

# Os Repositórios digitais no âmbito da Sociedade Informacional

## Digital Repositories within the scope of informational society

Sílvia Mendes Masson

Mestranda em Ciência da Informação  
Escola de Comunicação e Artes da Universidade de São Paulo

masson.idt@terra.com.br

### Resumo

A disseminação de repositórios digitais emerge no contexto da chamada sociedade informacional e do informacionalismo (Castells:2000), em meio ao advento de novas disciplinas e das tecno-ciências, que atuam na interdisciplinaridade, compartilham objetos de estudo e em que, freqüentemente, ocorre a apropriação mútua de conceitos e de terminologia, em virtude da dinâmica das mudanças científicas, tecnológicas e sociais em ritmo que se acelera à medida que se afastam das sociedades industriais. Este artigo pretende mergulhar na diversidade de conceitos e usos da expressão “repositório digital” e, com isso, refletir sobre como surge este serviço, ora considerado específico de uso na comunidade científica e de universidades, ora um serviço passível de ser utilizado em organizações privadas ou públicas, nas modalidades de repositórios institucionais e temáticos.

Palavras- chave: Repositório Digital. Arquivo/biblioteca. Iniciativa dos Arquivos Abertos. Auto-arquivamento

### Abstract

*The dissemination of digital repositories raises in the scope of informational society and informationalism (Castells: 2000), within the arrival of new subjects and techno-sciences, which acts in the interdisciplinarity, shares objects of study and, frequently, appropriates concepts and terminology mutually, in force of the dynamics of the technological and social changes, in an accelerating pace in proportion to the spacing from industrial societies.*

*This article aims to dive in the diversity of concepts and uses of the expression “digital repository” and, then, to reflect about on how this service/institution is born, sometimes considered specific to be used in scientific community and universities, sometimes a service which can be used in public or private organizations, creating the types of thematic and institutional repositories.*

Keywords: Digital Repository. Archives/Library. Initiative Open Archives. Self-archiving.

## 1. Introdução e justificativa para uma “arqueologia”<sup>1</sup> dos repositórios digitais

Com a explosão de informação e a ampliação do conhecimento técnico e científico, houve também uma explosão da terminologia necessária para comunicar o conhecimento. No último século, a emergência de novos saberes e a diversidade de termos, surgidos em razão das novas tecnologias e áreas do conhecimento, resultou na necessidade de ampliação do vocabulário terminológico, provocando, muitas vezes, polissemia, em razão da ampla variedade de conceitos e de termos que se tornam partilhados entre as disciplinas e áreas do conhecimento, porque as áreas interdisciplinares recorrem, com frequência, à apropriação mútua de conceitos e de terminologia, ou ao uso de um mesmo signo, porém com conteúdos semânticos distintos, nos diferentes campos referenciais científicos, cujas fronteiras não são precisamente demarcadas, como acontece com a ciência tradicional. Um outro fator que leva a “babelização” é a necessidade de usar metáforas, ou terminologia de outras áreas do conhecimento, notadamente das chamadas ciências exatas, adaptando-a para designar fenômenos e conceitos operatórios em ciências sociais e humanas.

É o caso da palavra “repositório”, objeto deste artigo, usado com frequência quando se trata de preservação e custódia de informação e/ou bens culturais. Partindo de uma “arqueologia” da palavra repositório – paráfrase que adotamos para construir a argumentação e as reflexões que serão seguidas -, investigar-se-ão os vários significados em que tem sido usada.

O termo “repositório”, no sentido dicionarizado, é aquilo “que é próprio para guardar alguma coisa; depósito” e, tem, como sinônimo, também os significados de “repertório, coleção”. (Buarque de Hollanda, s/d: 1219) O verbete, “repertório”, do latim “repertoriū”, é usado em sinonímia a “catálogo seletivo” e a “inventário”, em terminologia Arquivística, e também é terminologia biblioteconômica. O significado vernacular, no

---

<sup>1</sup> A metáfora se refere à proposta da Antropologia, em que a arqueologia tem como função analisar o objeto (neste caso, os repositórios digitais) como um instrumento para compreender o pensamento (neste caso, os conceitos operatórios), os valores e a própria sociedade (neste caso, a instituição ou serviço, repositório digital). Inspirado em Foucault, na designação da metodologia proposta pelo autor, em “As palavras e as coisas: uma arqueologia das ciências humanas”, em que o fenômeno cultural é considerado como “*realidade de estrutura e enquanto estrutura, os elementos são leitura do todo e o todo se lê na trama que eles desenham.*” (Foucault, s/d:VII) No entanto, a metáfora da palavra arqueologia, usada no artigo, se refere apenas ao sentido metafórico de “escavar e buscar vestígios” e não na complexa proposição de Foucault. O próprio Foucault, na contra-capá de outra obra, “A arqueologia do saber”, referindo-se à palavra arqueologia, dizendo que a havia deixado vazia: “*Palavra perigosa, pois parece evocar rastros caídos fora do tempo e petrificados, agora, em seu mutismo. Na verdade trata-se de descrever discursos. Não livros (na relação com seus autores), não teorias (com suas estruturas e coerência), mas os conjuntos, no mesmo tempo familiares e enigmáticos, que através do tempo, se tornam conhecidos (...)*” (Foucault, 1972).

dicionário, é registrado como “coleção, compilação, conjunto” (Idem: 1218), mantendo, portanto, similaridade de sentido com o significado terminológico. Isto ocorre, como nos referimos, com as terminologias das disciplinas em ciências sociais e humanas, porque freqüentemente são usadas palavras de uso corrente, registradas no dicionário e no uso coloquial, re-significadas e utilizadas na comunicação científica.

A palavra “repositório” aparece, como verbete, no Dicionário de Terminologia Arquivística (Camargo e Bellotto, 1996: 67): “*Repositório Ver: depósito (2).*” No verbete “depósito”, temos o sentido 1: “*Ação pela qual os documentos são colocados sob a custódia de uma instituição arquivística.*” No sentido 2, indicado pelo verbete “repositório”, assim está definido: “*Prédio ou área para armazenamento*”. (Idem, 1996: 22). No entanto, nos parece que o significado mais aproximado, pelo menos para o sentido que está implícito na locução “repositório digital”, se substituirmos “obra impressa” por “obra digital”, é o encontrado no verbete “*depósito legal: obrigação legal de remeter a bibliotecas, e excepcionalmente a arquivos públicos, um ou mais exemplares de cada obra impressa no país, estado ou município.*” (Ibidem, 1996: 22)

Na edição portuguesa, denominada “Dicionário do Livro: terminologia relativa ao suporte, ao texto, à edição e encadernação, ao tratamento técnico, etc” (Faria:1988) não consta o verbete repositório.

O “Dicionário de terminologia arquivística” (Alves et al:1993:31), publicado em Lisboa, pelo Instituto da Biblioteca Nacional e do Livro, também não registra o verbete repositório. Verificando o verbete “depósito legal” (com o qual encontramos a similaridade de função com repositório, ao falarmos da terminologia em Carmargo e Bellotto), no Dicionário do IBNL, encontramos os conceitos: “*Depósito legal 1. Entrega obrigatória e gratuita de um ou vários exemplares de toda e qualquer publicação impressa num País a uma ou mais instituições públicas para tal designada na lei. 2. Entrega imediata, obrigatória e gratuita de originais de determinados documentos de arquivo a uma instituição pública designada na lei para sua depositária (ex: leis, acordos, etc.) v. tb Aquisição p. 31.*” No segundo sentido, temos o significado de depósito que nos parece aproximar-se à sinonímia de repositório, e a repositórios digitais, se estivermos pensando apenas na designação de um “depósito” para os chamados “objetos digitais”. As especificidades de “repositório digital”, como serviço nascido no seio da comunicação científica, serão atingidas ao longo destas reflexões.

No “Diccionario de Terminología Archivística” (Ministerio de Cultura: Subdirección General de Los Archivos Estatales, 1995) da Espanha, não aparece o termo “repositório”, mas o verbete “depósito”, no segundo sentido, define: “(2) *Procedimiento de ingreso de fondos documentales en un archivo mediante una forma jurídica en virtud de la cual el depositario (el archivo) se compromete a guardarlos, sin que ello implique adquisición de uso, y, en su caso, a restituirlos (depósito voluntario). (...)*”. Ressalte-se que o depósito, no sentido de ingresso de fundos documentais em um arquivo, guarda semelhança com a idéia de ingresso de artigos, teses, livros e trabalhos, produzidos em meio digital ou digitalizados, no repositório digital.

O “Diccionario del archivero bibliotecario: terminología de la elaboración, tratamiento y utilización de los materiales propios de los centros documentales” apresenta o verbete “*repositorio archivístico: lugar donde se guarda un archivo*” (Ejarque, 2000: 390), e, no verbete, amplia: “*repositorio bibliográfico: biblioteca*” e “*repositorio documental: archivo*”. Nos conceitos assim definidos, faz sentido o paralelismo: “repositório digital: repositório de objetos digitais”, mas, sem dúvida, tautológico.

Não há o registro do verbete “*repositorio*”, no Diccionario enciclopédico de Ciencias de la Documentación de López Yepes (2004).

Para prosseguir na investigação se a função do depósito legal poderia ter paralelo com a função dos repositórios digitais, consultamos a Lei Federal (brasileira) n° 10.994, de 14 de dezembro de 2004, que trata da obrigatoriedade de todas as publicações serem encaminhadas ao Depósito Legal, na Fundação Biblioteca Nacional. O artigo 8º, esclarece que esta Lei não se confunde com o registro de obras intelectuais pelos autores ou cessionários, conforme o disposto, respectivamente, nos artigos 17 e 53, da Lei n° 5.988, de 14/12/1973.

No *site* da Biblioteca Nacional, há a citação do Artigo 2º da Lei n° 10.094/2004, com a definição de “depósito legal” como sendo uma exigência, por força de lei, de remessa à Biblioteca Nacional de um exemplar de todas as publicações produzidas em território nacional, por qualquer meio ou processo. O objetivo principal do Depósito Legal é assegurar a coleta, a guarda e a difusão da produção intelectual brasileira, visando à preservação e formação da Coleção Memória Nacional<sup>2</sup>. Neste caso, se a Lei já prevê o arquivamento, ou como está na Lei, o depósito, que torna a Coleção Memória Nacional o repositório de toda a informação publicada, no País. O repositório de informação digital, a

<sup>2</sup> Disponível em: <http://www.ibict.br/secao.php?cat=ISSN/Links> Acesso em 02/01/08.

seguir o raciocínio, estaria incluído, mas, na prática, não significa a mesma concepção do “repositório digital”, “temático” ou “institucional”, a que chegaremos mais adiante.

Cacaly (1997, 176-179), no “Dictionnaire encyclopédique de l’information et de la documentation”, no verbete “depósito legal”, traz uma detalhada explanação das funções e razões para o depósito legal e as diferenças de um país a outro. Sublinha que há dois objetivos principais: a conservação e o controle das publicações, incluindo os materiais audiovisuais e as edições eletrônicas em suporte material. Adverte que o depósito das publicações eletrônicas e dos audiovisuais deverá ser regulamentado em nível internacional. Segundo a lei francesa, são passíveis de depósito, em França:

“Le dépôt legal est régi en France par la loi 92-546 du 20 juin 1992 et par le décret 93 1429 de 31 décembre 1993. La loi de 1999 remplace celle du 21 juin 1943 qui concernait « les imprimés de toute nature, livres, périodiques, brochures, estampes, gravures, cartes postales illustrées, affiches, cartes des géographie et autres, les oeuvres musicales, photographiques, mises publiquement en vente, en distribution ou en location ou cedées pour la reproduction ».” (Cacaly :1997:179)

A Lei de “depósito legal”, no Brasil, é mais abrangente, ao incluir todas as publicações em qualquer meio ou processo, o que acaba por incluir as publicações em meio digital.

O repositório digital (conforme mais adiante será apresentado, no Quadro 2, que traz conceitos e definições por diversos autores e instituições) parece ter como objetivo preservar as publicações criadas em meio digital, ou as que são digitalizadas, referentes a artigos em periódicos, ou na *web*, os próprios periódicos, as atas de congresso e outros eventos, trabalhos de pesquisas, teses e dissertações para que constituam uma memória institucional e que possam estar disponível e acessível a quem precisar consultá-la, ou, no caso dos repositórios digitais temáticos, a memória do conhecimento em uma área especializada do conhecimento. Assemelha-se, assim, a uma antiga prática, surgida em França, por Francisco I, em 1537, que decretou a proibição da venda de qualquer livro sem que primeiro tivesse sido depositado um exemplar em sua biblioteca. O decreto não foi devidamente respeitado, mas o princípio de guardar para a memória coletiva foi estabelecido e utilizado em outros países, inclusive, em algumas ocasiões, este princípio foi usado como forma de praticar a censura, ou o controle e a vigilância sobre os conteúdos em circulação, e, por um outro lado, como forma de proteção aos direitos do autor. Ampliou-se, ao longo do tempo, a prática do depósito, com a abrangência para outros suportes de informação resultantes do avanço tecnológico, para além do livro, como os

audiovisuais e, mais recentemente, as publicações eletrônicas, assim como especifica o “*dépôt legal*”, praticado em França.

