dados define as relações enlaçando as várias partes e pode incluir metadados que as descrevam (exemplo, capítulo, seção, tabela).
- **Partes de Referência** (*reference parts*) – ampliam a publicação com referências a objetos que são "externos" à publicação ampliada e, portanto, podem ser compartilhados

com outras pessoas. Essas referências podem ser identificadores específicos da comunidade (por exemplo, identificadores *UKPubMed*), identificadores persistentes únicos (por exemplo, DOIs) ou URLs.

- **Partes Executáveis** (*executable parts*)- amplia a publicação com o contexto do experimento da pesquisa, o que inclui softwares, processos e dados para executar experiências. Isso permitirá a melhor interpretação e a validação por meio da repetição da experiência desenvolvida na pesquisa. Por exemplo, uma referência para um serviço web usado em um experimento, um fluxo de trabalho para ser executado por um determinado mecanismo ou, em geral, um código que pode ser dinamicamente executado por um determinado tempo de execução. Os estudos mais importantes sobre publicações ampliadas com partes executáveis para os propósitos de apoiar a revisão por pares, a validação de pesquisas e o reuso têm sido conduzidos no contexto dos ambientes virtuais de pesquisa (*Virtual Research Enviroments- VREs*¹⁷). Eles oferecem funcionalidades para o planejamento, execução e compartilhamento de experiências e são também chamadas de laboratórios eletrônicos. Alguns exemplos de laboratórios eletrônicos são IODA (SICIAREK; WISZNIEWSKI, 2011), SHARE (VAN GORP; MAZANEK, 2011) e D4Science.org.
- **Partes geradas** (*generated parts*) – amplia a publicação por meio de tabelas que podem mudar dinamicamente, dependendo das atualizações de dados de pesquisa de entrada fornecidos. Esse tipo de publicação ampliada contém partes de referência que se vinculam a bancos de dados externos (por exemplo, a referência para moléculas em um banco de dados químico). Tais partes podem ser processadas por uma aplicação do sistema de informação, de modo a gerar dinamicamente novas partes da publicação ampliada. Por exemplo, uma molécula pode ser processada para gerar sua renderização em 3D. Como tal, a renderização não é uma subparte estática da publicação, mas sim uma parte gerada dinamicamente.

Os elementos informacionais descritos, ao fazerem parte da publicação científica ampliada, podem melhorar a contextualização da pesquisa que os originou, propiciar a reprodutibilidade dos seus resultados e a validação do conhecimento científico produzido, dar uma maior visibilidade à publicação científica e, conseqüentemente, dar ao pesquisador o

¹⁷ Os VREs são plataformas digitais distribuídas que fornecem acesso a ferramentas, dados e serviços e permitem a colaboração entre pesquisadores, por meio de uma estrutura técnica que acessa uma infraestrutura de pesquisa mais ampla (CARUSI; REIMER, 2010).

reconhecimento científico. Ressalta-se, porém, que esses elementos informacionais devem estar efetivamente enlaçados com a publicação referente à pesquisa que os originou e que os enlaces devem ser perceptíveis ao ser humano e às aplicações baseadas em máquinas (ARAYA, 2014).

4.2 REQUISITOS E RECOMENDAÇÕES PARA A PUBLICAÇÃO AMPLIADA

Vehaar (2007) identificou uma série de dez requisitos e recomendações que estabelecem uma base para a formalização tecnológica e informacional da publicação ampliada, aumentando suas possibilidades de recuperação, acesso, uso e reuso, interoperabilidade e preservação. Esses requisitos são:

a) deve ser sempre possível especificar as partes componentes de uma publicação ampliada, conforme seção 4.1;

b) a publicação ampliada e seus componentes devem ser disponibilizados como recurso web identificados por Uniform Resource Identifiers (URI), o mesmo vale para seus componentes;

c) deve ser possível agregar objetos digitais compostos à publicação ampliada;

d) deve ser possível acompanhar as diferentes versões das publicações ampliadas como um todo e das suas partes constituintes, criando a possibilidade de acompanhamento do andamento da pesquisa;

e) deve ser possível registrar as propriedades básicas da publicação e dos outros recursos que estão adicionados a ela. Esse ponto aborda a possibilidade de se registrar as propriedades dos arquivos complementares, já que essas publicações são produzidas de forma colaborativa e interdisciplinar, trazendo confiabilidade para o recurso em questão.

f) deve ser possível registrar a autoria da publicação ampliada e dos outros recursos que a compõe. Onde o próprio Verhaar (2007) retrata que a autoria da publicação ampliada não implica automaticamente a autoria dos arquivos complementares ou dos conjuntos de dados associados a ela, dessa forma é importante poder registrar a autoria desses dados.;

g) deve ser possível assegurar a preservação de longo prazo das publicações ampliadas;

h) deve ser possível registrar as relações entre os recursos web que fazem parte da publicação ampliada. Isso traz a possibilidade do autor explicitar as relações dos arquivos que foram utilizados para a produção e desenvolvimento da sua pesquisa, mostrando aos seus

usuários a linha de raciocínio e quais abordagens e linhas de pensamento foram utilizadas para chegar até aqueles resultados.

i) As instituições que oferecem acesso a publicações ampliadas devem assegurar que elas possam ser descobertas. Assim, esse tipo de publicação deve ser visível por outros sistemas, para que seus conteúdos possam ser acessados por rastreadores web e mineração de dados, possibilitando uma melhor recuperação da informação, mesmo que o usuário comece sua pesquisa fora daquela base de dados e possa ser redirecionado a ela através de links de busca.

As instituições que oferecem acesso a publicações ampliadas devem assegurar que elas estejam disponíveis como documentos baseados no modelo *Object Reuse and Exchange* (OAI-ORE)¹⁸.

4.3 SISTEMAS DE INFORMAÇÃO PARA PUBLICAÇÃO AMPLIADA

Bardi e Manghi (2014) propõem uma classificação dos sistemas de informação para publicação ampliada existentes de acordo com seus objetivos funcionais. Nesse sentido, eles os caracterizam por proporcionar um conjunto de funcionalidades de gerenciamento e de consumo.

As funcionalidades de gerenciamento permitem a criação, a exclusão e a atualização de publicações ampliadas. Essas ações habitualmente são realizadas via interfaces usuário final. Elas os guiam sobre como fornecer partes da publicação e especificar as relações entre elas. Essas partes, dependendo do modelo de dados que o sistema comporta, podem ser fornecidas por meio de upload de arquivos a partir do sistema de arquivo do periódico ou pelo fornecimento de referências para arquivos (URL, DOI, etc.). Em alguns casos, essas partes podem ser identificadas ou geradas em tempo de execução por aplicativos ou podem ser produzidas por sistemas que procuram, por exemplo, fornecer as partes necessárias para compartilhar um experimento repetível.

As funcionalidades de consumo para o uso das publicações ampliadas se identificam, por exemplo, na leitura, no compartilhamento e na execução dessas publicações. Bardi e Manghi (2014) observam que essas funcionalidades atendem principalmente a quatro motivações científicas. Cada uma dessas motivações identificaria o sistema de informação capaz de propiciar a disseminação da publicação ampliada. As motivações seriam:

¹⁸ A *Open Archive Initiative* (OAI) define normas para descrição e intercâmbio de agregações de recursos da Web em seu projeto denominado *Object Reuse and Exchange* (OAI-ORE) (ARAYA, 2014). Maiores detalhes em: <http://www.openarchives.org/>

- a) a exportação de publicações ampliadas como pacotes de material suplementar;
- b) melhorar a legibilidade e compreensão da narrativa;
- c) a interconexão com dados de pesquisa;
- d) permitir a repetição de experimentos.

Os sistemas de informação que atenderiam às funcionalidades de consumo são (ARAYA, 2014; BARDI; MANGHI, 2014): sistemas de informação que atendem à motivação de ampliar a publicação por meio de pacotes com material suplementar; sistemas de informação que atende à motivação de melhorar a legibilidade e a compreensão do estudo científico; sistemas de informação para permitir a repetição de experimentos; e sistemas de informação para a interconexão com dados de pesquisa. Esses últimos serão os trabalhados nessa pesquisa e por isso serão melhor detalhados a seguir.

4.3.1 Sistemas de Informação para Interconexão com Dados de Pesquisa

Os sistemas de informação para interconexão com dados de pesquisa têm como principal objetivo funcional propiciar o enriquecimento da publicação ampliada com ligações para dados relevantes da pesquisa. Espera-se com isso fortalecer a citação desses dados e facilitar o seu reuso. Esses sistemas de informação atendem aos interesses de comunidades científicas, organizações e agências de financiamento que procuram estabelecer padrões e melhores práticas para a publicação e citação de conjuntos de dados e publicações na Web. Assim, os conjuntos de dados devem ser identificáveis, reusáveis e os pesquisadores que os produzem, devem ser cientificamente recompensados pelos seus esforços (BARDI; MANGHI, 2014).

Um exemplo desse tipo de sistema de informação é o *Europe PubMed Central* (Europe PMC)¹⁹ de acesso livre. Seu modelo de dados inclui: partes de texto estruturado (usadas no texto narrativo), partes embutidas com material complementar fornecido pelos autores para o editor e partes de referência que podem estar constituídas por uma lista de referências a bases de dados públicas como *UniProt*, *Protein Data Bank (PDB)* e *European Nucleotide Archive (ENA)* e por entradas a ontologias e taxonomias.

Logo, esse tipo de sistema de informação tem a motivação de poder possibilitar a criação de enlaces entre o artigo científico e conjuntos de dados para que, ao seguir os

¹⁹ <https://europepmc.org/>

enlaces, seja possível realizar a navegação por e o acesso às partes que constituam a publicação ampliada.

Sales (2014) e Sayão e Sales (2016) ressaltam que os próprios repositórios institucionais ou plataformas de periódicos científicos podem ter seus modelos de dados estendidos para incluir dados de pesquisa e outros objetos, podendo assim trabalhar com publicações ampliadas. Porém, na maioria dos casos, vem surgindo iniciativas onde os repositórios de dados são plataformas independentes que podem estabelecer links entre os seus recursos e os recursos dos repositórios de eprints, estabelecendo novas formulações de publicações acadêmicas.