A locução “repositório digital” é formada pelo substantivo repositório, entendido no seu sentido vernacular, como depósito e coleção, mais o adjetivo digital, que significa “tradução do mundo analógico, em dígitos”, resultando na expressão designativa de um serviço para atender à demanda de soluções para o tratamento da informação, intensivamente produzida no informacionalismo<sup>3</sup>. Do que foi exposto até aqui, percebe-se que os significados, tanto o dicionarizado, quanto os usados nas terminologias de áreas da Ciência da Informação, são aproximados, mas impõe construir o conceito, ou através de conceitos operatórios, e/ou através da fixação de objetivo e função de repositórios digitais, no contexto em que são instituídos.

Um exercício complementar, para a compreensão e a investigação da origem da expressão “repositórios digitais”, foi feito a partir de uma obra reconhecidamente importante para a Biblioteconomia, denominada “*A conturbada história das bibliotecas*”, de Battles (2003), em que o autor relaciona conhecimentos técnicos sobre os sistemas de classificação e de organização de uma biblioteca moderna com a pesquisa histórica sobre armazenamento dos livros. O objetivo é demonstrar os usos possíveis e distintos de conteúdos semânticos do vocábulo “repositório” compilados no Quadro 1, abaixo, em que se destacam os usos e os respectivos significados:

**Quadro 1: Os Significados do Termo “Repositório” na Obra de Battles**

Localização	Transcrição do trecho	Contexto/Significação
Capítulo 1 (Battles, 2003: 11 )	<i>“Mas a biblioteca – especialmente uma tão vasta – não é um mero repositório de curiosidades. É um mundo a um só tempo completo e incompletável, cheio de segredos”.</i>	O autor está a se referir à Biblioteca de Widener e a fazer outros comentários a respeito de particularidades do acervo. Aqui, o sentido é de coleção e de depósito, que abrigam curiosidades. Afirma que a biblioteca não é apenas um repositório de curiosidades, portanto, por dedução, biblioteca é também repositório.
Capítulo 2 (Battles, 2003: 30)	<i>“Na já então antiga cidade de Nínive, capital do império, ele [refere-se a Assurbanipal II] organizou uma grande biblioteca, que chegou a abrigar 25 mil placas. Embora essa biblioteca fosse antes de mais nada um arquivo, Assurbanipal tinha aspirações universais”.</i>	A propósito das bibliotecas da Mesopotâmia, o autor refere-se aos “livros”, gravados em argila, em escrita cuneiforme traçada por agulhetas, e que acabaram dando impulso à construção de bibliotecas e/ou arquivos, e ora usa biblioteca, ora usa arquivo para se referir ao “depósito” que abrigou as 25 mil placas de argila. Sublinhe-se a dificuldade em designar o “repositório” que abrigou as placas entre arquivo e biblioteca.
Capítulo 2	<i>“Outros arquivos e bibliotecas</i>	Este trecho segue o raciocínio da citação

<sup>3</sup> Castells assim designa a nova estrutura social, manifestada sob diversas formas, que depende da diversidade de culturas e instituições existentes em todo o planeta, associada à emergência de um novo modo de desenvolvimento, fundamentado na informação, e historicamente moldado pela reestruturação do modo capitalista de produção, no final do século XX. (Castells, 2000:17).



Localização	Transcrição do trecho	Contexto/Significação
(Battles, 2003: 31)	<i>espalhados pela Mesopotâmia exibiam níveis igualmente elevados de organização. Havia repositórios em que as placas eram guardadas em cestas numeradas, com os títulos gravados nas bordas da argila para facilitar a identificação”.</i>	anterior. Nele, o vocábulo repositório está usado como sinônimo de “depósito”, de lugar que serve para guardar alguma coisa, além de parecer importante também destacar que o repositório, ou o depósito para guardar as tabuinhas em argila (“documentos” ou “livros”), seriam “biblioteca e arquivo”.
Capítulo 2 (Battles:2003:36)	<i>“(…) como um repositório universal destinado à preservação da aprendizagem liberal, ainda que nossos caros mitos de origem nos levem a pensar assim. As bibliotecas estão envolvidas tanto na descoberta da verdade quanto em sua destruição, para satisfazer o espírito bárbaro dos príncipes, presidentes e aspirantes ao trono.”</i>	Em continuação à exposição da história das bibliotecas, fala-nos da influência de Alexandria, nos primeiros séculos de nossa era, em que o grande estoque de livros aí reunidos definiu uma nova concepção a respeito do valor do conhecimento, tendo como objetivo reunir tudo o que estivesse disponível e, ao patrocinar esse objetivo, os Ptolomeus confirmavam a intuição alexandrina de que o conhecimento era um bem a ser capitalizado e entesourado. O destino e a sobrevivência dos livros estavam mais relacionados à constante transformação da palavra, às necessidades e às preferências de leitores privados e de colecionadores, do que às chamadas, ao roubo e à censura. A política de aquisições dos Ptolomeus acontece sem que eles valorizassem a biblioteca. Neste ponto, o autor usa o vocábulo repositório com o significado de depósito, de lugar de armazenagem, e como sinônimo de biblioteca.
Capítulo 3 (Battles, 2003:69)	<i>(…) “Para os gregos e romanos, os livros tinham sido instrumentos, repositórios utilitários de conhecimento” (...),</i>	No contexto desta citação, o autor nos revela que acredita que muito da cultura livresca ocidental teve sua origem no Islã, uma vez que os muçulmanos, ao dominar outros povos, sempre demonstraram uma grande capacidade de assimilação: no século oitavo, aprenderam a confeccionar papel com prisioneiros chineses, com os escribas etíopes aprenderam o códice e elevaram a técnica de confeccionar capas em couro, em arte sofisticada. A citação é a continuidade desta idéia aqui sintetizada. Aqui usa a palavra “repositório” para se referir aos livros em rolo, por serem os rolos econômicos e simples. Os calígrafos e ilustradores do Islã transformaram o livro num objeto belo em que o conteúdo e a aparência suntuosa eram igualmente apreciados. Repositório, aqui, está relacionado ao suporte da informação/conhecimento, neste caso, os livros.
Capítulo 3 (Battles 2003:78-80)	<i>“Beddie lembra, que ‘na biblioteca de Lorsh’, no século X, de um total de 590 volumes, 98 eram de Agostinho; na de Bec, no século XII, havia 36 volumes de Agostinho; e dos 184 manuscritos existentes no Mosteiro de São Maurício, em Naumburg, 98 eram livros de Agostinhos. Esses números sugerem que as grandes</i>	Ao se reportar às bibliotecas da Idade Média, cita James Stuart Beddie, historiador que compilou catálogos medievais e que revela que as obras de Agostinho, depois da Bíblia, se constituíam na parte principal de uma biblioteca típica da Idade Média. Nesta citação, há o uso da palavra “repositório”, como metáfora para a mesma função dos livros, guardar conhecimento, e está a atribuir às bibliotecas a função de depósito, repositórios para guardar os

Localização	Transcrição do trecho	Contexto/Significação
	<i>bibliotecas serviam como repositórios de exemplares que seriam emprestados para cópia às bibliotecas menores. Talvez mostrem, também, que a tarefa de copiar livros, que era atribuída aos monges, servia tanto para instruí-los quanto para aumentar o tamanho das bibliotecas.”</i>	exemplares a serem emprestados para a feitura de cópias. Subsiste, em ambos os usos, a idéia de “coleção de conhecimento”, seja em formato de livro, ele próprio, “repositório de conhecimento”, ou em coleções de livros, “repositórios de livros” como sinônimo de biblioteca.

*(Elaborado, pela autora, a partir de Battles, 2003.)*

A palavra repositório é de uso freqüente, em seu significado vernacular, nas instituições que custodiam ou guardam acervos, ou em textos de comunicação científica, como destacamos em relação ao uso, como aparece na obra de Battles, cuja versatilidade polissêmica conservou sempre uma relação com a idéia de depósito ou coleção.

Ultimamente, o termo repositório tem sido freqüentemente usado para designar o armazenamento de objetos digitais, aparecendo na literatura de Ciência da Informação e Ciências da Comunicação, bem como nos *sites*, principalmente os da área jurídica<sup>4</sup> e da área da saúde.

Esta tentativa de realizar uma “arqueologia” do uso do vocábulo repositório, seja no significado vernacular, ou no terminológico de disciplinas no âmbito da Ciência da Informação, e do uso contemporâneo, em razão do surgimento de um número crescente de repositórios digitais, disponibilizados na *Net*, decorre de algumas indagações, referentes aos conceitos operatórios neles utilizados e que os distinguiriam, em tese, de outros serviços e/ou sistemas de informação.

Repositórios digitais são freqüentemente conceituados em relação às suas funções de reunir, preservar, dar acesso e disseminar o conhecimento de uma instituição científica, ou de uma área do conhecimento, aumentando sua visibilidade e se constituindo numa ferramenta de gestão do conhecimento científico. Assim conceituado, fica restrito apenas a tratar da informação científica, como geralmente está definido por grande parte dos autores que entendem ser objeto de estudo da Ciência da Informação tão somente a informação científica, em razão da exclusividade do tratamento biblioteconômico da comunicação científica.

<sup>4</sup> Repositório digital da área Jurídica, sem vínculo com o OAI-PMH: “A Biblioteca Digital Jurídica é um repositório de documentos jurídicos em inteiro teor, tais como capítulos de livros e periódicos, teses, dissertações, decisões jurisprudenciais, palestras e discursos.” Disponível em: [http://www.stj.gov.br/portal\\_stj/publicacao/engine.wsp?tmp.area=353](http://www.stj.gov.br/portal_stj/publicacao/engine.wsp?tmp.area=353) Acesso em 10/01/08.



Repositório digital é um serviço de informação? Ou um sistema de informação? É uma ferramenta de gestão do conhecimento, mas apenas do conhecimento científico? Ou é uma coleção de trabalhos de pesquisa de uma determinada área do conhecimento, ou de uma instituição, disponibilizada na *Net*? Biblioteca digital é um repositório. E repositório digital é um serviço da biblioteca digital?

A busca de respostas a estas indagações leva-nos a prosseguir as reflexões, agora, tendo como objeto o próprio repositório digital: como surge, em que contexto, qual a finalidade e como se consolidam os seus usos.

## **2. A interação entre Sociedade, Informação, Comunicação e Tecnologia: o contexto dos repositórios digitais na sociedade informacional**

Sociedade, informação e tecnologia constituem aspectos de uma interação complexa, cujos conceitos, isoladamente, exigem a evocação dos demais e ainda fazem emergir outros, como ciência, cognição, cultura, representação, e, entre as formas de representação, a linguagem e, através dela, a comunicação, como interação humana e social. Não é possível pensar a sociedade humana, sem conhecimento e informação, sem tecnologia e sem comunicação. Refletir sobre informação, além da comunicação, importa saber o contexto de produção, avaliar e conhecer as necessidades de armazenamento, recuperação, acessibilidade, geração de informação/conhecimento. Ao fazermos isto, necessariamente, estaremos também a falar de sociedade, cultura e tecnologia.

Castells (2002:6) pondera que a tecnologia não determina a sociedade, mas a incorpora, e a sociedade não determina a inovação tecnológica, mas a usa. Sociedade e tecnologia interagem, dialeticamente, porque a *“tecnologia é a sociedade e a sociedade não pode ser compreendida ou representada sem as suas ferramentas tecnológicas”*.

Ao falarmos de sociedade, informação e tecnologia, estamos também a falar de cultura. Morin (2003:159), ao conceituar cultura, abarca os aspectos que nos parecem importantes para o raciocínio que queremos aqui desenvolver. Para Morin, cultura é:

“a emergência fundamental própria da sociedade humana. Cada cultura concentra em si um duplo capital: por um lado, um capital cognitivo e técnico (práticas, saberes, saber-fazer, regras), por outro, um capital mitológico e ritual (crenças, normas, proibições, valores). É um capital de memória e de organização, como é o patrimônio genético para o indivíduo. A cultura dispõe, como o patrimônio genético, de uma linguagem própria (mas muito mais diversificada), que permite a rememoração, a comunicação, a transmissão deste capital de indivíduo para indivíduo e de geração em geração. O patrimônio hereditário dos indivíduos está gravado no código genético; o patrimônio

cultural herdado está gravado, em primeiro lugar, na memória dos indivíduos (cultura oral), depois escrito na lei, no direito, nos textos sagrados, na literatura e nas artes. Adquirida em cada geração, a cultura é continuamente regenerada. Constitui o equivalente a um Genos sociológico, ou seja, um engrama-programa, que garante a regeneração permanente da complexidade social.” (Morin, 2003:159)

Quanto à informação, há uma diversidade de definições e conceitos, mas, neste caso, usaremos a definição dada por Silva (2006: 25), porque harmoniza com o raciocínio, que está aqui sendo seguido, de que sociedade, informação, tecnologia e cultura modelam a comunicação e a sociedade na trajetória histórico-sociológica da humanidade. Para Silva, informação é:

“o conjunto estruturado de representações mentais e emocionais codificadas (signos e símbolos) e modeladas com/pela interação social, passíveis de serem registradas num qualquer suporte material (papel, filme, banda magnética, disco compacto, etc.) e, portanto, comunicadas de forma assíncrona e multi-direccionada.”