5 METODOLOGIA

Quanto aos objetivos, esta pesquisa tem caráter exploratório, por proporcionar familiaridade com o problema, maximizando o conhecimento do pesquisador em relação a este. Em geral, consiste no primeiro passo para se estudar sobre um campo o qual não se tem conhecimento suficientemente, sendo considerada como ponto inicial para estudos futuros (GIL, 2010). Essa pesquisa se configura também como bibliográfica, pois seu estudo e desenvolvimento foi feito a partir de documentos bibliográficos, como periódicos, artigos científicos publicados em eventos acadêmicos e livros. Utilizou-se como termos de busca “repositório institucional”, “repositórios de dados”, “dados de pesquisa”, “publicações ampliadas”, “repositórios de dados confiáveis”, “repositórios de publicações ampliadas”, “ciência aberta”, “ciência de dados”, “e-science” e combinações dessas, em português e em inglês. Os materiais recuperados foram filtrados/selecionados a partir da leitura do título e do resumo. Os selecionados foram lidos na íntegra e proporcionaram a construção do trabalho.

Ainda sobre o processo metodológico, acrescenta-se que a pesquisa enquadra-se na abordagem qualitativa (GIL, 2010) que caracteriza-se por ser “interpretativa, baseada em experiências, situacional e humanística” (STAKE, 2011, p. 41)., sendo consistente com suas prioridades de singularidade e contexto. Deste modo, ela procura assegurar que o propósito da investigação não é alcançar a generalização, mas fornecer exemplos situacionais à experiência do pesquisador (STAKE, 2011). Dentro dessa abordagem faz uso do método do estudo de caso que, de acordo com Yin (2015) investiga um fenômeno considerando seu contexto, realizando uma análise sob a conjuntura real.

A pesquisa possui como objeto de estudo os repositórios institucionais brasileiros de instituições, institutos e/ou centro de pesquisas que fazem uso de publicações ampliadas. Para identificá-los, tanto procurou-se na literatura consultada se havia menção a eles, como foram consultados os sites dos repositórios institucionais cadastrados no site do Opendoar²⁰, em Janeiro/2018 os repositórios institucionais cadastrados. Assim os 77 sites de repositórios encontrados foram verificados por meio de acesso e navegação, em busca de quais faziam uso de publicações ampliadas. Após esse processo apenas dois repositórios foram identificados que alegavam fazer uso de publicações ampliadas: o repositório CarpeDien²¹ do Instituto de

²⁰ Opendoar.org

²¹ <http://CarpeDien.ien.gov.br/>

Energia Nuclear (IEN) e o RIDI²² do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT).

Adicionalmente foi feito um levantamento na base de repositórios r3data.org e também na literatura por repositórios de dados relacionados a instituições e institutos de pesquisa já em funcionamento.

Os repositórios encontrados foram analisados a partir dos requisitos especificados por Verhaar (2007), a fim de verificar se eles efetivamente atendiam, não atendiam ou parcialmente atendiam²³, a condição de fazer uso de publicação ampliada, que são:

- a) Especificar as partes componentes da publicação.
- b) Ser disponibilizado como recurso web.
- c) Agregar objetos digitais referentes à publicação.
- d) Acompanhar diferentes versões da publicação e de seus componentes.
- e) Registrar as propriedades básicas da publicação.
- f) Registrar a autoria da publicação e dos recursos complementares.
- g) Possibilitar a preservação à longo prazo.
- h) Registrar as relações entre os recursos web que complementam a publicação.
- i) As Instituições que disponibilizam as pesquisas devem assegurar que as pesquisas sejam encontradas.

Os resultados obtidos foram consolidados, apresentados e discutidos à luz da literatura consultada.

²² <http://repositorio.ibict.br/>

²³ Foi considerado parcialmente atender também quando o repositório demonstrava indícios ou tinha recursos que colaboravam com a propriedade.

6 ANÁLISE DOS REPOSITÓRIOS DE PUBLICAÇÕES AMPLIADAS

Após a navegação nos repositórios levantados no site do Opendoar, verificou-se que, especificamente no âmbito brasileiro, foi difícil a localização de repositórios que trabalhem com publicações ampliadas. Também informações sobre repositórios de dados em instituições de pesquisa nacionais foram limitadas. Foi encontrado no cadastro do r3data.org apenas o Banco de Dados Científicos da Universidade Federal do Paraná (UFPR)²⁴, que 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. Esta base está em fase de integração com o Repositório Digital Institucional da referida instituição, o que poderá constituir, futuramente, uma iniciativa de publicação ampliada. E no referencial teórico houve menção ao projeto piloto do repositório de dados da UNESP (VIDOTTI et al., 2017). Também foi encontrado o trabalho de Sayão e Sales (2016), onde eles tratam o repositório CarpeDien, do Instituto de Energia Nuclear (IEN/CNEN), como enquadrado no campo de publicações ampliadas. Já o repositório do IBICT, em suas publicações vem dando indícios e estimulando o uso dessas publicações, dessa forma optou-se por analisar esses dois últimos repositórios para verificar se há uso efetivo de publicações ampliadas.

6.1 REPOSITÓRIO INSTITUCIONAL DO INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT)

No site do repositório, segundo o Ibict (2016, p.4), através do Guia de Depósito de Obra, se retrata como “uma base de dados para reunir e divulgar a produção científica dos pesquisadores do Instituto Brasileiro de Informação em Ciência e Tecnologia.” E ainda menciona que seus objetivos são “armazenar, preservar, divulgar, dar acesso à produção intelectual do IBICT, proporcionando uma maior visibilidade e maximizando os impactos da pesquisa como parte do movimento pelo Acesso Aberto à informação.”

Na sua apresentação, ele não se classifica como um repositório de publicações ampliada de forma explícita, mas navegando nas suas páginas é possível perceber indícios de intenções de disponibilização desse tipo de publicação. Por conta desses pequenos indícios,

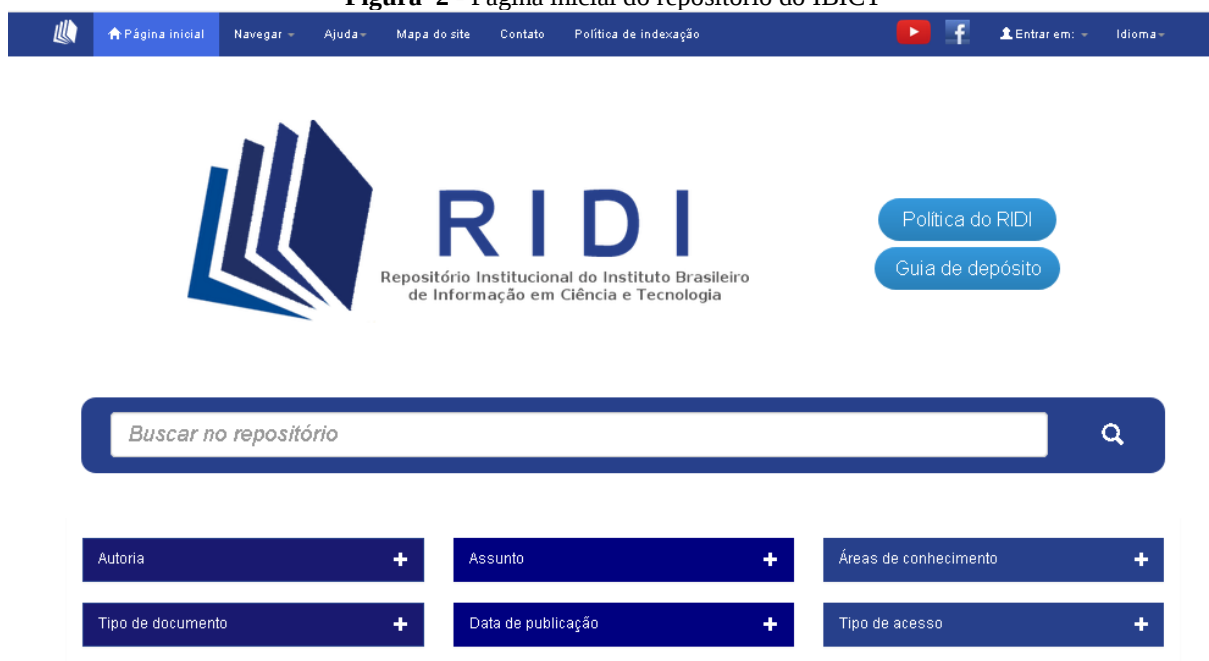
²⁴ <https://bdc.c3sl.ufpr.br/>

algumas pesquisas disponibilizadas pelo próprio Ibict começaram a tratar o seu repositório como um dos pioneiros no ramo de publicação ampliada no Brasil.

Após a verificação no site, foi visto que o Repositório Institucional Digital do Ibict (RIDI) disponibiliza arquivos referente aos resultados das pesquisas feitas nessa instituição. O site possui a possibilidade de navegação em três idiomas (português, inglês e espanhol), possui em média 730 arquivos publicados, como artigos, teses, conferências e livros, de diversas áreas do conhecimento.

Disponibiliza para seus usuários um guia de depósito, como um documento auxiliar, onde explica o que é e qual o objetivo do repositório, além de explicar como o usuário pode acessar, depositar arquivos (auto arquivamento) e fazer login no sistema. O mesmo foi analisado a partir do site <http://repositorio.ibict.br>.

Figura 2 - Página inicial do repositório do IBICT



Fonte: IBICT, 2017

Analisando a interface do repositório é percebida a intenção de guiar bem seus usuários, pois disponibiliza filtros em forma de botões, que abrem opções de resultados, como o nome de alguns autores, tipos de assunto, entre outros. E assim, auxilia a busca e navegação dos usuários junto as suas publicações. Posteriormente o repositório foi analisado (Quadro 2) a partir dos requisitos abordados por Verhaar (2007), verificando a forma que as pesquisas são disponibilizadas no repositório e qual o retorno conseguido pelos usuários, mostrando quais

requisitos o repositório consegue alcançar, e quais ainda deixam a desejar no contexto de publicações ampliadas.

Quadro 2 - Análise do repositório do RIDI

Requisitos	Análise		
	Atende	Há indícios da característica	Não atende
Especificar as partes componentes da publicação			X
Ser disponibilizado como recurso web		X	
Agregar objetos digitais referentes à publicação			X
Acompanhar diferentes versões			X
Registrar as propriedades básicas		X	
Registrar a autoria da publicação e dos recursos		X	
Possibilitar a preservação à longo prazo		X	
Registrar as relações entre os recursos web		X	
Assegurar que as pesquisas sejam encontradas.	X		

Fonte: Dados da Pesquisa, 2018.

Após a análise alguns pontos precisam ser destacados. Sobre o ponto “especificar as partes componentes da publicação”, a plataforma detalha a parte textual (de texto estruturado) e seus metadados (tais como: tipo, título, resumo, autor, idioma, palavras-chave, etc – é permitida a visualização completa do registro de metadados) e tem campo, que será descrito mais abaixo para a inserção de materiais suplementares (parte embutida). O ponto “ser disponibilizado como recurso web”, é alcançado quando a plataforma disponibiliza a URI de cada componente da publicação, porém, a URI só é disponibilizada para o recurso principal (Figura 3).

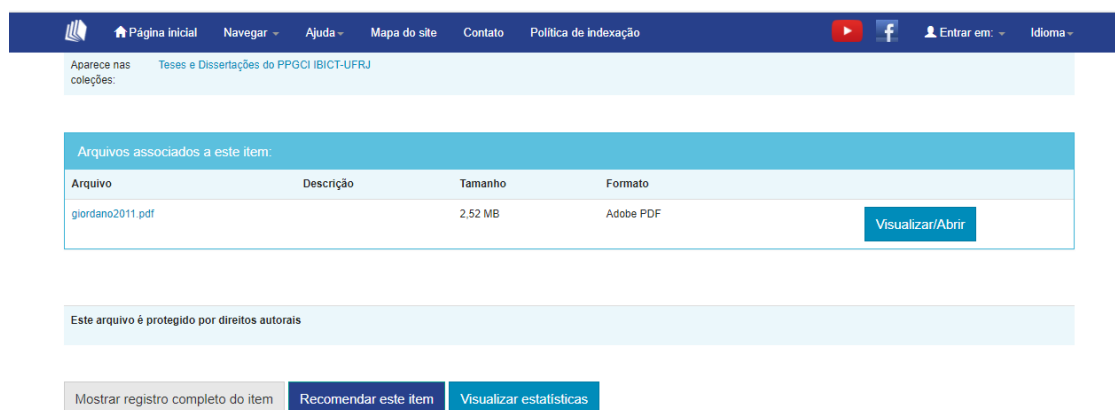
Figura 3 - URI de Recurso no Repositório RIDI

Use este identificador para citar ou linkar para este item: http://ridi.ibict.br/handle/123456789/889		
Registro completo de metadados		
Campo DC	Valor	Idioma
dc.creator	Rodrigues, Diogo Duarte	-
dc.date.accessioned	2016-10-10T17:32:15Z	-
dc.date.available	2016-10-10	-
dc.date.available	2016-10-10T17:32:15Z	-
dc.date.issued	2016-06-29	-
dc.identifier.uri	http://ridi.ibict.br/handle/123456789/889	-

Fonte: IBICT, 2017

Já o ponto “agregar objetos digitais referentes a publicação”, que é um requisito de grande relevância para se caracterizar como um repositório de publicações ampliadas, não é alcançado pelo RIDI (Repositório Institucional do Instituto Brasileiro de Informação em Ciência e Tecnologia), onde se enxerga um grande distanciamento para que esse repositório alcance os pontos que precisam para se tornar de fato um repositório de publicações ampliadas. Porém, alguns indícios já são vistos nessa plataforma, como um campo chamado “arquivos associados a este item” que representa um campo para disponibilizar arquivos que complementem a publicação principal, como mostra a Figura 4. Arquivos que foram utilizados para auxiliar na composição da pesquisa, sendo eles vídeos, imagens, outros textos ou pesquisas, entre outros. A disponibilização desse material complementar só tem a enriquecer mais ainda a evolução da pesquisa publicada, pois permite ao usuário ter acesso à esses arquivos para evoluir e agregar conhecimento junto aos resultados da pesquisa.

Figura 4 - Imagem dos Arquivos associados do site do Ibict



Fonte: IBICT, 2017

Junto ao ponto dos arquivos complementares, vem o requisito chamado de “acompanhar diferentes versões”, onde há a possibilidade e espaço para publicação do andamento da pesquisa, diferentes versões que mostram bem a evolução da pesquisa, não precisando esperar até obter todos os resultados para uma posterior publicação. O repositório RIDI, ainda não traz a possibilidade de acompanhar as pesquisas por meio de versões, a evolução poderia ser realizada, da forma como está por meio do campo para documentos associados.

“Registrar as propriedades básicas” é um ponto que aborda sobre possibilidade de registrar as propriedades dos arquivos complementares, junto com “registrar a autoria da publicação e dos recursos”, onde o repositório analisado, RIDI não disponibiliza essas

informações de forma completa, pois mostra as propriedades da pesquisa central, do recurso textual principal, mas não abre o espaço para registrar as propriedades dos recursos suplementares, nem seus respectivos autores.

O ponto “possibilitar a preservação a longo prazo” foi considerado parcialmente atendido, porque apesar da política de informação (IBICT, 2017) do repositório não mencionar a preservação da informação, ela trata de disponibilizar acesso a longo prazo, o que acabar por envolver a aplicação de alguma estratégia de preservação digital não explicitada.

Sobre o ponto “registrar as relações entre os recursos web”, o repositório não disponibiliza essas tais relações do conteúdo principal com outros recursos na Web. Porém, na visualização completa dos metadados, é possível verificar que os elementos dc.creator (autor) e dc.contributor (colaborador do recurso) fazem link com recursos Web externos ao repositório (Figura 5), no caso a base do currículo Lattes dos envolvidos. Sendo indícios iniciais dessa característica.

Figura 5 – Uso de URI externa no repositório RIDI

dc.creator.Lattes	http://lattes.cnpq.br/2409776496381886	pt_BR
dc.contributor.advisor1	Capeller, Ivan	-
dc.contributor.advisor1Lattes	http://lattes.cnpq.br/2258303291123908	pt_BR
dc.contributor.referee1	Pinheiro, Lena Vania Ribeiro	-
dc.contributor.referee1Lattes	http://lattes.cnpq.br/9613980184982976	pt_BR
dc.contributor.referee2	Pereira, Marcelo Fernandes	-
dc.contributor.referee2Lattes	http://lattes.cnpq.br/2957130184107059	pt_BR

Fonte: IBICT, 2017

Para possibilitar o último requisito, “assegurar que as pesquisas sejam encontradas”, o repositório mantém as suas publicações visíveis e recuperáveis tanto por máquinas quanto por usuários, de forma que possam para que os usuários possam ter acesso ao conteúdo completo de seus arquivos, possibilitando a disseminação da informação por um longo prazo. Também porque a plataforma é baseada nos conceitos de arquivos abertos, pois disponibiliza seus arquivos de forma aberta e livre, onde o acesso ao repositório é feito de forma gratuita, online e todos os arquivos estão no formato PDF, um formato aberto muito utilizado para disponibilização de arquivos.

Analisando o repositório com um olhar mais geral, dispensando a visão somente dos requisitos trazidos por Verhaar, outro ponto observado na plataforma foi referente a pesquisa feita a partir do autor, onde nos resultados há a possibilidade de navegar e encontrar outras pesquisas feitas pelo autor mencionado (sendo autor principal ou membro participante), através de links disponíveis pela própria plataforma, redirecionando o usuário à outras páginas, criando referências cruzadas.

Como desfecho da análise do repositório RIDI, diferente do encontrado na literatura, pode-se afirmar que, de modo geral, o repositório ainda não se encaixa no campo de repositórios de publicações ampliadas, mesmo já tendo alguns indícios de que pretende evoluir seu modelo de publicações.

6.2 REPOSITÓRIO CARPEDIEN: DADOS E INFORMAÇÕES EM ENERGIA NUCLEAR – INSTITUTO DE ENGENHARIA NUCLEAR.

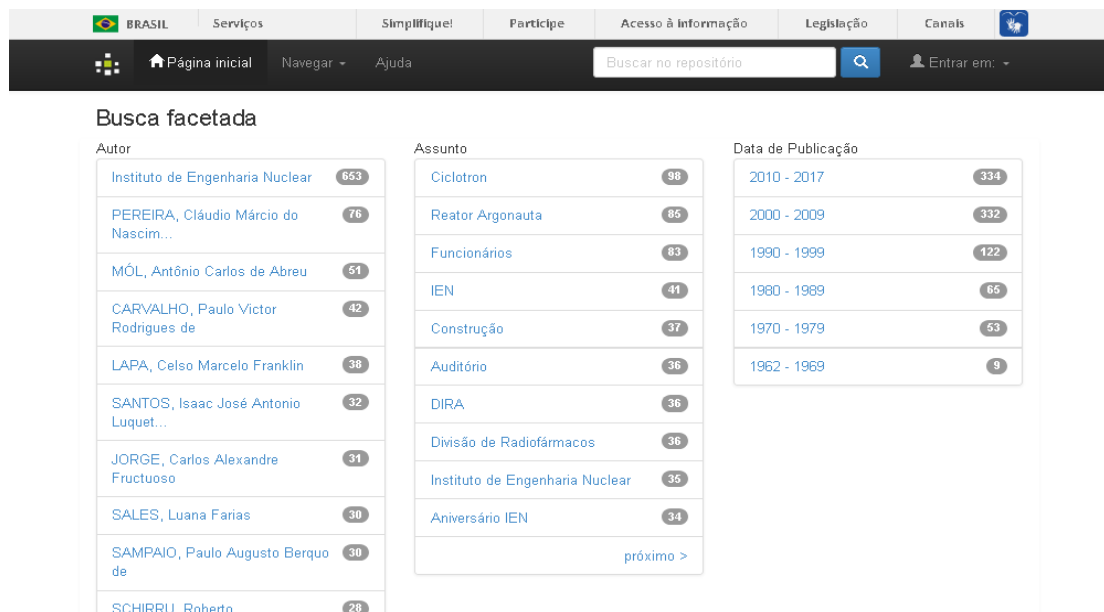
No próprio site do repositório, na aba “ajuda”, não aparece uma descrição bem definida do repositório, mas relata que a plataforma “captura, distribui e preserva produtos de pesquisa digital.” Além de mencionar que disponibiliza arquivos em vários formatos digitais. E oferece um conteúdo que cresce diariamente, a medida que novas pesquisas vão sendo adicionadas ao repositório. Dentre todas as informações oferecidas na própria plataforma, ela não é apresentada como um repositório de publicações ampliadas, diferente do exposto em Sales (2014) e Sayão e Sales (2016), que já o caracterizam assim em suas pesquisas.