Informação, nesta concepção, como fenômeno humano e social, é abrangida pela cultura e ambas são indissociáveis e delas decorrem, como emanção prática, as manifestações dos fluxos da informação e dos registros que transportam a informação e o conhecimento, que acompanham o homem e suas necessidades de produzir, transformar, utilizar, comunicar, transmitir e preservar o conhecimento/informação.

A capacidade de comunicação humana, por suportes em diversos meios tecnológicos, atravessando o tempo e o espaço, desempenharam papel capital na evolução da humanidade, substituindo a tradição oral, e criando uma memória coletiva fora das mentes dos membros individuais do grupo, garantindo que estará em lugar seguro, disponível à consulta quando for preciso.

Os repositórios da memória, como instituições inseridas na sociedade, agem, reagem e interagem às tecnologias e à cultura, respondendo às evoluções e revoluções tecnológicas. O desenvolvimento tecnológico, a partir da revolução industrial, torna-se a mola propulsora de um processo de tecnificação do mundo e de cientificação das atividades sociais, que ficou conhecido como revolução tecnológica.

Segundo Castells (2002: 36) evocando a outros autores,

“o que caracteriza a actual revolução tecnológica não é a centralidade do conhecimento e da informação, mas a aplicação deste conhecimento e informação na produção de conhecimento e de dispositivos de processamento/comunicação da informação, num ciclo de realimentação cumulativo entre a inovação e o seu uso.”<sup>5</sup>

---

<sup>5</sup> O autor teve a preocupação de definir o que entende por conhecimento e informação. Adotou a definição de conhecimento dada por Daniel Bell (1973:175): “*Conhecimento: um conjunto de declarações organizadas sobre factos ou ideias, apresentando um julgamento ponderado ou resultado experimental que é transmitido a outros por intermédio de algum meio de comunicação, de alguma forma sistemática. Assim, diferencio*

A partir de uma adaptação da análise clássica das revoluções científicas de Kuhn, Carlota Perez, Christopher Freeman e Giovanni Dosi – citados por Castells (2002: 86) – criaram o conceito de paradigma tecnológico que auxilia a compreender a transformação tecnológica atual, uma vez que interage com a economia e a sociedade e representam a base material da sociedade de informação e o paradigma da tecnologia da informação. Castells (2002:87) destaca as características deste novo paradigma:

1º) A informação é a sua matéria-prima, porque “*são tecnologias para agir sobre a informação, não apenas informação para agir sobre a tecnologia*”, como aconteceu nas revoluções tecnológicas anteriores;

2º) O novo meio tecnológico molda todos os processos da existência individual e coletiva, pois a informação é uma parte integral de toda a atividade humana. Ou seja, grande “*capacidade de penetração dos efeitos das novas tecnologias.*”;

3º) *Lógica de redes* que pode ser implementada em qualquer sistema ou conjunto de relações, usando as novas tecnologias. Esta lógica de redes é necessária para estruturar o não-estruturado, preservando a flexibilidade, porque o não-estruturado é a força motriz da inovação na atividade humana;

4º) O paradigma da tecnologia da informação é baseado na *flexibilidade*, em razão de sua capacidade de reconfiguração, perfeito para uma sociedade caracterizada pela constante mudança e fluidez organizacional. Castells (2002: 88) alerta que a flexibilidade pode ser uma força libertadora tanto quanto uma tendência repressiva, se as regras forem sempre ditadas pelos poderes instituídos;

5º) A crescente “*convergência de tecnologias específicas para um sistema altamente integrado*”, no qual as trajetórias tecnológicas antigas ficam literalmente impossíveis de serem distinguidas: microeletrônica, telecomunicações, optoeletrônica e computadores estão, agora, integrados nos sistemas de informação.

O paradigma da tecnologia de informação não evolui para um fechamento, como um sistema, mas para a sua abertura como uma rede de acessos múltiplos. Abrangência, complexidade e disposição em forma de rede são os seus principais atributos.

---

*conhecimento de notícias e entretenimento.*” Para o conceito de informação, preferiu a definição operacional de Porat (1977:2): “*Informação são dados que foram organizados e comunicados*”. Referências apud Castells: Bell, Daniel. The coming of post-industrial society: a venture in social forecasting. Nova Iorque: Basic Books, 2ª edição. Porat, Marc. The information economy: definition and measurement. Washington, DC: US Department of Commerce, Office of Telecommunication, publicação 77-12(1).

A partir dos anos 70, do século XX, até a esta data, a informática integra vários modos de comunicação, em rede interativa, o hipertexto e a metalinguagem, reunindo no mesmo sistema as modalidades escrita, oral e audiovisual da comunicação humana. Castells (2002: 432) assim sintetiza:

“A potencial integração de texto, imagem e sons no mesmo sistema, interagindo a partir de múltiplos pontos, num tempo escolhido (real ou passado) numa rede global, em condições de acesso livre e a preço módico, muda de forma fundamental o carácter da comunicação. E esta molda, decididamente, a cultura (...)”

As tecnologias do século XX e XXI propiciam vivermos temporalidades distintas em espaços superpostos: o real, o virtual e a ficção<sup>6</sup>. A noção de tempo e espaço ganha novas percepções e dimensões e se vê acrescida do ciberespaço e da realidade virtual, na perspectiva de uma comunicação humana imediata, eficiente e ubíqua<sup>7</sup>.

O extremo dinamismo iniciado na modernidade resulta dessa separação entre espaço e tempo e dá condição ao processo que Giddens designa de “*desencaixe*”, ou seja, o “*deslocamento*’ das relações sociais de contextos locais de interação e sua reestruturação através de extensões indefinidas de tempo-espaço.”(Giddens, 1991:25)

A vida social moderna ganha a capacidade de conectar o local e o global de formas impensáveis em sociedades mais tradicionais, afetando a vida de milhões de pessoas ao mesmo tempo. A vida social pós-moderna ganha a capacidade de viver e conviver no espaço de fluxos de Castells, ou no ciberespaço de Lèvy (Lévy, 1999: 92), “*o espaço de comunicação aberto pela interconexão mundial dos computadores e das memórias dos computadores.*”

---

<sup>6</sup> A manifestação da realidade é percebida pelo ser humano pelos cinco sentidos e o real torna-se compreensível pela racionalidade: o que podemos ver, ouvir e sentir, através dos sentidos é real.

O imaginário é aquilo que criamos a partir de uma representação da realidade, consciente ou não, é uma representação simbólica trabalhada pela construção de símbolos e a atribuição de significados, a partir da realidade, mas por serem produto da imaginação, os símbolos construídos pelo imaginário não exigem comprovação, comparação ou verificação com o real.

A ficção é uma criação do imaginário a partir de uma simulação do real, podendo ser compreendida como uma interpretação fantasiosa do real. No caso da ficção científica, há inspiração nas descobertas científicas e tecnológicas reais, mas permitem a liberdade de projetar o curso do futuro e o potencial de impacto, sem o comprometimento com a realidade, sendo passível de conservar apenas tênues conexões com os fatos científicos reais e com tecnologias projetadas, mas não viabilizadas.

<sup>7</sup> Tofler (s/d 15ª ed:111/120), na obra “A terceira onda”, faz uma abordagem em que analisa a sincronização do comportamento humano aos ritmos da máquina, gerando a obsessão pelo tempo e pelo relógio, durante o industrialismo, e um tempo linear, diferentemente da concepção circular de tempo, repetitivo, característico das sociedades pré-industriais. O espaço, na Primeira Onda, gerou uma civilização restrita espacialmente. Na Segunda Onda, o espaço correspondia à imagem do tempo, pois quanto mais limites e horários, no tempo, mais e mais fronteiras apareciam para estabelecer limites de espaço. Castells trata do mesmo tema, mas parte dos conceitos de Inni de espaço dos fluxos e espaço dos fixos.

Com a Internet, dão-se as relações entre o concreto e o virtual, o real e o atual, as noções de espaço e tempo adquirem outra dimensão, as relações sociais formatam-se de uma outra maneira, em que o modelo da cibercultura traz consigo as cibercomunidades, constituídas por elementos diversos e interligados por *links*, constituindo o “hipertexto mundial interativo”. Novas relações se constituem a partir da troca, simbólica ou não, e da velocidade imediata do conhecimento, real ou imaginário, em que trafegam inúmeras informações e possibilidades de troca, sejam elas associadas ao poder, à economia, à cultura, ao social, à ciência ou ao conhecimento, e a outras tantas possibilidades.

### **2.1. Comunicação Científica e a Publicação Eletrônica**

A produção de revistas científicas cresceu significativamente durante todo o século XIX, em função do aumento do número de pesquisadores e de pesquisa. Durante o século XX, o crescimento permaneceu intenso, incrementado pelo fato das revistas científicas serem também publicadas por editores comerciais, pelo Estado e por universidades.

Com a expansão da quantidade de periódicos e do aumento da demanda em virtude do maior número de pesquisadores e de pesquisas, os editores de revistas científicas passaram a controlar o mercado editor, e as comunidades científicas passaram a ter dificuldade em adquirir as publicações e manter os acervos das bibliotecas atualizados, ainda que fossem os próprios cientistas e pesquisadores que fornecessem o material a ser editado.

A tecnologia digital veio para criar uma nova ordem na edição e publicação da comunicação científica: o surgimento das publicações científicas em meio eletrônico e a aproximação e interação da comunidade científica, pela web, em novas escalas de tempo e espaço.

Entende-se por publicação eletrônica

“qualquer tecnologia de distribuição de informação em uma forma que possa ser acessada e visualizada pelo computador e que utilize recursos digitais para adquirir, armazenar e transmitir informação de um computador para outros.” (Stanek, 1995, apud Sabbatini, 1999:1)

A própria definição de publicação eletrônica já prevê a comunicação, o compartilhamento ampliado, porque o computador e a rede internética promovem e facilitam a ampliação da comunicação e da interação, não só da informação propriamente editada, mas da comunidade que a acessa, a usa e a transforma em novo conhecimento.

O formato das revistas científicas permaneceu inalterado durante três séculos, mas, o avanço da tecnologia provocou alterações, a partir da década de 60, com as microformas

em substituição da cópia em papel, e, a seguir, dos anos 70 em diante, com os avanços da editoração eletrônica, que permitiram melhorar a editoração das revistas, até alcançar o suporte automatizado para todas as etapas envolvidas na publicação. De início, digitalizaram-se as revistas impressas em papel, passando a tê-las em suporte de papel e em formato digital (*online*). A tendência, no momento, é o surgimento de revistas eletrônicas editadas em meio eletrônico, formatadas com os recursos que a tecnologia digital oferece, como *hiperlinks*, recursos de hipermídia, mecanismos de buscas, etc.

A convergência da microcomputação e da rede de computadores resultou numa revolução nas formas e métodos de como a informação é gerada, armazenada, processada e transmitida, no meio acadêmico, reproduzindo as profundas alterações que aconteceram na sociedade como um todo, conforme está contextualizado neste artigo, ao falarmos da interação entre sociedade, cultura, informação/conhecimento/comunicação e tecnologia.

A publicação eletrônica da comunicação científica apresenta modelos de complementação em relação às outras formas de comunicação científica e é um fenômeno inexorável e irreversível.

## **2.2. A Interoperabilidade e a Preservação Digital**

O impacto da tecnologia digital na comunicação tecno-científica sublinhou as questões relacionadas com o gerenciamento, a perenidade dos registros do conhecimento científico e sua acessibilidade. As estratégias e a metodologia para a preservação digital envolvem procedimentos de manutenção e recuperação de dados, no caso de perdas acidentais, e a garantia da mídia e da estabilidade do conteúdo, da acessibilidade através do tempo. As mídias são suportes transitórios e exigem a constante transferência para novas mídias, mais atualizadas tecnologicamente, assim como o *software* e *hardware* que se tornam obsoletos em curto espaço de tempo, e, ainda que versões sucessivas de programas possam ser compatíveis, os fabricantes de *software* normalmente não garantem a compatibilidade por longo período, assim como o *software* proprietário, que é protegido, não disponibiliza o código fonte, o que inviabiliza a conversão de dados.

Preservação digital ou arquivamento digital é o

“planejamento, alocação de recursos e aplicação de métodos de preservação e tecnologias necessárias para que a informação digital de valor contínuo permaneça acessível e utilizável por longo prazo, considerando-se neste caso longo prazo, o tempo suficiente para preocupar-se com os impactos de mudanças tecnológicas. A preservação digital aplica-se tanto a documentos “nato-digitais” quanto a documentos convertidos do



formato convencional para o formato digital” (Hedestrom, 1997/1998, apud Thomaz e Soares, 2004:2)

Os requisitos para a preservação digital envolvem a fixação dos limites do objeto a ser preservado, preservação da presença física, do conteúdo, da apresentação, da funcionalidade, da autenticidade, da possibilidade de localizar e rastrear o objeto digital ao longo do tempo, preservação da proveniência e do contexto. (Bullock, 2001, baseado nos conceitos do modelo OAIS, apud Thomaz e Soares, 2004: 5-6)

Boeres e Arellano (2005: 3) afirmam, citando Lazorchak (2004), que as primeiras considerações sobre a preservação digital partiram da Arquivologia e que foi a comunidade arquivística a primeira a apoiar o desenvolvimento dos repositórios digitais, fundamentada no princípio arquivístico da “custódia responsável”, no princípio de respeito a ordem original (que possibilita a preservação das relações da estrutura orgânico-funcional e garante a organicidade e o contexto de produção) e no ciclo vital dos documentos.