O repositório (Figura 6) possui a possibilidade de navegação em dois idiomas (português e inglês) e possui 1444 arquivos publicados, como artigos, teses, conferências e livros, referente às áreas específicas do conhecimento como: química, tecnologia química e ciência em geral. A análise dessa plataforma foi feita a partir do site: <http://CarpeDien.ien.gov.br>.

Figura 6 - Página do Repositório CarpeDien

Fonte: CNEN, 2018.

Olhando para a interface do repositório, percebe-se que ela é de fácil manuseio. A busca oferecida é a simples e uma busca chamada de facetada (Figura 7). Para facilitar a percepção das novas adições no repositório, na página inicial são exibidas chamadas para as últimas publicações feitas no site, dando possibilidade ao usuário de acessar as pesquisas mais atuais da base.

Figura 7 - Busca Facetada no site do repositório CarpeDien

Fonte: CNEN, 2018.

O repositório CarpeDien, assim como o RIDI, foi analisado a partir dos requisitos abordados por Verhaar (2007) e o resultado pode ser visualizado no Quadro 3.

Quadro 3 - Análise do repositório do CarpeDien

Requisitos	Análise		
	Atende	Há indícios da característica	Não Atende
Especificar as partes componentes da publicação		X	
Ser disponibilizado como recurso web		X	
Agregar objetos digitais referentes à publicação			X
Acompanhar diferentes versões			X
Registrar as propriedades básicas		X	
Registrar a autoria da publicação e dos recursos			X
Possibilitar a preservação à longo prazo		X	
Registrar as relações entre os recursos web		X	
Assegurar que as pesquisas sejam encontradas.	X		

Fonte: Dados da Pesquisa, 2018.

Sobre o primeiro ponto discutido “especificar as partes componentes da publicação”, a plataforma detalha a parte textual (de texto estruturado) e seus metadados (tais como: tipo, título, resumo, autor, idioma, palavras-chave, etc), permitindo, inclusive a visualização completa de todos os metadados cadastrados.

O ponto “ser disponibilizado como recurso web” é alcançado quando a plataforma disponibiliza a URI de cada publicação, porém, no CarpeDien a URI só é disponibilizada para o recurso principal (Figura 8). Porém, chama a atenção o fato de que são apresentadas duas URI distintas na mesma página, uma em destaque acima do registro completo de metadados do recurso e outra no elemento de metadados dc.identifier, quando esse identificador deveria ser único em ambos os lugares, como no caso do RIDI (Figura 3).

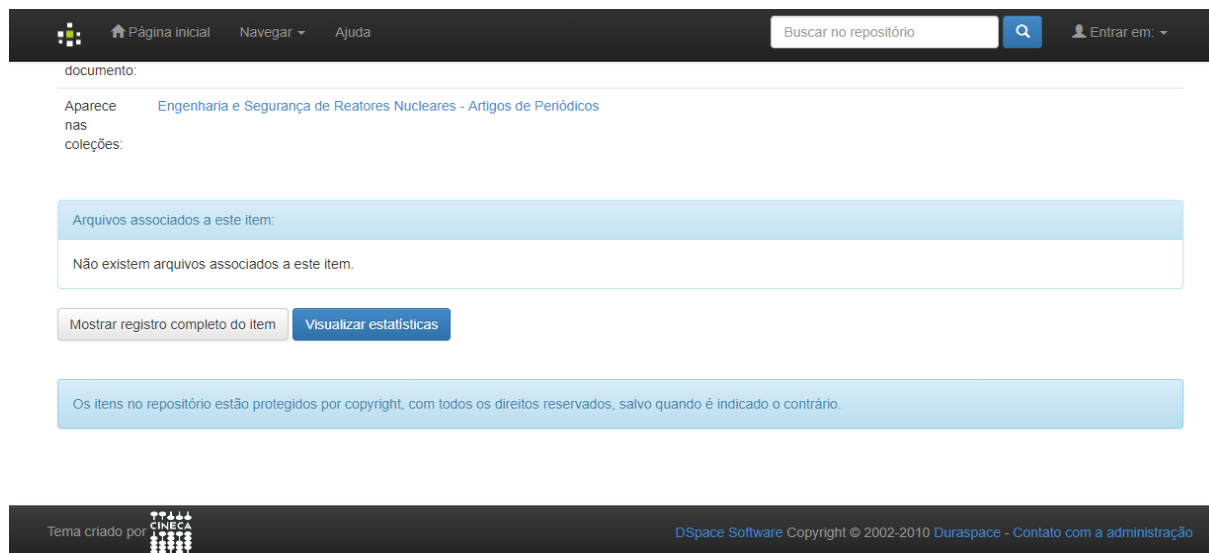
Figura 8 - URI de Recurso do Repositório CarpeDien

Use este identificador para citar ou linkar para este item: http://carpedien.ien.gov.br:8080/handle/ien/649		
Registro completo de metadados		
Campo DC	Valor	Idioma
dc.contributor.author	SALES, Luana Farias	-
dc.contributor.author	SAYÃO, Luís Fernando	-
dc.date.accessioned	2014-05-20T15:31:09Z	-
dc.date.available	2014-05-20T15:31:09Z	-
dc.date.issued	2013	-
dc.identifier.uri	http://hdl.handle.net/ien/649	-

Fonte: CNEN, 2018.

O “agregar objetos digitais referentes à publicação”, o CarpeDien, da mesma forma que o RIDI, permite associar arquivos à uma pesquisa ou documento (Figura 9). Porém eles não possuem URI própria, logo, esse item é atendido parcialmente.

Figura 9 - Associação de arquivos complementares no CarpeDien



Fonte: CNEN, 2018

Sobre “acompanhar diferentes versões”, nesse repositório não foi encontrado nenhuma opção de versionamento disponível. Sobre os pontos “registrar as propriedades básicas” e “registrar a autoria da publicação e dos recursos” no CarpeDien, não há disponibilização das autorias complementares, nem de informações/propriedades adicionais sobre elas, esses conteúdos são preenchidos apenas no recurso principal.

Sobre o ponto “possibilitar a preservação à longo prazo”, no próprio repositório não foi encontrada políticas ou informações sobre a questão da preservação. Porém, em artigos publicados sobre o mesmo ele é considerado uma ferramenta de curadoria e preservação. Logo, esse foi considerado um ponto parcialmente atendido, mesmo não sendo explicitadas as estratégias adotadas.

No ponto “registrar as relações entre os recursos web”, o repositório não disponibiliza formas de relacionar o conteúdo principal com outros recursos na Web. Porém, na visualização completa dos metadados, é possível verificar que os elementos dc.creator (autor) faz link com recursos Web externos ao repositório (Figura 10), no caso a base do currículo Lattes dos envolvidos. Sendo indícios iniciais dessa característica.

Figura 10 - Uso de URI externa no repositório CarpeDien

dc.creator.email	lsayao@ien.gov.br	pt_BR
dc.creator.curriculum	http://lattes.cnpq.br/9090064478702633	pt_BR
dc.creator.curriculum	http://lattes.cnpq.br/3422623122948389	pt_BR
dc.relation.references	SALES, Luana. F. ; SAYÃO, Luís Fernando . Repositório do Instituto de Engenharia Nuclear: curadoria digital, publicações ampliadas e gestão de pesquisa. In: Congresso Brasileiro de Biblioteconomia e Documentação, 2013, Florianópolis. Anais do XXV CBBB, 2013. v. 25. Trabalho apresentado.	pt_BR

Fonte: CNEN, 2018

O último ponto, “assegurar que as pesquisas sejam encontradas” é atendido, visto que as publicações do CarpeDien são disponibilizadas de forma livre e são encontradas pelos usuários mesmo quando fazem uso do google. Ressalta-se que, durante a análise do repositório ocorreu de, algumas vezes, o site do repositório que se propõe a estar em uma plataforma online, aberta e disponível, se encontrou com problemas e fora do ar, impossibilitando que o usuário conseguisse navegar pelo repositório ou ter acesso ao seu conteúdo.

Como desfecho da análise do repositório, diferente da visão de Sales e Sayão (2014; 2016), o repositório CarpeDien precisa avançar ainda para ser considerado um repositório de publicações ampliada, pois o mesmo ainda não disponibiliza os principais critérios para se encaixar nesse campo, como a disponibilização dos recursos complementares e suas relações.

7 CONSIDERAÇÕES FINAIS

Considera-se que a preservação e disseminação da memória são preocupações da área de Ciência da Informação, e essa preocupação tem acompanhado o crescimento exponencial da produção científica, trazendo a tona a necessidade de amadurecimento de novas formas de armazenamento e articulação de documentos e o aprimoramento de metodologias, técnicas e ferramentas. Também traz a tona a necessidade de relacionar as informações e produtos da ciência, tais como os artigos, seus dados, teses e dissertações, entre outros, de forma que um o contexto de desenvolvimento da pesquisa possa ser compreendido e sua evolução acompanhada da melhor forma. Por isso mesmo, novas técnicas e ferramentas tanto para armazenamento de informações, tais como os repositórios de dados e as publicações ampliadas surgiram e vem se popularizando.