Além destas considerações, outras foram adicionadas como a necessidade do uso de padrões internacionais de arquivamento e metadados para proporcionar uma gestão mais efetiva de estratégias de preservação de documentos eletrônicos. Tem sido uma discussão aberta, desde 1990, com a participação de diversas comunidades e com a forte presença de duas importantes instituições de abrangência internacional e experiência no desenvolvimento de padrões: A *International Organization for Standardization (ISO)* e o *Consultive Commitee for Space Date Systems (CCSDS)*.

Os arquivos e instituições de preservação de documentos participaram de oficinas internacionais de trabalho, na Europa (Inglaterra e França) e nos EUA, e o ponto de partida deste processo foi o “*desenvolvimento de um modelo de referência para estabelecer conceitos e termos comuns, fornecer um esquema para esclarecer as entidades significativas e relacionamentos entre entidades em um ambiente de arquivos*” e estabelecer padrões de apoio, divulgados pela publicação do *Reference Model for an Open Archival Information System (OAIS) – Red Book*, em maio de 1999, e segunda versão do *Red Book*, em julho de 2001, e do *Blue Book*, em janeiro de 2002, e a aprovação da ISO 14.721 *Space data an information transfer systems — Open Archival Information System (OAIS)*<sup>8</sup> *Reference model*, em fevereiro de 2003. (Thomaz e Soares (2004:13)

---

<sup>8</sup> “An Open Archival Information System (or OAIS) is an archive, consisting of an organization of people and systems, that has accepted the responsibility to preserve information and make it available for a Designated Community”.

Na transliteração, “*archival*” é traduzido como adjetivo relativo a arquivo ou a Arquivística e o OAIS, Modelo de Referência de Sistema de Informação de Arquivamento Aberto, que estabeleceu o quadro conceitual para definir e implementar repositório digital, desenvolvido pelo *Consultative Committee da Space Data*.

Segundo Bullock (1999, apud Thomaz e Soares, 2004: 7 e 12), estes manuais têm em comum as seguintes recomendações e objetivos: 1º) reconhecimento da responsabilidade inicial do produtor na preservação de seus documentos; 2º) ampliar a consciência e a compreensão dos conceitos relevantes para a preservação de objetos digitais, especialmente entre instituições não arquivísticas; 3º) definir terminologias e conceitos para descrever e comparar modelos de dados e arquiteturas de arquivos; 4º) identificar as responsabilidades da instituição arquivística; 5º) adotar diretrizes adequadas para a seleção dos objetos digitais a serem preservados e ampliar o consenso sobre os elementos e os processos relacionados à preservação e acesso à informação digital; 4º) proteger os itens arquivados de alteração intencional e não intencional; 5º) fornecer descrição do contexto incluindo histórico de criação, transferência e uso, e registros de auditoria, criando um esquema para orientar a identificação e o desenvolvimento de padrões; 6º) descrever de forma completa os objetos digitais.

### **2.2.1. OPEN ARCHIVES: um modelo para interoperabilidade das publicações digitais e OPEN ACCESS: um modelo para a interoperabilidade humana**

A OAI – *Open Archives Initiative*<sup>9</sup> é uma organização, não governamental, cujo objetivo é desenvolver e promover soluções de interoperabilidade, de tal forma, que facilite o acesso

---

Tradução livre, feita pela autora do artigo: "Um *Sistema de Informação de Arquivamento Aberto* é um arquivo que consiste em uma organização de pessoas e sistemas, que aceitaram a responsabilidade de preservar a informação e torná-la disponível para uma comunidade designada." Mais uma vez nos deparamos com a polissemia da palavra “arquivo”.

“*Open Archival Information System (OAIS) Reference Model: Developed by the Consultative Committee on Space Data, a conceptual framework and reference tool for defining a digital repository. It provides a model of the environment, functions, and data types for implementing a digital repository.*”

Tradução livre, feita pela autora deste artigo: “**Modelo de Referência para Sistema de Informação de Arquivamento Aberto**: desenvolvido pelo *Consultative Committee do Space Data*, é um modelo conceitual e um instrumento de referência para definir um repositório digital. Ele provê um modelo do ambiente, das funções e tipos de dados para implementar um repositório digital.”

Disponível em: [en.wikipedia.org/wiki/Open\\_Archival\\_Information\\_System](http://en.wikipedia.org/wiki/Open_Archival_Information_System) - Acessado em 09/04/08.

<sup>9</sup> “Iniciativa surgida em 1999 con el objetivo de desarrollar y promover estándares de interoperabilidad para la difusión y recuperación de publicaciones científicas en Internet. Esta iniciativa surge en el seno de las comunidades de *e-prints*, como alternativa a la distribución de información científica en manos de los tradicionales editores comerciales. La interoperabilidad se favorece con el establecimiento del protocolo OAI-PMH, que facilita la comunicación de metadatos codificados en el formato Dublín Core, entre servidores proveedores y clientes recolectores.”

eficiente aos conteúdos em formato digital e na *web*, através de uma maneira convencional (protocolo), e o recolhimento, de forma automatizada, dos metadados, mas não necessariamente os seus conteúdos.

*Archive* é uma palavra em inglês que, em Arquivística, na transliteração para o português, comporta algumas dificuldades na tradução do significado, apesar da aparente semelhança dos vocábulos em língua inglesa e portuguesa. Em Arquivística, “*archive*” refere-se a “arquivo histórico”, dentro da concepção custodial e patrimonialista que divide a vida e a guarda dos documentos, de acordo com a teoria das três idades – (arquivos correntes, intermediários e históricos), ou seja, “repositório” de documentos “históricos”, fonte primária para o pesquisador. Para a guarda dos documentos administrativos na fase ativa e semi-ativa, o termo designativo é “*records*”.

*Open Archives*, no entanto, que literalmente transliterado seria “arquivos abertos”, tem como sentido denotativo e conotativo a idéia de “arquivo” como sinônimo de conjunto de dados processados informaticamente, pois recordamos que arquivo é um termo polissêmico em nosso idioma e em terminologia Arquivística, como ficou sublinhado no início deste artigo. *Open* (aberto) significa abertura do ponto de vista da arquitetura do sistema, ou seja, de como as máquinas se comunicam. Terminologicamente, aberto não quer dizer uso gratuito e indiscriminado, mas sim que permite chegar ao conteúdo e constatar as condições para acessá-lo, seja através de pagamento, ou exclusivo de uma comunidade, ou que esteja em sigilo temporário, aguardando patente ou publicação prioritária e depois de um prazo, liberada.

Já o movimento “*Open Access*”, Acesso Livre ou Acesso Aberto - que na sigla O.A. pode provocar equívocos em razão das mesmas letras nas abreviaturas (“O.A. - *Open Archives*”) - surge em decorrência da Convenção de Santa Fé e da preocupação com o compartilhamento do conhecimento, da concepção de que a inclusão científica leva a uma maior inclusão social e maior acesso à informação. Kuramoto (Acessado em 1º/07/07), citando Stevan Harnard, afirma que 15% da produção científica mundial se encontra em repositórios de acesso livre, e os estudos mostram que os trabalhos depositados em repositórios de acesso livre têm um incremento cerca de 300%, em média, no fator de impacto, em relação aos repositórios de acesso restritivo. No mesmo *blog*, Kuramoto referindo-se à sua palestra na 59ª Reunião Anual da SBPC, no encontro aberto “*Publicar ou Perecer: Acesso Livre é sobreviver!*” concluiu dizendo que “*o acesso livre é*

*dependente, não apenas da interoperabilidade tecnológica, mas muito da interoperabilidade humana”.*

O Acesso Livre, ou Acesso Aberto, e os repositórios digitais, que o adotam, viabilizam a concretização de uma antiga aspiração, o livre fluxo de informação e sua distribuição para um público ampliado. O ideal de compartilhamento do conhecimento produzido pela humanidade, por redes de saber, não é contemporâneo, pois pode ser detectado na elaboração da *Encyclopédie*, de Diderot e D’Alembert, de Paul Otlet e seu grupo na Bélgica, criando a “Documentação” e propondo o *Mundaneun* e o *RBU Repertoire Bibliographique Universel*, da Memex imaginada por Vanevar Bush, até chegar ao *Open Archives* e *Open Access* e os repositórios digitais, no século XX e XXI. As TIC – Tecnologias da Informação e Comunicação, finalmente, propiciam um estágio muito avançado de viabilização dessa trajetória iniciada há tantos séculos.

Entre as muitas definições de “acesso livre”, há a de Suber (2003, apud Baptista et al, 2007:5), referindo-se “à *acessibilidade ampla e irrestrita a conteúdos disponíveis em formato digital, no sentido em que remove barreiras de preço e de permissão, tornando a literatura disponível com o mínimo de restrições de uso*”.

A adesão ao movimento do “Acesso Livre” é resultado de uma conjuntura de fatores: 1) a conscientização e reação de pesquisadores ao modelo de negócios de editora comerciais de revistas científicas, que recebem os artigos, às vezes cobrando valores significativos para publicar trabalhos de pesquisa e condicionando a publicação a rigoroso crivo de avaliação, uma vez que há o interesse dos pesquisadores, porque assim têm maiores oportunidades de visibilidade e citação dos seus trabalhos; 2) altos custos para a aquisição das assinaturas de publicações ou para o acesso eletrônico a elas, para integrar os acervos das instituições científicas; 3) crescente conscientização do aumento de impacto provocado pela disponibilização de documentos científicos livres de barreiras ao acesso, disseminando pesquisas e aumentando o índice de impacto de citação dos autores e dando visibilidade e prestígio às instituições às quais pertençam; 4) pesquisas mantidas e financiadas, por órgãos de fomento ou órgãos públicos, devem ser acessíveis sem restrição. Um outro aspecto evocado tem a ver com a preservação da informação digital de forma a garantir a memória institucional e o acesso, ainda que faltem, nos repositórios digitais em funcionamento, maiores e melhores definições de uma política e metodologia de preservação.

O modelo “Arquivo Aberto” (*Open Archives*) foi concebido a partir das experiências de Ginsparg, em 1991, no Laboratório Nacional de Los Álamos, nos EUA, que desenvolveu e implantou um repositório digital (arXiv – <[http://arxiv.org/.](http://arxiv.org/)>), na área de ciência da computação, de física e matemática.

Mueller (2006:32) assim fala sobre repositório digital:

“O objetivo é disponibilizar textos apresentados em eventos e outros canais paralelos aos periódicos tradicionais e também artigos submetidos, mas ainda não publicados em periódicos tradicionais. Os próprios autores depositam seus textos e os responsáveis pelos repositórios têm a tarefa de evitar que material irrelevante seja depositado.”

Kuramoto (2006:94) afirma que repositórios foram criados com o objetivo de contornar as dificuldades encontradas no sistema de comunicação científica (custos, demora na avaliação e publicação, etc). E se refere a eles desta forma:

“Assim, os repositórios digitais, também denominados *e-prints*, surgiram como alternativas ao tradicional sistema de comunicação científica. No entanto, esses repositórios não eram dotados de um conselho editorial que promovesse a avaliação prévia dos trabalhos pelos seus pares (*peer review*). Em seu lugar, o pacote de *software* que administra esse repositório era dotado de um moderador, o qual não tinha o papel de *peer review* como existente no sistema tradicional da comunicação científica, mas apenas o de ser um filtro, de forma a garantir que o *paper* depositado pertencia ao escopo daquele repositório.”

Kuramoto informa, ainda, que esse sistema de filtragem permite que os leitores publiquem seus comentários a respeito dos trabalhos depositados e o autor pode depositar uma nova versão, incorporando as melhorias decorrentes dos comentários, numa modalidade nova de uma antiga tradição: os colégios invisíveis, agora permitindo a integração de membros na escala mundial, pelo acesso através da *Net*, e constituindo os colégios virtuais.

O sucesso dessa experiência propiciou que outras áreas científicas criassem repositórios. Em 1999, Paul Ginsparg, Rick Luce e Herbert Van de Sompel convocaram uma reunião de responsáveis por repositórios de *e-prints* acadêmicos, na Convenção de Santa Fé, da qual se originou a *Open Archives Initiative (OAI)*. Esta iniciativa, segundo Björk (2005, apud Mueller, 2006: 32), contempla quatro canais mais importantes para o acesso aberto: 1º) periódicos científicos eletrônicos com avaliação prévia pelos pares; 2º) servidores de *e-prints* para áreas específicas (repositórios temáticos); 3º) repositórios institucionais de universidades específicas e 4º) auto-arquivamento em páginas pessoais dos autores.

A principal linha da OAI foi a definição de aspectos técnicos e de suporte organizacional para uma estrutura de publicação aberta para edições comerciais e de acesso livre, em

repositórios que tivessem entre si alto nível funcional de interoperabilidade. Um arquivo de *e-prints* (denominação que é sinônimo de repositório digital para documentos digitais no formato de comunicação científica no estágio de “interação/revisão” pela comunidade científica) deve ter essencialmente um mecanismo de submissão; um sistema de armazenamento a longo prazo; uma política de gestão para a submissão e preservação de documentos e uma interface aberta que permita a terceiros coletar os metadados dos respectivos arquivos. São necessários também provedores de dados que são os gestores de arquivos *e-prints* com, no mínimo, mecanismos de submissão para o auto-arquivamento dos trabalhos e *papers*, sistema de armazenamento a longo prazo e mecanismos de exposição de metadados do arquivo para facilitar a sua coleta por terceiros, ou provedores de serviços. Além do provedor de dados, é necessário provedores de serviços que são as instituições ou serviços de terceiros que implementam os serviços com valor agregado, a partir dos dados coletados junto aos arquivos de *e-prints*, ou repositórios digitais. (Kuramoto, 2006: 94)

O modelo *Open Archives* estabeleceu um protocolo<sup>10</sup> de comunicação para a realização da coleta de metadados a partir de determinado provedor de dados, pelo provedor de serviços que deve utilizar um programa chamado *Harvester* (mecanismo de colheita) do protocolo OAI-PMH (*Open Archive Initiative Protocol for Metadata Harvesting*).

Ao mesmo tempo, os repositórios/servidores OAI são um modelo conceitual para metadados de preservação e manutenção de acesso à informação digital por longo prazo, adotados por uma organização/instituição. Esta afirmação está relacionada com a garantia de migração para softwares mais recentes, mas não é formulada em uma política de preservação da memória, como existe, por exemplo, na Arquivística, com os critérios, mesmo que polêmicos, para as avaliações e tabelas de temporalidade. Os repositórios institucionais adotam uma política de preservação, provendo o ambiente necessário para a disseminação e processos de reprodução, assim como criando o compromisso institucional com a preservação (OAIS), mas em que não há proposição de uma metodologia para alcançá-la, exceto a garantia de arquivamento e obrigatoriedade da adoção dos protocolos e abertura dos metadados dos arquivos depositado, já referido anteriormente.

Diversas iniciativas de construção de repositórios e bibliotecas digitais foram desenvolvidas, destacando-se as bibliotecas digitais de teses e dissertações em todo o

---

<sup>10</sup> Protocolo: conjunto de regras que administra a comunicação entre os diferentes componentes de um sistema computacional. (Flamino, 2006: 89).



mundo e as publicações eletrônicas periódicas disseminadas na *web*, com a utilização do pacto de *software* Sistema Eletrônico de Editoração de Revistas (Seer), originado a partir do *software Open Journal Systems*, além de diversos outros repositórios institucionais e temáticos.

### **3. Repositórios Digitais: armazenamento e compartilhamento de publicações eletrônicas**

A disseminação de repositórios digitais emerge no contexto da chamada sociedade informacional e do informacionalismo (Castells:2000), em meio ao advento de novas disciplinas e das tecno-ciências, que atuam na interdisciplinaridade, compartilham objetos de estudo e em que, freqüentemente, ocorre a apropriação mútua de conceitos e de terminologia.

A polissemia de conceitos e de terminologia, freqüente nas áreas do conhecimento surgidas a partir do século XIX, em razão da interdisciplinaridade e da transdisciplinaridade dos saberes, parece ser uma decorrência natural da apropriação mútua de conceitos e das terminologias dessas áreas do conhecimento que se localizam nos limites ou têm os limites diluídos entre outras disciplinas.<sup>11</sup>

No entanto, é necessário distinguir interdisciplinaridade de ausência de definição de campo e objeto de estudo, como também distinguir a diferença de profissionais que atuam em campos interdisciplinares, dos profissionais sem formação, ou com formação de graduação em uma disciplina que compõe uma área de natureza inter e transdisciplinar e que esteja mais preocupado com a vigilância do exercício da profissão, delimitada por territórios rígidos de atuação, e uma defesa de mercado de trabalho, equívoca, como a que acontece entre arquivistas e bibliotecários, incompatível com a necessidade de tratamento da informação por uma equipe multidisciplinar para tratar da gestão da informação no paradigma da sociedade informacional.

---

<sup>11</sup> “Quando Japiassu faz a distinção entre interdisciplinaridade e transdisciplinaridade enfatiza que a primeira caracteriza-se ‘pela intensidade das trocas entre os especialistas e pelo grau de integração real das disciplinas, no interior de um projeto específico de pesquisa’ (Japiassu, 1976), enquanto a transdisciplinaridade, em citação de Piaget, equivaleria a uma etapa posterior e superior, ‘que não se contentaria em atingir interações ou reciprocidade entre pesquisas especializadas, mas que situaria essas ligações no interior de um sistema total, sem fronteiras estabelecidas entre as disciplinas’. (Piaget apud Japiassu, 1976, p 75)”. (Pinheiro, 1999:161).

É apropriado esclarecer que apesar de a pós-modernidade abrigar paradoxos, a multi e interdisciplinaridades referidas não são estímulos à falta de rigor e de critérios científicos na investigação e na pesquisa. É necessário aprender a lidar com a interdisciplinaridade e com as novas formas de fazer ciência, que o progresso técnico-científico instalou, coerentes e compatíveis com o paradigma tecnológico da sociedade informacional.

Em Ciência da Informação e em Informática, vários termos usados para exprimir conceitos e compor a terminologia têm significados diversos em outros contextos de uso, desde a significação vernacular até a que assumem no âmbito em que são adotadas. O mesmo vocábulo com descodificação polissêmica, como acontece com a terminologia da Biblioteconomia e da Arquivística - ou Arquivologia. E, mais recentemente, a Informática recorre a metáforas para criar uma terminologia própria e “captura” termos, como “arquivo”, “biblioteca”, e “repositório” para sua terminologia/jargão, nascidos de re-significações, dificilmente descodificadas, sem distúrbios, pelos leigos e os profissionais das outras áreas. Só para ficar num exemplo: o termo “biblioteca”, aparece na expressão “Biblioteca de Conexão Dinâmica”, na extensão DLL (*Dynamic Link Library*), que significa um conjunto de funções e rotinas de programa que podem ser acessadas dinamicamente por um programa, na medida que surge a necessidade de usá-las.

Outro aspecto a considerar é o de que, ao estudarmos um fenômeno, selecionamos certos aspectos do fenômeno, e realizamos uma “abstração”, através de um sistema do pensamento para interpretar esses aspectos, resultando na criação de termos ou conceitos, para comunicar os resultados. A criação de termos ou conceitos de uma determinada disciplina é o seu sistema conceitual e o usamos para representar o fenômeno, ou aspectos do fenômeno que estamos investigando. É importante ter em mente que conceito não é o fenômeno, é uma construção lógica a partir de impressões sensoriais, percepções, ou mesmo experiências bem complexas, ou seja, não existe sem o quadro de referência estabelecido e sem um sistema teórico.

A conceitualização é essencial porque se constitui na base de toda a comunicação e pensamento humanos, mas a ciência exige, para além disso, maior precisão, por isso o cientista não pode deixar de considerar o caráter abstrato da conceitualização. *Grosso modo*, para chegar a conceitualização, partimos das definições possíveis de que algo é, e, a partir dessas “definições”, passamos para a conceitualização e, daí, ao conhecimento. À diferença das Ciências Naturais, as Ciências Sociais têm, como característica inerente às suas especificidades, uma gama de variáveis a serem consideradas e que alteram e

dificultam o estabelecimento de princípios universais de causa e efeito e desses aspectos decorrem as questões retrocitadas sobre conceitos e terminologia nas disciplinas dessa área do conhecimento.

A área das Ciências Sociais Aplicadas, em que figura a Ciência da Informação é uma área consiliente<sup>12</sup>, como está no prefácio de publicação do IBICT/CNPq (Pinheiro, 1999: 9), que reúne artigos com estudos teóricos sobre a Ciência da Informação. Discutir a cientificidade da Ciência da Informação não é o foco desta análise, mas é necessário contextualizar e registrar o caráter polêmico da área do conhecimento que estuda, ou deveria estudar, o tratamento da informação nos repositórios digitais. Pinheiro (1999: 155) desenvolve os seus argumentos no artigo “*Campo interdisciplinar da ciência da informação: fronteiras remotas e recentes*”, defendendo o seu ponto de vista, afirmando que a Ciência da Informação

“tem seu próprio estatuto científico, como ciência social que é, portanto, interdisciplinar por natureza, e apresenta interfaces com a Biblioteconomia, Ciência da Computação, Ciência Cognitiva, Sociologia da Ciência e Comunicação”, entre outras áreas, e suas raízes, em princípio, vêm da bifurcação da Documentação/Bibliografia e da Recuperação da Informação. E seu objeto de estudo, por si mesmo, na complexidade de categoria abstrata, é de difícil apreensão”.

Apesar de não nos determos na polêmica e não evocarmos a diversidade de autores e suas concepções relativas ao caráter disciplinar da Ciência da Informação, optou-se aqui pela concepção de Silva e Ribeiro (2002: 79/90) e Silva (2006), cujas referências são as adotadas para a realização destas reflexões e análise das diversas tentativas de definições que, nos parece, ainda não permitem chegar à conceitualização dos repositórios digitais. Para Silva (2006: 140/141)

“Ciência da Informação é uma ciência social que investiga os problemas, temas e casos relacionados com o fenómeno info-comunicacional perceptível e cognoscível através da confirmação ou não das propriedades inerentes à génese do fluxo, organização e comportamento informacionais (origem, colecta, organização, armazenamento, recuperação, interpretação, transmissão, transformação e utilização da informação). Ela é trans e interdisciplinar, o que significa estar dotada de um corpo teórico-metodológico próprio construído, dentro do paradigma emergente pós-custodial, informacional e científico, pelo contributo e simbiose da Arquivística, da Biblioteconomia/Documentação, dos Sistemas de Informação e Organização de Métodos.”

---

<sup>12</sup> Consiliente é “termo criado por William Whevell em 1840 para indicar um ‘salto conjunto’ do conhecimento entre e através das disciplinas, por meio da ligação de fatos e de teorias, para criar novas bases explanatórias. (Wilson, Edward O. Consiliente, the unity of knowledge. New York: A.A. Knoff, 1998, 332 p)” (Apud Pinheiro, 1999:9)

O compartilhamento do objeto de estudo, a informação, por diversas disciplinas, leva também ao partilhar de terminologia e nomenclatura, como dissemos, o que torna fundamental a elaboração de dicionários e glossários, constantemente revistos para a apreensão da dinâmica da evolução científica e disseminação do conhecimento gerado.

“Nas ciências sociais (...), além de, numa mesma comunidade de especialistas coexistirem pensadores (cientistas) adotando diferentes paradigmas, são usadas palavras e expressões tomadas da linguagem natural, de uso comum ou emprestadas de outras áreas, sendo-lhes atribuído, contudo, novo conteúdo conceitual, sem que este fato esteja claro para os ouvintes, leitores fora de seu estreito círculo e, em alguns casos, até mesmo dentro do tal círculo.” (Gomes apud Galvão, 1998: 48)

Do ponto de vista epistemológico e teórico-metodológico, repositórios digitais não têm sido objeto de estudo da Ciência da Informação, embora estejam frequentemente presentes em artigos que o tratam como uma importante mídia da comunicação científica, relacionado com a OAI. Estes artigos e trabalhos parecem ser mais dedicados à defesa e à divulgação das implementações até agora realizadas, do que preocupados em estudar os repositórios digitais, no campo da Ciência da Informação. Os artigos, as comunicações, as teses e dissertações que tratam dos repositórios digitais são, em sua maioria, escritos por bibliotecários que objetivam estudos de mediação e acesso, incluindo as reflexões em suas investigações, como modalidade de comunicação científica, no âmbito da Biblioteconomia. A locução “repositório digital”, que reúne o substantivo e adjetivo, cujos significados já exploramos na “arqueologia” feita de início, resulta na expressão designativa de um serviço/ferramenta/instituição para atender à demanda de soluções para o tratamento da informação digital e a comunicação científica, intensivamente produzida e praticada, no contexto da sociedade informacional e, portanto, em contexto sócio-cultural-tecnológico do informacionalismo e que viu surgir vários campos de conhecimento inter e transdisciplinares.

A abordagem da análise, que estamos a construir, é fundamentada no conceito de Ciência da Informação, área inter e transdisciplinar, definido por Silva e Ribeiro (2002) e Silva (2006), e a usaremos como fundamento teórico-metodológico para estudar o objeto científico “repositório digital”.

As instituições que desde a Antigüidade custodiavam a informação, os Arquivos e Bibliotecas, armazenavam a memória do conhecimento humano ou de suas atividades, e se constituíram nos repositórios de informação em suportes analógicos, de diversas naturezas de acordo com as técnicas de cada tempo, e tiveram suas funções nem sempre claramente

distinguidas e freqüentemente confundidas em ambigüidades terminológicas, como pudemos destacar, inclusive no Quadro 1, tendo como referência Battles (2003).

A polissemia das palavras “biblioteca” e “arquivo”, leva-nos, novamente, à necessidade de definirmos, a depender do contexto do que é que estamos a tratar.