Porém, apesar de muitos autores considerarem as publicações ampliadas uma tendência para a comunidade acadêmica, atualmente, os repositórios que trabalham com a política de publicações abertas ainda não adotaram essa nova forma de publicação. Talvez por limitações técnicas, falta de conhecimento ou esperando um maior amadurecimento do uso dessa nova abordagem para a comunicação científica. A partir da investigação feita, é possível afirmar que ainda há um longo caminho a ser percorrido para a implantação e efetivo uso de publicações ampliadas, caminho esse que vem se delineando com a preocupação primeira de criação e estímulo ao uso de repositórios de dados de pesquisa que, em alguns momento, possam ser integrados a bases de eprints ou a elas relacionados. De fato, percebe-se que, tal qual houve um estímulo e investimento nas instituições para criação de seus repositórios de teses e dissertações e, posteriormente, de seus repositórios institucionais para disseminação do conteúdo produzido em seu contexto, começa a crescer o movimento de estímulo a criação e uso de repositórios de dados, onde, em breve, começarão a surgir as primeiras iniciativas. Por enquanto, apenas o protótipo da UNESP (VIDOTTI et al., 2017) e a Base de Dados Científicos da Universidade Federal do Paraná foram identificados. Adicionalmente, uma ação que vai contribuir com esse crescimento dos repositórios de dados é a exigência que as agências de fomento comecem a implementar de que o pesquisador deve gerenciar, preservar e tornar públicos os dados das pesquisas desenvolvidas, apresentando como isso será feito em planos de gestão de dados de pesquisa.

O uso de repositórios de publicações ampliadas pode trazer avanços para os pesquisadores e para a Ciência: o primeiro é a oportunidade de poder acompanhar de perto a

evolução das pesquisas, pois é possível visualizar as diversas versões da pesquisa até a sua conclusão. Isso também traz um ganho de tempo entre uma publicação e outra, pois a pesquisa pode ser postada e complementada posteriormente, não precisando esperar tanto tempo para disseminar as informações referentes à pesquisa. Além disso, agiliza a busca por detalhes da pesquisa, encontrando em um único lugar todas as informações necessárias sobre a mesma e se tendo, inclusive a possibilidade de reuso de dados para o avanço da Ciência.

Porém, como visto neste trabalho, durante o período de levantamento bibliográfico e coleta de dados da pesquisa efetivamente nenhum repositório de publicações ampliadas foi encontrado no contexto brasileiro. E, até os repositórios que são mencionados como contemplando publicações ampliadas, ainda estão no início do caminho, faltando as principais características de como disponibilizar junto ao documento final os documentos secundários/complementares que foram utilizados para a produção do mesmo ou a ele estão relacionados.

Com essa investigação, conclui-se que os objetivos geral e específicos da pesquisa foram alcançados, contribuindo assim para apresentar em que nível as instituições de pesquisa brasileiras se encontram com relação ao uso de publicações ampliadas e, paralelamente, na criação e uso de repositórios de dados de pesquisa. Assim, os resultados mostram que, apesar da tendência que se fortalece cada vez mais no contexto da ciência aberta de se criarem repositórios de dados abertos de pesquisa, esses repositórios, quando existem, ainda não estão integrados às publicações a eles relacionadas. Também destaca-se que ainda é insipiente a quantidade de repositórios de dados abertos de pesquisa brasileiros. Cenário esse que deve se modificar em alguns anos, devido às exigências crescentes das instituições de fomento a pesquisa pela criação e execução de planos de gestão de dados de pesquisa.

Indicam-se como trabalhos futuros, fazer o levantamento e avaliação dos repositórios que fazem uso de publicação ampliada ou o uso feito de repositório de dados de pesquisa fora do Brasil, analisando suas características e forma de funcionamento, de forma a apresentar contribuições para as futuras iniciativas brasileiras. Assim como repetir essa investigação em um intervalo de tempo, para analisar o crescimento do número de repositórios que atendam o enfoque da pesquisa.

REFERÊNCIAS

- ABAGLI, S.; APPEL, A. L.; MACIEL, M. L. **E-Science e ciência aberta: questões em debate.** In: XIV Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB 2013). Santa Catarina, 2013. Disponível em: <<http://enancib.sites.ufsc.br/index.php/enancib2013/XIVenancib/paper/view/168>>. Acesso em: 17 jan. 2018.
- ABBOT, D. **What is digital curation?** Edinburgh, UK: Digital Curation Centre, 2008. Disponível em: <<http://www.dcc.ac.uk/resources/briefing-papers/introduction-curation/what-digital-curation>>. Acesso em: 20 jul. 2017.
- ARAYA, E. R. M. **Comunicação científica: agregação, compartilhamento e reúso de elementos informacionais.** Tese — Marília: Universidade Estadual Paulista, 2014. 130 p.
- ARELLANO, M. A. M. **Crerios para a preservação digital da informação científica.** 2008. 354 f. Tese (Doutorado em Ciência da Informação) – Universidade de Brasília, Brasília, 2008. Disponível em: <<https://core.ac.uk/download/pdf/11884842.pdf>>. Acesso em: 10 jul. 2017.
- BANDEIRA, J. et al. Dados abertos conectados para a educação. In: IV Congresso Brasileiro de Informática na Educação. IV Jornada de Atualização em Informática na Educação. 2015. **Anais...** Alagoas, 2015. Disponível em: <<http://www.br-ie.org/pub/index.php/pie/article/view/3551>>. Acesso em: ago. 2017
- BANON, G. J. F.; RIBEIRO, M. L.; BANON, L. C. Contribuição ao livre acesso da literatura científica em ciência espacial: Implementação da política de auto-arquivamento na biblioteca digital do INPE. In: Seminário Internacional de Bibliotecas Digitais Brasil, 1 (SIBDB), 2007, São Paulo. **Anais...**, 2007. Disponível em: <<http://www.bibliotecadigital.unicamp.br/document/?down=23559>>. Acesso em: 20 jul. 2017.
- BAPTISTA, A. A.; COSTA, S. M. S.; KURAMOTO, H.; RODRIGUES, E. **Comunicação científica: o papel da Open Archives Initiative no contexto do acesso livre.** Encontros Bibli: Revista Eletrônica Biblioteconomia Ciência da Informação, Florianópolis, n. esp., 1. sem. 2007.
- BARDI, A.; MANGHI, P. **Enhanced Publications: Data Models and Information Systems.** LIBER Quarterly, [S.l.], v. 23, n. 4, p. 240-273, abr. 2014. Disponível em: <http://liber.library.uu.nl/index.php/lq/article/view/8445/9825>. Acesso em: 20 set. 2017.
- BERLIN **Declaration on Open Access to Knowledge in the Sciences and Humanities.** Berlin, 2003. Disponível em: <http://www.fu-berlin.de/sites/open_access/weiteres/Veranstaltungen/oa_berlin/poster/Berlin-Declaration_Simone-Rieger_MPIWG.pdf> . Acesso em: 20 jul. 2017.
- BERTIN, P. R. B.; VISOLI, M. C.; DRUCKER, D. P. **A gestão de dados de pesquisa no contexto da s-Science: benefícios, desafios e oportunidades para organizações de P&D.** Ponto de Acesso, Salvador, v.11, n.2, 2017. Disponível em: <<https://portalseer.ufba.br/index.php/revistaici/article/download/21449/15200>>. Acesso em: 14 jan. 2018.
- BLATTMANN, U. WEBER, C. DSpace como repositório digital na organização. **Revista ACB: Biblioteconomia em Santa Catarina**, Florianópolis, v.13, n.2, p.467-485, jul./dez., 2008. Disponível em: <<https://revista.acb.org.br/racb/article/view/593>>. Acesso em: 2 jul. 2017.
- BORGES, M. M. **A esfera: comunicação acadêmica e novos media.** Coimbra, 2006. Disponível em: <<https://estudogeral.sib.uc.pt/handle/10316/8557>>. Acesso em: jun. 2017.
- BORGMAN C. **Research data: who will share what, with whom, when, and why?** In: China-North American Library Conference, 5.Beijing, 17 aug. 2010. Disponível em:

<<http://works.bepress.com/borgman/238>> Acesso em: 20 set. 2017.

BRITO, V. M. A preservação da memória científica da Fiocruz: a visão de quem faz ciência. In: V Encontro Nacional de Pesquisa em Ciência da Informação (V ENANCIB), **Anais...**, Belo Horizonte (MG), 2003. Disponível em:

<<http://enancib.ibict.br/index.php/enancib/venancib/paper/viewFile/1928/1069>>. Acesso em: 02 nov. 2017.

CAFÉ, L. et al. Repositórios institucionais digitais: nova estratégia para publicação científica na Rede. In: Encontro Nacional de Informação em Ciências da Comunicação, 12, 2003, Belo Horizonte.

Anais... Belo Horizonte: [PUC MG], 2003. p. 1-12. Disponível em:

<http://www.intercom.org.br/papers/nacionais/2003/www/pdf/2003_ENDOCOM_TRABALHO_cafe.pdf>. Acesso em: jun 2017.

CAFE, L.; LAGE, M.B. Auto-arquivamento: uma opção inovadora para a produção científica.

DataGramaZero - Revista de Ciência da Informação, v.3, n.3, 2002. Disponível em:

<<http://basessibi.c3sl.ufpr.br/brapci/index.php/article/view/0000001051/f7cb919a0523ec46871ac8ff9c03139b>>. Acesso em: jul. 2017.

CATARINO, M. E. Simple Knowledge Organization System: construindo sistemas de organização do conhecimento no contexto da Web Semântica. **Informação & Tecnologia (ITEC)**, João Pessoa, v. 1, n. 1, p. 17-28, jan./jun. 2014. Disponível em:

<<http://periodicos.ufpb.br/index.php/itec/article/view/19307>>. Acesso em: ago. 2017.

CORUJO, L. M. N. Repositórios Digitais e Confiança: Um exemplo de repositório de preservação digital: o RODA. 2014. **Master thesis**, Faculdade de Letras da Universidade de Lisboa. [Thesis], 2014. Disponível em:

<<http://eprints.rclis.org/30375/>>. Acesso em: 20 set. 2017.

CUNHA, D. R.; SOUZA, D.; LÓSCIO, B. F. **Linked Data**: da Web de Documentos para a Web de Dados. Escola Regional Ceará, Maranhão e Piauí, 2011.

DELGADO, L. de A. N. **História oral** - Memória, tempo, identidades. Ed. Autentica, 2006.

DOBRTATZ, S. et al. **NESTOR catalogue of criteria for trusted digital repositories**. 2006.

Disponível em: <<http://www.dcc.ac.uk/resources/repository-audit-and-assessment/nestor>>. Acesso em: 20 set. 2017.

DODEBEI, V. Ensaio sobre memória e informação. **Morpheus** (UNIRIO. Online), v. 9, p. 227-244, 2016.

DODEBEI, V.; GOUVEIA, I. Memória do futuro no ciberespaço: entre lembrar e esquecer.

DataGramaZero: Revista de Ciência da Informação. v. 9, n. 5, out 2008.

DODEBEI, V. **Novos meios de memória: livros e leitura na época dos weblogs**. Enc. Bibli. R. Eletr. Bibliotecon. Ci. Inf. Florianópolis, 2009.

DONELLY, M. et al. **DRAMBORA Interactive**: User Guide. Glasgow, 2009.

FENTRESS, J.; WICKHAM, C. **Memória social**; novas perspectivas sobre o passado. Lisboa: Teorema, 1994. 278 p.

FERREIRA, C. A. S. **Preservação da Informação Digital**: uma perspectiva orientada para as bibliotecas [em linha]. Coimbra : [s.n], 2011. Dissertação de mestrado. Disponível em:

<<http://hdl.handle.net/10316/15001>>. Acesso em: 20 set. 2017.

FERREIRA, J.; AMARAL, A. Memória eletrônica e desterritorialização. **Revista Sociologia**, n. 4, p. 137-166, abr. 2004

FERREIRA, M. **Introdução à preservação digital** - Conceitos, estratégias e actuais consensos. Guimarães, Portugal: Escola de Engenharia da Universidade do Minho., 2006. Disponível em: <<https://repositorium.sdum.uminho.pt/bitstream/1822/5820/1/livro.pdf>>. Acesso em: 15 jun. 2017.

FERREIRA, M.; SARAIVA, R.; RODRIGUES, E. **Estado da arte em preservação digital**. Repositório Científico de Acesso Aberto de Portugal. 2011. Disponível em: <<http://repositorium.sdum.uminho.pt/handle/1822/17049>>. Acesso em: 25 jun. 2017.

FERREIRA, V. B.; VILLALOBOS, A. P. de O.; MOURA, M. P. **O modelo E-Science nos institutos nacionais de ciência e tecnologia de nanotecnologia**: evidências de práticas colaborativas. In: XVII Encontro Nacional de Pesquisa em Ciência da Informação (XVII ENANCIB), Bahia, 2016. Disponível em: <<http://www.ufpb.br/evento/lti/ocs/index.php/enancib2016/enancib2016/paper/view/4091>>.. Acesso em: 18 jan. 2018.

FIRMINO, H. N. M. **Organização e publicação dos termos do website da ANACOM sob uma perspectiva Linked Open Data**. Minho, 2013. Dissertação de mestrado. Disponível em: <<http://www.portaldoconhecimento.gov.cv/handle/10961/3367>>. Acesso em: 20 set. 2017.

FONSECA, J. J. S. **Metodologia da pesquisa científica**. Fortaleza: UEC, 2002. Disponível em: <http://leg.ufpi.br/subsiteFiles/lapnex/arquivos/files/Apostila_-_METODOLOGIA_DA_PESQUISA%281%29.pdf>. Acesso em: 18 mar. 2018.

GIL, A. C. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2002. Disponível em: <https://professores.faccat.br/moodle/pluginfile.php/13410/mod_resource/content/1/como_elaborar_projeto_de_pesquisa_-_antonio_carlos_gil.pdf>. Acesso em: set. 2017.

GONDAR, J. Cinco proposições sobre memória social. **Morpheus**: revista de estudos interdisciplinares em memória social, Rio de Janeiro, v.9, n.5, 2016.

GONDAR, J. Memória individual, memória coletiva, memória social. **Morpheus** – Revista Eletrônica em Ciências Humanas, ano 5, n. 09, 2006.

GONDAR, J. **Quatro proposições sobre memória social**. In: GONDAR, J.; DODEBEI, V. (Orgs.). O que é memória social? Rio de Janeiro: Contra Capa Livraria, 2009. p. 11-26.

GOUVEIA JUNIOR, M. Memórias e seus suportes: da fala à virtualização e suas necessidades por próteses e lugares. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v.9, n.2, p.64-74, jan./jun.2012. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/1911/pdf_20>. Acesso em: 3 jan. 2018.

HAKOPOV, Z.N. **Digital Repository as Instrument for Knowledge Management**. International Atomic Energy Agency, Vienna, 2016. Disponível em: <<http://eprints.rclis.org/29046/>>. Acesso em: jun. 2017.

HALBWACHS, M. **A Memória Coletiva**. São Paulo : Vértice, 1990. 189 p.

HALBWACHS, M. **Memória Coletiva e Experiência**. Tradução: Maria Luisa Sandoval Schmidt e Miguel Mahfoud. Psicologia USP, São Paulo p. 285-298, 1993. Disponível em: <<http://www.revistas.usp.br/psicousp/article/download/34481/37219>>. Acesso em: 15 dez. 2017.

HALBWACHS, M.. **A memória coletiva**. São Paulo : Vértice, 1990. 189 p.

HEY, A. J. G. (ED.). **The fourth paradigm**: data-intensive scientific discovery. Redmond, Washington: Microsoft Research, 2009. 251 p.

HEY, T.; TANSLEY, S.; TOLLE, K. Jim Gray on eScience: a transformed scientific method. In: HEY, T.; TANSLEY, S.; TOLLE, K. (Org.). **The Fourth Paradigm**: data-intensive scientific Discovery, 2009. Disponível em: <http://www.languagelog.ldc.upenn.edu/myl/JimGrayOnE-Science.pdf>. Acesso em: 20 jul. 2017.

HUNTER, J. Scientific publication packages – a selective approach to the communication and archival of scientific output. **The International Journal of Digital Curation**, v.1, n. 1, 2006. Disponível em: <http://www.ijdc.net/index.php/ijdc/article/view/8>. Acesso em: 20 jul. 2017.

INNARELLI, H. C. **Preservação de documentos digitais**: confiabilidade de mídias CD-ROM e CD-r. 174f. 2006. Dissertação (Mestrado) Faculdade de Engenharia Mecânica, Universidade Estadual de Campinas, 2006.

KOCHE, J. C. **Fundamentos de metodologia científica**: teoria da ciência e iniciação à pesquisa. Petrópolis, RJ: Vozes, 2011.

KURAMOTO, H. Informação Científica: proposta de um novo modelo para o Brasil. **Ciência da Informação**, Brasília, v.35, n.2, 2006. Disponível em: http://www.scielo.br/scielo.php?pid=S0100-19652006000200010&script=sci_abstract&tlng=pt. Acesso em: jul. 2017.

KURAMOTO, Hélio. Informação científica: proposta de um novo modelo para o Brasil. **Ciência da Informação**, [S.l.], v. 35, n. 2, aug. 2006. ISSN 1518-8353. Disponível em: <http://revista.ibict.br/ciinf/article/view/1144>. Acesso em: 28 may 2018.

LAGOZE, C.; SOMPEL, H. V. de. Ore User Guide – primer. **Open Archive Initiative**, 2008. Disponível em: <http://www.openarchives.org/ore/1.0/primer>. Acesso em: 20 jul. 2017.

LAMPERT, S. R.; FLORES, D. **O repositório digital como instrumento para preservação e acesso ao patrimônio arquivístico documental**. In: 7º Simpósio. Convenção do Patrimônio Imaterial: 10 anos depois [2003-2013]. **Anais...** Pelotas, 2013. Disponível em: <https://pt.scribd.com/document/279834364/O-Repositorio-Digital-Como-Instrumento-Para-Preservacao-e-Acesso-Ao-Patrimonio-Arquivistico-Documental>. Acesso em: jun. 2017.

LE COADIC, Y. F. **A ciência da informação**. 2. ed. Brasília, DF: Brinquet de Lemos livros, 2004.

LE GOFF, J. **História e memória**. Tradução: Bernardo Leitão. São Paulo: UNICAMP, 1990. Disponível em: <http://memorial.trt11.jus.br/wp-content/uploads/Hist%C3%B3ria-e-Mem%C3%B3ria.pdf>. Acesso em: 15 dez. 2017.

LEITE, F. C. L. **Como gerenciar e ampliar a visibilidade da informação científica brasileira**. Brasília: Ibict, 2009.

LIMA, M. H. T. de F. **Consequências do movimento pelo livre acesso – open access – e o direito à informação científica**. In: Implantação e gestão de repositórios institucionais: políticas, memória, livre acesso e preservação. Salvador: EDUFBA, 2009.

LYNCH, C. A. Institutional Repositories: essential infrastructure for scholarship in the digital age. **ARL Bimonthly Report**, n. 26. 2003. Disponível em: <http://www.arl.org/newsltr/226/ir.html>. Acesso em: 20 set. 2017.

LYNCH, C. **Institutional repositories: essential infrastructure for scholarship in the digital age**. 2003. Disponível em: <https://www.cni.org/publications/cliffs-pubs/institutional-repositories-infrastructure-for-scholarship>. Acesso em: 10 jun. 2017.

LYNCH, C. O quarto paradigma de Jim Gray e a construção do registro científico. In: HEY, T.; STEWARD, T.; TOLLE, K. (Org.). **O quarto paradigma**: descobertas científicas na era da eScience. Tradução Leda Beck. São Paulo: Oficina de textos, 2011. p. 187-193.

MARCONDES, Carlos Henrique. “Linked data” – dados interligados - e interoperabilidade entre arquivos, bibliotecas e museus na web. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, v. 17, n. 34, p. 171-192, maio/ago. 2012.

MÁRDERO ARELLANO, M. Á. **Critérios para a preservação digital da informação científica**. 354f. Tese (Doutorado em Ciência da Informação) -Universidade Federal de Brasília, Departamento de Ciência da Informação, 2008.

MARINHO, R. de B. **Arquitetura de informação para a web: projetando a experiência do usuário no Portal de Periódicos CAPES**. 2012. 142 p. Dissertação (Mestrado em Ciência da Informação) – Universidade Federal da Bahia. Salvador, 2012. Disponível em: <http://repositorio.ufba.br/ri/bitstream/ri/7825/1/Rafael%20Marinho%20-%20Dissertacao.pdf>. Acesso em: 05 dez. 2017.