Biblioteca pode designar:

“serviço criado organicamente numa determinada entidade e/ou uma instituição cultural (...) destinada a incorporar e tornar acessível informação editada e posta a circular pelo mercado editorial-livreiro, bem como publicada e distribuída por entidades com objectivos e actividades específica (Laboratórios científicos e farmacêuticos, Unidades Industriais dos mais diversos ramos, Instituições Culturais, Associações Políticas, Cívicas e Humanitárias, etc.) ou designar “sistema (...) de informação recebida por uma entidade activa criada e vocacionada para facilitar e promover o acesso e difusão desses conteúdos acumulados. Nesta segunda acepção, Biblioteca não é a instituição, nem o serviço, não é o continente, mas o conteúdo (conjunto de livros, revistas jornais, filmes, documentários, etc), consistindo, assim, num sistema de informação organizado, cuja estrutura (entidade que o cria) não produz a informação destinada a acesso e difusão, antes a recebe e colhe nas mais diversas latitudes e, em concreto, no mercado editorial e informacional, no espaço global de “conteúdos”, intelegível, numa concepção sistêmica (...)” (Silva, 2006: 138-139)

Biblioteca Digital:

“sistema de informação que mantém e proporciona acesso remoto a uma ou várias colecções de publicações digitais, ainda que tais colecções sejam o seu aspecto mais visível, a biblioteca digital é constituída por um conjunto de elementos e processos (humanos, tecnológicos, normativos, económicos e materiais) que tornam possível a sua existência.” (Lópes Yepes, 2004, apud Silva, 2006: 139)

Etimologicamente, biblioteca é uma coleção de livros. Conceitualmente, é uma reunião de coleções organizadas e guardadas num espaço físico fechado. Sua função é a coleta, organização e disseminação de informação, geral e/ou especializada, no âmbito de uma comunidade específica. Silva et al (2006a: 265) apresentam, no artigo “*Bibliotecas digitais: uma nova cultura, um novo conceito, um novo profissional*”, em coletânea editada pelo IBICT, a diversidade e a dificuldade conceitual para designar os repositórios informacionais no meio digital, sublinhando que esse conceito deve manter o sentido de “celeiro de conhecimentos”, e citam as diversas propostas de designação: biblioteca sem papel, biblioteca do futuro, biblioteca eletrônica, biblioteca virtual e biblioteca digital, após pesquisarem o conceito, em periódicos, apontam a imprecisão que acontecia em 2004 (ano em que acontece a pesquisa), destacando que a formulação do conceito estava quase sempre nas ferramentas e nos processos, sem mencionar os atores envolvidos, como os produtores de conhecimento, os gerenciadores – profissionais de informação e analistas – e os usuários. (Silva, Sá, Furtado, 2004:3 apud Silva et al, 2006a: 171).

Na síntese conceitual a que chegaram, concluíram que:

“bibliotecas digitais são repositórios de informação disponíveis nas redes referenciais (ou de texto completo, de imagens fixas ou móveis e de sons), acessíveis de forma remota, de qualquer lugar do planeta. Esta condição está estreitamente relacionada com o conceito de recuperação da informação, desenvolvimento de ferramentas para esse fim, assim como ao estabelecimento e adoção de padrões que permitam a intercambialidade.” (Silva et al, 2006a: 273.)

Passemos agora ao Arquivo, que é assim definido no Dicionário de Terminologia Arquivística:

Significado 1: “Conjunto de documentos que, independentemente da natureza ou do suporte, são reunidos por acumulação ao longo das atividades de pessoas físicas ou jurídicas, públicas ou privadas.”

Significado 2: “Entidade administrativa responsável pela custódia, pelo tratamento documental e pela utilização dos arquivos sob sua jurisdição.”

Significado 3: “Edifício em que são guardados os arquivos.”

Significado 4: “Móvel destinado à guarda de documentos”

Significado 5. “Em processamento de dados, conjunto de dados relacionados, tratados como uma totalidade.”(Camargo e Bellotto, 1996: 5)

Silva (2006), na “Terminologia Essencial”, dentro da perspectiva que propõe para a Ciência da Informação, entende arquivo em duas acepções:

“serviço criado organicamente numa determinada entidade e/ou uma instituição cultural (Arquivo de âmbito nacional, distrital ou municipal, público ou privado) destinada a incorporar e tornar acessível informação produzida/recebida por terceiros; e sistema semi-fechado de informação produzida/recebida por uma entidade activa (ou desactivada) no decurso de sua atividade em cumprimento dos seus objectivos gerais e específicos. Não se deve confundir Serviço de Arquivo (um departamento orgânico de uma entidade e/ou uma instituição mais complexa) com Arquivo=Sistema de Informação consistindo em toda a informação/documentação, nos seus diferentes tipos e registada em diversos suportes, criada, recebida e acumulada (consagra, no sistema, a dimensão memória) por uma entidade, isto é, por uma instituição ou organização (de onde emana a estrutura orgânico-funcional do sistema)”. (Silva, 2006: 137/138)

Chamamos a atenção para a polissemia do termo arquivo, na concepção da terminologia arquivística, a ponto de criar redundâncias na própria definição de “arquivos<sup>\*</sup> que guardam arquivos<sup>\*\*</sup>”: arquivos<sup>\*</sup> - designando documentos - guardados em arquivos<sup>\*\*</sup> - espaço físico onde se guardam documentos acumulados por uma pessoa física ou jurídica ao longo de suas atividades.

Na terminologia proposta para a Arquivística, como uma transdisciplina no campo da Ciência da Informação, o conceito de Arquivo é perfeitamente ajustável ao que se conceitua como repositório digital institucional - informação produzida como atividade-fim e atividade-meio de uma instituição, inclusive a acadêmica e universitária-, e



repositório digital temático - informação produzida por uma entidade criada para colecionar informação com o objetivo de atender a um tema, ou qualquer outro critério de seleção. O repositório digital temático poderia ser um sub-sistema, um serviço de arquivo, dentro de uma instituição e do seu correspondente sistema de informação que pode abrigar biblioteca, repositório digital e outros sub-sistemas que prestam serviço de informação a essa instituição/organização, ou uma instituição criada para o fim de guardar informação digital sobre temas determinados, como fazem as bibliotecas especializadas.

Se adotarmos a definição de Ciência da Informação, proposta por Silva (2006: 140), o objeto informação social, sem outras adjetivações se é científica, ou técnica, etc., é o objeto de estudo da CI, incluindo o estudo das propriedades da informação, o fluxo e o comportamento informacionais, o que implica também em estudar a origem, a coleta, a organização, o armazenamento, a recuperação, a interpretação, a transmissão, a transformação e a utilização da informação, portanto a informação nos repositórios digitais é objeto de estudo da Ciência da Informação, inter e transdisciplinarmente.

Bibliotecas e arquivos, nos conceitos retrocitados, em Ciência da Informação, proposto por Silva (2006), cuja significação esteja relacionada a sistemas de informação, cumprem função semelhante àquela que atualmente é dada aos repositórios digitais, no aspecto de armazenamento. Portanto, arquivos e bibliotecas, como sistemas de informação (incluindo a informação em suporte digital) lidam com o mesmo objeto – informação social - e têm como objetivo dar tratamento a essa informação, de acordo com as funções das quais decorreu, em conformidade com a estrutura orgânico-funcional da organização que a produziu e do uso que se fará dela. Se o suporte é o digital, deverão adaptar os procedimentos compatíveis com os recursos que a automação permite e/ou exige, seja na produção, nos fluxos, no armazenamento e na recuperação, mas isso não altera a fundamentação teórico-metodológica em que se fundam as unidades de informação e os sistemas de informação que lhes correspondam.

Se a informação é produto decorrente das atividades de uma organização (pessoa física ou jurídica), destinar-se-á, ao Arquivo, porque corresponde ao produto (informação) das atividades executadas por uma estrutura orgânico- funcional que visa a um objetivo de uma organização, estabelecido por contrato social, lei/decreto, ou decorre das atividades de uma pessoa física. Se, por outro lado, a informação digital foi colecionada em razão de um tema escolhido, de uma área do conhecimento, de uma atividade humana, o sistema de informação estará no âmbito de competência das bibliotecas digitais, ou repositórios

digitais, porque, em ambos os casos, a aquisição do acervo se processa de forma similar ao das bibliotecas, ao adquirirem e atualizarem os seus acervos e coleções. O destino do objeto informação para as unidades de informação não está no suporte, mas na forma e finalidade em que e para quê a informação é produzida. A atenção ao suporte acontece, desde sempre, com a preocupação da conservação e preservação da informação nele inscrita, de acordo com as suas características físico-químicas, e foi assim no armazenamento e ordenação das tabuinhas de argila, dos rolos de papiro, dos manuscritos, dos livros, plantas, vídeos, gravações, etc. Por que é então que surgem os repositórios digitais, nas modalidades institucional e temático, em vez da informação ser destinada a arquivos ou bibliotecas, mesmo que na modalidade bibliotecas digitais? Talvez porque os repositórios digitais, diante do paradigma tecnológico e do estágio técnico que alcançaram as TIC, trabalham com a informação em um suporte que permite a agregação de um diferencial importante, advindo dos recursos que a web e o paradigma tecnológico viabilizam.

No Quadro 2, estão apresentados conceitos e definições de repositórios digitais, que possibilitam constatar a pluralidade de percepções, sendo possível verificar que os conceitos vão desde tentativas de dizer o que é ou o que faz, variando entre ser uma ferramenta, um serviço ou uma instituição, cujo objetivo principal é a preservação e disseminação da informação, em geral, restrita à informação científica e o seu uso na comunidade científica e nas universidades, numa modalidade alternativa e aperfeiçoada de comunicação científica, assim como são os periódicos e publicações eletrônicas. No entanto, é um recurso, seja como um serviço, ou ferramenta, passível de ser utilizado em organizações privadas ou públicas, nas modalidades de repositórios institucionais e temáticos.

Do que consta no Quadro 2, é possível inferir que os repositórios digitais são caracterizados e definidos como armazém de objetos digitais ou informação digitalizada, isto é, restritivamente guardam informação (destacadamente a informação científica), codificada em código alfa-numérico e em suporte digital.

O repositório institucional tem a conceitualização fundamentada no fato de reunir a informação produzida e acumulada, por uma entidade/instituição, ou seja, uma função característica da definição de arquivo no conceito de “conjunto de informação acumulada ao longo das atividades de uma instituição”, ou arquivo como “entidade responsável pela custódia, tratamento e utilização da informação”. Seguindo, ainda, os demais sentidos de

“arquivo”, resta a possibilidade de repositório digital ser um “espaço” (virtual), ou “lugar” cibernético, em que a informação é depositada. Já o repositório digital temático, tem a sua designação decorrente de produzir e acumular informação temática, similar à biblioteca especializada. Parece-nos, então, que deveriam ser apenas repositórios digitais temáticos, uma vez que, em geral, há uma predominância de serem definidos como “coleção” de informação digital. O repositório institucional, em verdade, não deveria ser assim designado, pois ou é um serviço de informação que compõe o sistema de informação de uma instituição/organização, ou então deveria armazenar e tratar toda a informação produzida, nas atividades-fim e atividades-meio de uma instituição e, neste caso, será o sistema de informação dessa instituição/organização.

A obra “Comunicação e Produção Científica: contexto, indicadores e avaliação”<sup>13</sup>, que reúne textos produzidos por pesquisadores do Núcleo de Produção Científica (NPC) da ECA/USP, na área de pesquisa da Comunicação Científica, traz um Glossário referente a Termos e Conceitos da Área de Comunicação e Produção Científica, em que Lara (2006:410) define “Repositório”, sem a adjetivação de digital, como um servidor.

Servidor é um sistema de computação que fornece serviços a uma rede de computadores; esses serviços podem ser de diversa natureza, como arquivos e correio eletrônico. O termo servidor é usualmente aplicado a computadores completos, embora um servidor possa equivaler a um *software* ou a partes de um sistema computacional. A definição de Lara, para repositório, dá a um termo vernacular, com sentido amplo, um conteúdo semântico específico e faz uma silepse da locução, porque não agrega o adjetivo digital, e o define como “servidor, acessível em rede”. Assim a palavra “repositório” torna-se terminologia, ou seja, dentro de um sistema conceitual. No glossário, significa servidor, acessível em rede e vinculado a provedor de dados. Acontece, no entanto, que ao remeter para “arquivos abertos” e “arquivo de *ePrint*”, nos leva a concluir que todos são repositórios digitais, no

---

<sup>13</sup> Ainda em relação a esta obra, a orelha da capa da publicação apresenta texto, assinado por Maria Immacolata Vassallo de Lopes, que destaca a compreensão e o uso do conhecimento advindo da Comunicação Científica, como uma necessidade contemporânea, por ser “*essa área emergente dentro das ciências da comunicação e da informação*”. A obra foi concebida e viabilizada através de canais de comunicação que envolveram, além dos meios tradicionais, os meios baseados nas tecnologias de comunicação e informação das redes eletrônicas, o que operacionalizou a constituição de um “colégio invisível eletrônico”, experiência destacada, pelos organizadores, na apresentação da obra, como positiva e estimulante. A experiência contou, ainda, com a leitura e acompanhamento crítico dos Profs. Drs. Antônio Agenor Briquet de Lemos e Isaac Epstein. No Prefácio, Epstein esclarece: “*Esta publicação, proposta pelo Núcleo de Produção Científica (NPC) sediado na ECA/USP, foi elaborada por um grupo composto por docentes e doutores de vários programas de Pós-Graduação do país. Os textos aqui expostos, pela variedade e riqueza dos temas, competências de seus autores e riqueza da bibliografia referida é mais que oportuna como instrumento de pesquisa dos investigadores da área*”. (Población (Org), 2006)

sentido de que armazenam informação digital, e que somente o “arquivo de *ePrint*” apresenta a diferença de ter a informação passível de ser alterada, após a publicação em meio digital, por se constituírem em “literatura cinzenta”. Todos os conceitos mantêm similaridade ou aproximação com os conceitos de repositório digital, quando se vinculam à interoperabilidade, à Iniciativa de Arquivos Abertos e ao auto-arquivamento.