MARTINS, E. **O repositório: imagem de marca e objeto de aprendizagem em meio digital**. In: III Conferência do IPCB sobre o livre acesso ao conhecimento científico. 2013 - Universidade Lusófona. Auditório Comenius, 2013. Disponível em <<http://repositorio.ipcb.pt/handle/10400.11/1739>>. Acesso em: jun. 2017.

MAYERNIK M et al. The data conservancy instance infrastructure and organization service for research data curation. **D-Lib Magazine** Sep/Oct 2012; 18(9/10). 2012. Disponível em: <http://www.dlib.org/dlib/september12/mayernik/09mayernik.html>. Acesso em: 20 set. 2017.

MELO, J. O. de S. F. **Metodologia de Avaliação de Qualidade de Dados no Contexto do Linked Data**. 2017. 111 f. Dissertação (Mestrado em Ciência da Informação) – Universidade Estadual Paulista, Marília, 2017. Disponível em: <<https://repositorio.unesp.br/handle/11449/150870>>. Acesso em: 10 ago. 2017

MIRANDA, A. C. U. C.; CARVALHO, A. A. V.; RAMOS, A. L. S. M. Portal de periódicos da capes: uma política pública de acesso à informação científica e tecnológica. **Biblionline**, v.11, n.1, p.185-204, 2015. Disponível em: <<http://periodicos.ufpb.br/index.php/biblio/article/view/25742>>. Acesso em: ago. 2017.

MONTEIRO, F.; BRASHER, M. Organização da informação em repositórios temáticos: o uso da modelagem conceitual. In: VIII ENANCIB - Encontro Nacional de Pesquisa em Ciência da Informação, 8, 2007, **Anais...** Salvador: ANCIB, 2007. Disponível em: <<http://www.enancib.ppgci.ufba.br/artigos/GT2--261.pdf>>. Acesso em: jun. 2017.

MUCHERONI, M. L.; SILVA, F. J. M. DA; PALETTA, C. F. Entre a publicação ampliada e a multimodalidade. XVI Encontro Nacional de Pesquisa em Ciência da Informação - XVI ENANCIB, 16. 2015. **Anais...** João Pessoa-PB: ANCIB, 2015. Disponível em: <http://www.ufpb.br/evento/ti/ocs/index.php/enancib2015/enancib2015>. Acesso em: 20 set. 2017.

MURRAY-RUST P. Open data in science. **Serials Review**, v. 34, n. 1, p. 52-64, 2008.

NEHMY, R. M. Q. et al. A ciência da informação como disciplina científica. **Perspectivas em Ciência da Informação**, Belo Horizonte, v.1, m.1, p.9-25, jan./jun., 1996.

NEVES, A.; GALINDO, M.; CUNHA, P. **Onde as ruas não têm nome: um olhar sobre o Ciberespaço**. Recife: UFPE, 1997.

NORA, P. **Entre Memória e História: A problemática dos lugares**. Tradução Yara Aun Khoury. Proj. História, São Paulo, 1993. Disponível em: <<https://revistas.pucsp.br/index.php/revph/article/viewFile/12101/8763>>. Acesso em: 15 dez. 2017.

OLIVEIRA, A. C. S. de; SILVA, E. M. **Ciência Aberta: dimensões para um novo fazer científico**. Inf. Inf., Londrina, v.21, n.2, 2016. Disponível em: <<http://www.uel.br/revistas/uel/index.php/informacao/article/view/27666>>. Acesso em: 15 jan. 2018.

ONLINE COMPUTER LIBRARY CENTER/ RESEARCH LIBRARY GROUP OCLC/RLG.

Implementing preservation repositories for digital materials: current practice and emerging trends in the cultural heritage - A Report by the PREMIS Working Group. September 2004, 66 p. Disponível em: <http://www.oclc.org/research/activities/past/orprojects/pmwg/surveyreport.pdf>. Acesso em: 15 jul. 2017.

OPEN ARCHIVES INITIATIVE. **Objective Reuse and Exchange**. Disponível em: <http://www.openarchives.org/ore>. Acesso em: 20 set. 2017.

PAVAN, C.; STUMPF, I. R. C. Avaliação pelos pares nas revistas brasileiras de ciência da informação: procedimentos e percepções dos atores. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, v.14, n.28, p.73-92, 2009. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2009v14n28p73>. Acesso em: jul. 2017.

PESET, F.; FERRER-SAPENA, A.; SUBIRATS-COLL, I. Open data y Liked open data: su impacto en el área de bibliotecas y documentación. **El profesional de la información**, 2011.

PIEDRA, N. et al. Linked Data en repositorios digitales de latinoamérica. In: IV Congresso Brasileiro de Informática na Educação, 4, 2015, **Anais...**Recife,2015.

POLLAK, M. Memória e identidade social. Tradução Monique Augras. **Estudos Históricos**, Rio de Janeiro, v. 5, n. 10, p. 200-215, 1992. Disponível em: <http://bibliotecadigital.fgv.br/ojs/index.php/reh/article/download/1941/1080>. Acesso em: 16 dez. 2017.

RIBEIRO JUNIOR, D. I. Repositórios de dados para E-Science: Open Data, Linked Data e suas tecnologias. **Ciência da Informação**, Brasília, DF, v.42 n.2, 2013. Disponível em: <http://revista.ibict.br/ciinf/article/view/1386/1564>. Acesso em: jun. 2017.

RIBEIRO, F. **Gestão da Informação / Preservação da Memória na era pós-custodial:** um equilíbrio precário. In: Faculdade de Letras da Universidade do Porto, Departamento de Ciências e Técnicas do Património; Coimbra, 2005.

RICHARDSON, R. J. **Pesquisa Social: métodos técnicos**. São Paulo: Atlas, 1999.

RLG/OCLC. **Trusted digital repositories:** attributes and responsibilities. Mountain View,CA. : RLG, OCLC, 2002.

RODRIGUES, F. de A.; SANT'ANA, R. C. G. Publicação ampliada no contexto de teses e dissertações. XVII Encontro Nacional de Pesquisa em Ciência da Informação – XVII ENANCIB 2016, 17, 2016. **Anais...** Salvador: ANCIB, 2016. Disponível em: <http://200.20.0.78/repositorios/handle/123456789/3610?show=full>. Acesso em: 20 set. 2016.

RODRIGUES, M. DE L. T. S. **Preservação digital de longo prazo:** Estado da arte e boas práticas em repositórios digitais. Lisboa, 2003. Disponível em: <http://dited.bn.pt/8927/37/Tese_Preservacao_Digital_Lurdes_Saramago.pdf>. Acesso em: jul. 2017.

RONDINELLI, R. C. **Gerenciamento arquivístico de documentos eletrônicos:** uma abordagem teórica da diplomática arquivística contemporânea. 4. ed. Rio de Janeiro: Editora FGV, 2005.

ROSSI, P. **O passado, a memória, o esquecimento:** seis ensaios da história das idéias. Tradução Nilson Moulin. São Paulo: UNESP, 2010. 240p.

RUEDA, V. M. da S.; FREITAS, A. de; VALLS, V. M. Memória Institucional: uma revisão de literatura. **CRB-8 Digital**, São Paulo, v..4, n.1, p.78-89, abr. 2011.

SALES, L. F. **Integração semântica de publicações científicas e dados de pesquisa**: proposta de modelo de publicação ampliada para a área de Ciências Nucleares. Tese—Rio de Janeiro: Universidade Federal do Rio de Janeiro, 2014. 265 p.

SALES, L. F. **Integração semântica de publicações científicas e dados de pesquisa**: proposta de modelo de Publicação Ampliada para a área de Ciências Nucleares. Tese (Doutorado em Ciência da Informação) – Universidade Federal do Rio de Janeiro, Instituto Brasileiro de Informação em Ciência e Tecnologia, Programa de Pós-Graduação em Ciência da Informação, Rio de Janeiro, 2014. 265 f.

SALES, L. F.; SAYÃO, L. F. Enhanced publication: a new model of scientific publication for the nuclear area. **IEN – Progress Report** 2013-2014, v. 2, n. 2015, p. 1, 2015. Disponível em: <http://revistas.ien.gov.br/index.php/ienprogressreport/article/view/131>. Acesso em: 20 set. 2017.

SALES, L. F.; SAYÃO, L. F. Modelo triádico de relações: um protótipo de modelagem conceitual para a área nuclear. XIII Encontro Nacional de Pesquisa em Ciência da Informação – XIII ENANCIB 2012, 13, 2012. **Anais...** Rio de Janeiro: ANCIB, 2012. Disponível em: <http://www.enancib2012.icict.fiocruz.br/>. Acesso em: 20 set. 2017.

SALES, L. F.; SAYÃO, L. F.; SOUZA, R. F. Modelo triádico de relações: um protótipo de modelagem conceitual para a área nuclear. XIV Encontro Nacional de Pesquisa em Ciência da Informação - XIV ENANCIB 2013, 14, 2013. **Anais....** Florianópolis: ANCIB, 2013. Disponível em: <http://enancib2013.ufsc.br/index.php/enancib2013/XIVenancib>. Acesso em: 20 set. 2017.

SALES, L. F.; SAYÃO, L. F.; SOUZA, R. F. Publicações ampliadas: um novo modelo de publicação acadêmica para o ambiente de e-science. 2013. In XIV Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB 2013), 14, **Anais...** Florianópolis, Brasil, 2013. [Presentation]. Disponível em: <http://eprints.rclis.org/20666/>. Acesso em: 15 ago. 2017.

SALES, L. F.; SOUZA, R. F.; SAYÃO, L. F. Publicação ampliada: um novo modelo de publicação científica voltada para os desafios de uma ciência orientada por dados. XV Encontro Nacional de Pesquisa em Ciência da Informação – XV ENANCIB 2014, 15, 2014. **Anais...** Belo Horizonte: ANCIB, 2014. Disponível em: <http://ridi.ibict.br/handle/123456789/824>. Acesso em: 20 set. 2017.

SANT'ANA, R. C. G. A importância do papel do profissional da ciência da informação nos processos de recuperação de conteúdos digitais estruturados. In: GUIMARÃES, J. A. C.; FUJITA, M. S. L. (Org.). **Ensino e pesquisa em biblioteconomia no Brasil**: a emergência de um novo olhar. Marília : Fundepe ; São Paulo : Cultura Acadêmica, 2008. p. 145-154.

SANT'ANA, R.C.G.; BONINI NETO, A. Análise de dados sobre produção de leite: uma perspectiva da Ciência da Informação. **Informação e Tecnologia**. v.1, n.1, jan/jun, 2014. Disponível em: <http://periodicos.ufpb.br/ojs/index.php/itec/article/view/19493/10863>. Acesso em: 20 ago.2017.

SANTOS, A. R. dos. **Metodologia científica**: a construção do conhecimento. 6 ed. rev. Rio de Janeiro: DP&A, 2006.

SANTOS, H. M. dos; FLORES, D. Repositórios digitais confiáveis para documentos arquivísticos: ponderações sobre a preservação em longo prazo. **Perspectivas em Ciência da Informação**, v.20, n.2, 2015. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/2341>>. Acesso em: jun. 2017.

SANTOS, H. M.; HEDLUND, D. C.; FLORES, D. Padronização dos formatos de arquivo: um caminho para preservar os documentos arquivísticos digitais. **Biblionline**, v.11, n.1, p.158-172, 2015.

SARMENTO, A. G. da S. Preservar para não restaurar. In: Simpósio Internacional de Propriedade Intelectual, Informação e Ética, 2003, Florianópolis. **Anais...** Florianópolis: Associação Catarinense de Bibliotecários, 2003. Disponível em:

<<http://siabi.trt4.jus.br/biblioteca/acervo/Produção%20Intelectual/Preservar%20para%20não%20restaurar.pdf>>. Acesso em: jun. 2017.

SAYÃO, L. F. **Bases de dados: a metáfora da memória científica**. Ci. Inf., Brasília, v. 25, n. 3, p. 314-318, set./dez. 1996

SAYÃO, L. F. Repositórios digitais confiáveis para a preservação de periódicos eletrônicos científicos. **Ponto de Acesso**, Salvador, v.4, n.3, p.68-94, dez., 2010. Disponível em: <<http://basessibi.c3sl.ufpr.br/brapci/index.php/article/view/0000009600/642405662ccac716ff70ad502506bf75>>. Acesso em: jun. 2017.

SAYÃO, L. F.; SALES, L. F. Curadoria digital: um novo patamar para a preservação de dados digitais de pesquisa. **Informação & sociedade: estudos**, João Pessoa, v. 22, n. 3, p. 179-191, set. dez. 2012. Disponível em: <http://www.ies.ufpb.br/ojs/index.php/ies/article/view/12224/8586>. Acesso em 20 nov. 2015.

SAYÃO, L. F.; SALES, L. F. Dados abertos de pesquisa: ampliando o conceito de acesso livre. **Revista Eletrônica de Comunicação, Informação & Inovação em Saúde**, v. 8, n. 2, p. 76–92, 2014. Disponível em: <http://www.reciis.icict.fiocruz.br/index.php/reciis/article/download/611/1252>. Acesso em: 20 set. 2016.

SAYÃO, L.F. Interoperabilidade das bibliotecas digitais: o papel dos sistemas de identificadores persistentes - URN, PURL, DOI, Handle System, CrossRef e OpenURL. **TransInformação**, Campinas, v.19, n.1, p. 65-82, jan./abr., 2007. Disponível em: <http://www.scielo.br/scielo.php?pid=S0103-37862007000100006&script=sci_abstract&tlng=pt>. Acesso em: ago. 2017.

SAYÃO, L.F.; SALES, L. F. Curadoria Digital: um novo patamar para preservação de dados digitais de pesquisa. **Inf. & Soc.:Est.**, João Pessoa, v.22, n.3, p. 179-191, set./dez. 2012.

SAYÃO, L.F.; SALES, L. F. Dados abertos de pesquisa: ampliando o conceito de acesso livre. **RECIIS – Rev. Eletron. de Comun. Inf. Inov. Saúde**. v. 8, n. 2, p. 76-92, jun. 2014.

SCHAFER, M. B.; CONSTANTE, S. N. E. Políticas e estratégias para a preservação da informação digital. **Ponto de Acesso**, v.6, n.3, 2012.

SENA, N. K. Open archives: caminho alternativo para a comunicação científica. **Ciência da Informação**, Brasília, v.29, n.3, p.71-78, set./dez. 2000.

SHINTAKU, M.; SUAIDEN, E. Repositório Institucional como componente de sistema de informação gerencial para universidades. **Biblos: Revista do Instituto de Ciências humanas e da informação**, v.29, n.1, 2015. Disponível em: <<https://www.seer.furg.br/biblos/article/view/4905>>. Acesso em: jun. 2017.

SILVA, F. B. M. E. Ciência da informação: uma ciência para a informação científica e tecnológica?. **Informação & Sociedade: Estudos**, v. 17, n. 3, p. 93-96, 2007. Disponível em: <http://www.brapci.inf.br/v/a/7627>. Acesso em: 28 Maio 2018.

SMITH, M. et al. DSpace: An Open Source Dynamic Digital Repository. **D-Lib Magazine**, v. 9, n. 1, January, 2003. Disponível em: <<http://dlib.org/dlib/january03/smith/01smith.html>>. Acesso em: 20 set. 2017.

SOUZA, Marylu Ferreira de. **Comunicação da informação científica em novos espaços de memória**. Dissertação (Mestrado) – Recife: Universidade Federal de Pernambuco, 2012. 148p.

- TARGINO, M. das G. **Comunicação Científica**: uma revisão de seus elementos básicos. v. 10, n. 2, 2000. Informação & Sociedade: Estudos. Disponível em: <<http://www.ies.ufpb.br/ojs2/index.php/ies/article/view/326/248>>. Acesso em: 20 dez. 2017
- TARGINO, M. das G. O óbvio da informação científica: acesso e uso. **Trasinformação**, Campinas, v. 19, n. 2, p. 95-105, maio/ago., 2007.
- THOMAZ, K. P. Repositórios digitais confiáveis e certificação. **Arquivística.net**, Rio de Janeiro, v.3, n.1, p.80-89, jan./jun., 2007.
- TOLLE, K.; TANSLEY, S.; HEY, T. Jim Gray e a eScience: um método científico transformado. In: HEY, T.; STEWARD, T.; TOLLE, K. (Org.). **O quarto paradigma**: descobertas científicas na era da eScience. Tradução Leda Beck. São Paulo: Oficina de textos, 2011. p. 17-29
- TRISKA, R.; CAFE, L. Arquivos abertos: subprojeto da biblioteca digital brasileira. **Ciência da Informação**, v.30, n.3, 2001.
- VAZ, G. J. **e-Science na Embrapa**. Campinas: Embrapa Informática Agropecuária, 2011. Disponível em: <<https://www.embrapa.br/busca-de-publicacoes/-/publicacao/921022/e-science-na-embrapa>>. Acesso em: 18 jan. 2018.
- VERHAAR P. Report on object models and functionalities. **DRIVER II**, 2007. Disponível em: https://openaccess.leidenuniv.nl/bitstream/handle/1887/16018/Report_on_Object_Models_and_Functionalities.pdf?sequence=1. Acesso em: 20 jul. 2017.
- VIANA, C. L. de M.; ARELLANO, M. A. M. **Repositórios Institucionais baseados em dspace e eprints e sua viabilidade nas instituições acadêmico-científicas**. 2006, Brasília. Disponível em: <http://eprints.rclis.org/8834/1/Trabalho_SNBU_RI_DSpace_EPrints_IES.pdf> . Acesso em: 09 nov. 2017.
- VIANA, C. L. M.; ARELLANO, M. A. M.; SHINTAKU, M. Repositórios institucionais em ciência e tecnologia: uma experiência de customização do DSpace. In: Simpósio Internacional de Bibliotecas Digitais, 2005, São Paulo. **Anais...** 2005. Disponível em: <<http://eprints.rclis.org/7168/>>. Acesso em: 12 jul. 2017.
- VIDOTTI, S. A. B. G.; CONEGLIAN, C. S.; ROA-MARTÍNEZ, S. M.; ARAKAKI, F. A.; BRANDT, M. B.; FERREIRA, A. M. J. F. DA C. Repositório de Dados de Pesquisa para Grupo de Pesquisa: um projeto piloto. XVIII Encontro Nacional de Pesquisa em Ciência da Informação - XVIII ENANCIB, 18. 2017. **Anais...** Marília-SP: ANCIB, 2017. Disponível em: <http://enancib.marilia.unesp.br/index.php/xviiienancib/ENANCIB/paper/view/388>. Acesso em: 03 jan. 2018.
- WEITZEL, S. da R. Fluxo da informação científica. In: POBLACION, D. A.; WITTER, G. P.; SILVA, J. F. M. da. (Orgs). **Comunicação & produção científica**: contexto, indicadores e avaliação. São Paulo: Angellara, 2006a. P. 81- 114.
- WEITZEL, S. da R. 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: <<http://seer.ufrgs.br/index.php/EmQuestao/article/view/19>>. Acesso em: jul. 2017.
- WEITZEL, S. da R. 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. 2006b.
- WEITZEL, S.; LEITE, F. C. L.; ARELLANO, Miguel A. M. E-LIS: um repositório digital para a biblioteconomia e ciência da informação no Brasil. In: XV SNBU, 15, 2008, Campinas. **Anais...** Campinas: Unicamp, 2008. Disponível em: <<http://repositorio.unb.br/handle/10482/12145>>. Acesso em: jul. 2017.

WEITZEL, S. da R. **Fluxo da informação científica**. In: POBLACION, Dinah Aguiar; WITTER, Geraldina Porto; SILVA, José Fernando Modesto. (Org.). Comunicação e produção científica: contexto, indicadores e avaliação. São Paulo: Angellara, 2006. p. 81-114.

WOUTERSEN-WINDHOUWER, S. et al. **Enhanced Publications Linking Publications and Research Data in Digital Repositories**. Amsterdam University Press, Amsterdam, 2009.