No Quadro 2, repositório digital aparece definido ora como coleção, ora como arquivo. Em qual dos sentidos de arquivo? É no sentido percebido em Informática? Se fosse assim porque a designação original foram “*open archives*” e não “*open files*”? Tanto coleção como arquivo são vocábulos que aparecem nas definições reproduzidas no Quadro 2, seja na afirmação de que repositórios institucionais são constituídos por “coleção de arquivos digitais”, ou “arquivos digitais de coleções”, o que evidentemente não é a mesma coisa, pois a ordem, neste caso, altera o significado. Ou, ainda, como “*coleções digitais que armazenam [...] a produção intelectual das comunidades universitárias*”, o que resulta em circularidade e redundância.

**Quadro 2: Conceitos de repositório digital, repositório institucional e repositório temático**

<b>Autor</b>	<b>Obra</b>	<b>Definição/Conceito</b>
Rodrigues et alii	RepositoriUM (Acesso em: 2007)	<i>Repositórios institucionais são “coleções digitais que armazenam, preservam, divulgam e dão acesso à produção intelectual de comunidades universitárias.”</i>
Linch (2003 apud Carvalho et al)	[2006:3] <i>O repositório institucional como recurso de apoio à gestão do conhecimento da Universidade Católica de Brasília</i>	<i>Repositório institucional é “um conjunto de serviços que a universidade oferece aos membros de sua comunidade para a gestão e disseminação de materiais digitais criados pela universidade e membros de sua comunidade.”</i>
Crow (2002:16 apud Carvalho et al)	[2006:3] <i>O repositório institucional como recurso de apoio à gestão do conhecimento da Universidade Católica de Brasília</i>	<i>“[...] repositório institucional é um arquivo digital de produtos intelectuais criados por professores de uma instituição, por uma comunidade de pesquisadores e estudantes.”</i>
Para Crow, Linch, Rodrigues (2007:1), segundo Carvalho et al (2006:3)	[2006:3] <i>O repositório institucional como recurso de apoio à gestão do conhecimento da Universidade Católica de Brasília</i>	<i>“os repositórios institucionais possibilitam reunir, preservar, dar acesso e disseminar boa parte do conhecimento da instituição, contribuindo para aumentar a visibilidade de sua produção científica.”</i>
Universidade do Minho	Serviço de Documentação da Universidade do Minho <sup>14</sup>	Repositórios institucionais são “Sistemas de informação que servem para armazenar, preservar e difundir a produção intelectual de uma dada instituição, normalmente uma comunidade universitária. Podem ser criados e mantidos de forma individualizada, ou por grupos de instituições que trabalhem numa base cooperativa.”
Latin American Knowledge Harvester	Disponível em: <a href="http://laih.unm.edu/lakh/pt/glossary.html#b20">http://laih.unm.edu/lakh/pt/glossary.html#b20</a> Acesso em 10/01/08	“Repositório Institucional (RI), como um conceito, deve capturar e colocar à disposição de um usuário a maior quantidade de pesquisas de uma instituição (por exemplo, uma universidade). Na primeira instância, poderia incluir materiais tais como artigos de pesquisa e versões eletrônicas de outro tipo de documentos, tais como teses ou dissertações; mas também pode incluir muitas das fontes digitalizadas geradas no trabalho diário institucional, como são os documentos administrativos, notas de aula ou qualquer objeto de aprendizagem”. (Citado e traduzido da Wikipedia).”
Crow (2002), apud Leite (2006:88)	<i>Gestão do conhecimento científico acadêmico: proposta de um modelo conceitual. [Dissertação de Mestrado]</i>	<i>Repositórios digitais são “coleções digitais que capturam e preservam a produção intelectual da comunidade de uma única universidade ou de uma comunidade multiuniversitária”</i>

<sup>14</sup> Disponível em: [http://lusospace.sdum.uminho.pt:8080/pt/dspace\\_about.jsp](http://lusospace.sdum.uminho.pt:8080/pt/dspace_about.jsp) - Acessado em 10/01/08.

- Lara (2006:410) *Glossário na obra “Comunicação e Produção Científica: contexto, indicadores e avaliação”*
- Repositório é um “servidor acessível em rede que pode processar as solicitações exigidas pelo Protocolo The Open Archives Initiative Protocol for Metadata Harvesting – OAI-PMH. Um repositório é gerenciado pelo provedor de dados para expor os metadados aos colheidores (harvesters).” [Remete para arquivos abertos – open archives]
- Lara (2006:390/391) *Idem*
- “(…) (Open Archives) Arquivos ou repertórios de trabalhos científicos em forma digital (ePrints) disponibilizados para o acesso público via ftp ou http. Incluem versões digitais preliminares de documentos científicos como também artigos aceitos para publicação pelo processo tradicional de revisão por pares. Constituem um modelo alternativo e eqüitativo de disseminação de bases de dados de conhecimentos como também fórum privilegiado de debate e de difusão da produção científica. Os arquivos abertos nasceram da Open Archives Initiative – OAI, com a finalidade de desenvolver padrões de interoperabilidade visando a facilitar a disseminação de conteúdos. The Open Archives Initiative – OAI, possui código compartilhado para campos de metadados (exemplo: “date”, “author”, “title”, “journal” etc), de forma que os textos completos, mesmo em diferentes formatos, são interoperáveis entre si. Os metadados de cada documento podem ser coletados e todos os documentos podem ser pesquisados e recuperados juntos tal como se estivessem em uma coleção global, acessível a todos.” (Lara, 2006:) [Remete para arquivo de ePrint]
- Lara (2006: 390) *Idem*
- Arquivo de Eprint (eprint archive) é “Coleção ou repositório digital de documentos eletrônicos, muitas vezes identificados como literatura cinzenta. São arquivados pelo autor e circulam à margem das publicações tradicionais, podendo ser objeto de atualização mesmo após passarem pelo processo de revisão pelos pares. Os arquivos de ePrints compartilham os mesmos metadados, tornando seus conteúdos interoperáveis entre si. Seus metadados podem ser coletados em arquivos virtuais globais simultaneamente sem precisar pesquisar um a um. É como se fosse uma base de dados comercial, mas com acesso ao texto completo para qualquer usuário. Os arquivos de ePrint foram erroneamente denominados no Brasil, de arquivos abertos. No domínio da Comunicação Científica, os arquivos de ePrint são considerados um modelo alternativo de comunicação, onde os pares tornam-se seus próprios editores. Termos relacionados: Arquivos abertos; auto-arquivamento.”
- Weitzel (2006:5 e 6) *Reflexões sobre os repositórios institucionais. In Intercom – Sociedade Brasileira de Estudos Interdisciplinares da Comunicação, XXIX Congresso*
- Repositório digital é “um arquivo digital que reúne uma coleção de documentos digitais”, constituídos por publicações científicas eletrônicas, como revistas científicas, anais de eventos, relatórios de pesquisa etc., que podem ser acessados tal qual ocorre em uma biblioteca. A seguir, afirma que podem ser categorizados em repositórios institucionais, que



- Brasileiro de Ciências da Comunicação. UnB, 6 a 9 de setembro de 2006.* são aqueles que se “referem à organização e acesso à produção científica de uma instituição”, e os temáticos, “que se referem a uma área”.<sup>15</sup>
- Leite (2006:84) *O repositório institucional como recurso de apoio à gestão do conhecimento da Universidade Católica de Brasília* Repositórios digitais são “uma conseqüência da aplicação da Internet e de tecnologias emergentes de informação e comunicação no contexto da comunicação científica, somada ao movimento mundial de acesso livre à informação científica, fundamentado no modelo de arquivos abertos, levando ao surgimento de modelos alternativos que provocaram mudanças paradigmáticas no processo de comunicação científica.”
- Viana, C. L.M; Márdero Arellano, M.A.; Shintaku, M. (s/data) Repositórios institucionais em ciência e tecnologia: uma experiência de customização do Dspace. “Um repositório digital é uma forma de armazenamento de objetos digitais que tem a capacidade de manter e gerenciar material por longos períodos de tempo e provar o acesso apropriado. Essa estratégia foi possibilitada pela queda nos preços no armazenamento, pelo uso dos padrões como o protocolo de coleta de metadados da Iniciativa dos Arquivos Abertos (OAI-PMH), e pelos avanços no desenvolvimento dos padrões de metadados que dão suporte ao modelo de comunicação dos arquivos abertos.”
- Weitzel (2006a:126) *Os repositórios de e-prints como nova forma de organização da produção científica: o caso da área das ciências da comunicação no Brasil. (Tese Doutorado em CI) São Paulo: USP/ECA, 2006* Na pesquisa, “repositório de e-prints [é] um tipo de repositório digital, que pode ser temático ou institucional, mas que utiliza obrigatoriamente o software E-prints desenvolvido pela University of Southampton, cujos textos que reúne já foram publicados ou não, e estão sujeitos à crítica por meio de comentários de outros pesquisadores.”
- Weitzel (2006a: 119) *Idem* “Os provedores de dados são os repositórios digitais, propriamente ditos, incluindo os repositórios de e-prints. Também são as publicações on-line individuais, isto é, periódicos científicos, teses e dissertações, anais de eventos etc, contendo, preferencialmente, os textos completos dos trabalhos já publicados como os seus dados descritivos (metadados)”.
- Ginspard, Paul (apud Café e Lage, 2002) *Auto-Arquivamento: uma opção inovadora para a produção científica.* Repositório de e-prints, inicialmente concebido para a área da Física, no Laboratório Nacional de Los Álamos, Novo México, depois ampliado para as áreas de Matemática e Ciências da Computação. Destina-se à comunicação científica e foram criados como um meio de compartilhar em uma mesma comunidade os resultados de pesquisas, relatórios preliminares, trabalhos ainda não revisados (*non-peer reviewed*), etc. Com o aumento desses tipos de repositórios disponíveis na web, sentiu-se a necessidade de criar uma estrutura técnica e estabelecer padrões de tecnologias de informação e comunicação que viabilizassem a interoperabilidade entre eles.

(Fonte: Elaborado pela autora a partir dos autores citados na referência, ao final deste artigo.)

<sup>15</sup> O Reposcom (<http://reposcom.portcom.intercom.org.br>) e o E-LIS (<http://eprints.relis.org>) são exemplos de repositórios temáticos e institucional respectivamente. (Weitzel.a, 2006:139).

Linch define como um “conjunto de serviços que a universidade oferece aos membros de sua comunidade para a gestão e disseminação de materiais digitais [...]”. Neste caso, há uma restrição de repositório institucional à comunidade universitária e o identifica como prestador de serviços que têm por objetivo a gestão e disseminação de “materiais digitais”, evitando assim o uso das palavras documento ou informação.

Os conceitos, elaborados por Lara, não permitem distinguir as diferenças entre arquivo aberto, arquivos *e-Print* e repositório, neste caso, sem o adjetivo “institucional”.

**Repositório** é um “servidor acessível em rede que pode processar as solicitações exigidas pelo Protocolo [...] OAI-PMH” gerenciado pelo provedor de dados que expõe os metadados aos usuários pesquisadores. Neste caso, a definição de repositório fica dependente de um Protocolo específico: o OAI-PMH – *The Open Archives Initiative Protocol for Metadata Harvesting*, o que condiciona o termo e conceito à adoção do Protocolo OAI-PMH. Com o uso de outro *software* ou adoção de outro protocolo, já não será mais um repositório?

No caso da definição de arquivos abertos, por Lara, assemelha-se muito aos conceitos aqui coletados sobre o repositório institucional. No entanto, o uso dos termos “arquivos” e “repertórios”, colocados como sinônimos pela conjunção “ou”, que pode ser usada para equivalência ou significado adverso, confundem a descodificação do que se pretende definir. De início, poder-se-ia inferir que seriam repositórios de versões digitais preliminares, mas ao longo da explanação, a Autora inclui também artigos aceitos para publicação pelo processo tradicional de revisão por pares e, neste caso, parece servir perfeitamente para definir repositório institucional, como é feito em outras definições que integram o Quadro 2.

A definição do arquivo de *ePrint* parece ser similar ao conceito de repositório digital, conforme conceitualizaram os demais autores. O uso do vocábulo “coleção” seguido da conjunção “ou” “repositório digital de documentos eletrônicos” permite inferir que uma reunião de documentos eletrônicos não se constitui uma coleção. Os documentos eletrônicos, nesse conceito de arquivo de *e-print*, podem incluir um estágio de elaboração da informação, denominado de literatura cinzenta, ou seja passível de ser revista, comentada e alterada pelo autor. Aliás esta é uma propriedade dos textos produzidos em meio eletrônico e que altera a forma de fenômeno info-comunicacional no paradigma das novas tecnologias.

Contudo, parece não fazer muito sentido falar em literatura cinzenta, uma vez que repositórios institucionais se propõem a ser um fórum em que o autor poderia dar publicidade aos seus trabalhos, independente da submissão formal a critérios de avaliação e aprovação por seus pares, apesar de permitir a posterior alteração e evoluções da produção científica, de forma dinâmica e a mantê-la sob constante atualização e revisão.

No endereço eletrônico<sup>16</sup> do Serviço de Documentação da Universidade do Minho, repositório institucional é assim definido:

*“Sistemas de informação que servem para armazenar, preservar e difundir a produção intelectual de uma dada instituição, normalmente uma comunidade universitária. Podem ser criados e mantidos de forma individualizada, ou por grupos de instituições que trabalhem numa base cooperativa.”*

Aqui também o conceito, exceto pela última oração do período, cabe tanto nos conceitos de arquivo ou de bibliotecas, independente de ser a informação em suporte digital ou analógico. A dificuldade em definir e conceituar é, entre outras, a de optar pela síntese e deixar importantes funções e características de lado, ou tentar cercar todas as funções e características do objeto descrito, ampliando tanto que, de tão extenso e pormenorizado, o conceito ou a definição tornam-se demasiadamente restritivos e específicos.

Numa abordagem, dentro do campo inter e transdisciplinar da Ciência da Informação, as funções dos arquivos e bibliotecas, conforme foram aqui conceituados, seriam as instituições mais adequadas para abrigar os repositórios digitais e deixaria de ter sentido as modalidades institucional e temática, como retro-referimos. Integrados aos arquivos e bibliotecas, e ambos compondo o sistema de informação institucional, de acordo com a natureza, origem e uso da informação que armazena, os repositórios digitais funcionariam com as principais características que lhes são próprias: o auto-arquivamento, como uma variação dos depósitos legais de obras; a interoperabilidade, para a garantia da disseminação e acesso, através da adoção de protocolos de arquivos abertos e, idealmente de acesso aberto, e a preservação de informação produzida em meio digital, constituindo-se, ao mesmo tempo, em garantia de visibilidade institucional e da memória institucional, porém em um conceito de memória diverso daquele previsto, tradicionalmente, pelos arquivos e bibliotecas.

A criação de “repositório digital” como uma nova “instituição” específica para tratar da informação em meio digital, poderia ser criticada como sofisticamente originada da

---

<sup>16</sup> [http://lusodspace.sdum.uminho.pt:8080/pt/dspace\\_about.jsp](http://lusodspace.sdum.uminho.pt:8080/pt/dspace_about.jsp).

premissa de que a informação em meio digital, em razão do suporte, é outro objeto, por isso o repositório tem que ser qualificado com o adjetivo digital. No entanto, o objeto é o mesmo, ou seja, a informação, independente do suporte ser argila, pergaminho, papel impresso, gravação eletrônica etc., como decorre da evolutiva progressão de substituição ou agregação de tecnologias, foi sempre “armazenada” por arquivos e bibliotecas, não sendo o suporte, mas a forma como foi produzida e como será utilizada que determina o tratamento a lhe ser dado, por arquivos ou bibliotecas.

Por que, então, foram criados os repositórios digitais?

Weitzel (2006:115) destaca fases na emergência dos repositórios digitais, entendidos como decorrência da aplicação do “ *modelo OA* [Open Archives – Arquivos Abertos], *apesar das mudanças estruturais no fluxo da informação científica, não rompe paradigmas*”, porque se ajusta às novas tecnologias, no que chamamos aqui de sociedade informacional e ao informacionalismo. A primeira fase, denominada Pré-OAI, é a que se inicia com a criação do primeiro repositório, o ArXiv.org, e passou a ser o incentivo para a criação de “arquivos públicos digitais” (original denominação dos repositórios digitais), sendo, depois, num segundo momento, com a militância de Stevan Harnard a alterar a cadeia de publicação, proposta a idéia da auto-publicação e do acesso livre, tornando a comunicação, nas comunidades científicas, mais ágil, a custos menores e mais efetiva, inclusive contando com a revisão e aprovação pelos pares, pela rede da Net, em acesso e disseminação mais abrangentes e rápidos, que permitiam a alteração e revisão pelo autor de uma maneira mais flexível e atualizada constantemente. A fase seguinte, iniciada com o Congresso de Santa Fé, é a da constituição da Iniciativa dos Arquivos Abertos, em 1999, marcada pelo encontro de soluções técnicas e operacionais para serem usadas para os arquivos de *e-prints*, tanto para os já existentes quanto para os que foram sendo criados, dotando-os de provedores de dados, ou seja, as publicações *online* e os repositórios digitais, os provedores de serviços de dados que compõem a infra-estrutura capaz de assegurar a auto-sustentabilidade, interoperabilidade dos diferentes sistemas e o acesso livre para todos interessados em pesquisas e em baixar arquivos da produção científica. E a fase atual que é a da consolidação e difusão desta modalidade de gestão do conhecimento científico.

Com relação à Biblioteconomia e a Arquivística, praticadas desde a Antigüidade, evoluem e pode também serem reconhecidas fases, conforme Silva et al (1999:28) designaram: partem da fase sincrética e pré-custodial que dura até o século XVIII, passando para uma fase técnica e custodial, em que são estabelecidos os princípios técnicos e práticos para

funcionarem, e, somente a partir de 1980, entram na fase científica e pós-custodial, voltadas para a disseminação do conhecimento e para o acesso à informação.

Tanto os arquivos quanto as bibliotecas podem ser definidos como coleções<sup>17</sup> de documentos conservados para fins de utilização e, a partir do século XX, há uma crescente elaboração de métodos de gestão para dar a conhecer melhor os acervos e oferecer ao usuário uma ampla gama de serviços que lhe permitam usufruir o direito irrestrito de acesso à informação disponível.

A informação torna-se mais importante do que o suporte e desde o surgimento da Documentação, no século XIX, com Paul Otlet e La Fontaine, integram-se os serviços de documentação (proporcionando acesso ao teor dos documentos mediante índices, catálogos e outras formas de classificação) e os serviços de informação (extraído dos documentos, e de outras fontes, os itens de informação que interessam ao usuário) aperfeiçoam-se crescentemente para atender ao aumento do volume de informação, à demanda dos usuários e aos progressos da informática.

Uma questão relevante, sempre evocada ao se justificar a criação de repositórios digitais, se refere a considerá-los uma estratégia fundamental para a preservação dos documentos digitais (às vezes chamados objetos digitais), mas que se trata de informação, em suporte digital, objeto de estudo da Ciência da Informação, inter e transdisciplinar.

Reconhecida a natureza peculiar do suporte digital, que exige procedimentos para manter a acessibilidade e autenticidade da informação, através do tempo, e que inclui investimentos e procedimentos técnicos específicos, realizados por uma equipe multidisciplinar, sublinhamos que cuidar da memória institucional não é prerrogativa dos repositórios digitais. A política de conservação de informação é cada vez mais reconhecida e praticada, em arquivos e bibliotecas, ainda que seja freqüente a falta de inversão de recursos, há uma crescente evolução na conscientização e aprofundamento do conhecimento, mais recentemente acrescido pela necessidade de tratamento do suporte digital, o que envolve combinações específicas dos componentes de *hardware*, *software*, mídia e pessoal técnico.

---

<sup>17</sup> A Arquivística Tradicional não aceita este termo, pois o descodifica como sendo uma “reunião artificial de documentos que, não mantendo relação orgânica entre si, apresentam alguma característica comum.” (Bellotto e Camargo, 1996:17). O objetivo é demarcar a diferença de acervo entre arquivo e biblioteca, pelo primeiro acumular “conjunto de documentos reunidos por acumulação ao longo das atividades de pessoas físicas e jurídicas” (Idem:1996: 5), e a segunda acumular documentos em consequência de uma política de aquisição, o que caracterizaria o ato de colecionar a partir de um tema, ou assunto, e constituir as coleções que gere.

No entanto, a questão da preservação digital, adotada nos repositórios digitais, tem a ver muito mais com a interoperabilidade e abertura dos metadados, para garantia de acesso através de *hardwares* e *softwares* em diversas tecnologias, do que com o conceito de preservação e de uma política de guarda da informação. Isto fica denotável no que Thomaz (2007:88) destaca, ao afirmar que repositório digital confiável é aquele que garanta o acesso confiável, por longo prazo, à sua comunidade-alvo dos recursos digitais sob sua administração, agora e no futuro. Mas, o prazo não é determinado e nem se orienta como determiná-lo; é apenas adjetivado por um evasivo “longo”, e ainda que em alguns artigos se faça a menção à metodologia arquivística para a avaliação da informação custodiada, nada mais detalhado foi encontrado nas referências e autores, consultados para a elaboração desta pesquisa, ao contrário dos arquivos e bibliotecas que contam com critérios de avaliação, ainda que polêmicos e carregados de muita subjetividade, para o armazenamento e descarte da informação que armazenam e tratam.

Outro argumento utilizado, na defesa do papel dos repositórios digitais na preservação da informação, é a justificativa de que os repositórios digitais institucionais garantem a preservação da memória institucional. Novamente, não encontramos na literatura pesquisada, procedimentos que garantam a prática de uma política, com prazos definidos e critérios de avaliação, ou a sistematização da guarda que comprovem essa função. O conhecimento científico é uma informação com valor de uso, muitas das vezes, sujeito à obsolescência, em que os fluxos de informação, que acompanham a noção de tempo e espaço, da cibercultura e do espaço dos fluxos no informacionalismo, não é da mesma natureza da percepção feita pelos arquivos e bibliotecas que, historicamente, assumiram o papel de preservar a memória da humanidade, na fase vista atrás, designada de custodial. A natureza do uso da informação pela comunidade científica obedece a razões que são diferentes daqueles evocadas pelos arquivos e bibliotecas, o que pode levar a critérios subjetivos, submissos apenas à percepção particular do pesquisador, por isso falível, do que é uma informação útil e passível de ser preservada. Repositórios digitais respondem a uma demanda essencialmente prática, da informação que armazenam, e o armazenamento em meio digital e sua atualização envolvem custos e, embora, dentro de uma outra concepção de espaço ocupado pela armazenagem mais compacta do que a analógica, haverá sempre um limite passível de ser alcançado e necessitando de ser manejado.

Desde a Biblioteca de Alexandria, os seres humanos tentam coletar o conhecimento do mundo. A angústia da perda do conhecimento acumulado atravessa os tempos e,



principalmente, em momentos de mudança de paradigma tecnológico ou sócio-cultural, há uma acentuação dessa preocupação.

A *Encyclopédie*, concebida em consequência dos ideais da Revolução Francesa, é pensada como *O Livro*, “*um ato livresco radicalmente novo de tomada de posse do mundo, a expressão triunfal dos avanços do saber*”, uma síntese útil e manejável dos saberes existentes, uma forma de compressão do conhecimento, utilizável, prática, uma “biblioteca” ao alcance das mãos e capaz de salvar, do desastre, uma cultura e uma tecnologia. (Goulemot, 2006: 262-263).

No século XIX, Paul Otlet e Henri La Fontaine preparam o conceito de globalização, com o conceito de “mundialismo”, sonham em facilitar o acesso do maior número de pessoas à informação, graças a um complexo conjunto de bibliotecas, conectadas por canais telegráficos e telefônicos. Otlet sempre teve em mente fazer com que o conhecimento registrado fosse disponibilizado para quem o necessitasse e com isso contribuir para o enriquecimento intelectual da humanidade e, ao reconhecer que a pesquisa e a recuperação da informação dos documentos, realizada por uma equipe permanente qualificada, era fundamental. Desta forma, antecipou o que estava por vir, e que o avanço das tecnologias tornou possível, em termos de comunicação em rede, armazenamento, acesso e disseminação da informação/conhecimento.

Não é de hoje, a preocupação com o volume do conhecimento acumulado e a falta de capacidade de controle sobre ele, recrudescido a cada inovação tecnológica na história da humanidade. As tecnologias digitais revelaram, como nunca, o volume exponencial de informação/conhecimento acumulados pela humanidade, agravado pelo fato de a informação agora estar encriptada no código alfa-numérico, condicionada aos *softwares* e as evoluções tecnológicas que acontecem numa velocidade muito maior do que jamais antes havia acontecido.

Os repositórios digitais emergem das demandas da comunidade científica, surgida na área das ciências físicas e naturais, no contexto em que o cruzamento dos fatores conjunturais propiciaram essa emergência, algo ainda recente, e que passou a extrapolar para outros usos, e se constitui em desafio a ser estudado pela Ciência da Informação.

À guisa de conclusão deste artigo, porém para ter o efeito de prosseguir as reflexões aqui realizadas, é apresentado, a seguir, uma análise para a inserção dos repositórios digitais às cinco características do paradigma tecnológico, propugnado por Perez, Freeman e Dosi, numa adaptação da clássica análise das revoluções científicas de Kuhn, e referido por

Castells (2000:86-94), enumeradas anteriormente neste artigo, ao fundamentar a interação entre as TIC, sociedade e economia (capitalismo) e o conceito de sociedade informacional e o informacionalismo:

**1ª Característica:** Os repositórios digitais se caracterizam pelas tecnologias da informação agindo sobre a informação, seja na produção, coleta, disseminação e preservação da informação, seja alterando as práticas editoriais clássicas, ainda que não através de uma total ruptura, porque, como diz Weitzel (2006), as publicações formais continuam a existir e não são substituídas, apenas ganham um outro suporte para a divulgação. O movimento de acesso livre é outra importante ação das tecnologias sobre a informação e o acesso ao conhecimento. Os repositórios digitais constituem um serviço de informação que conjuga o armazenamento, a disseminação (publicação eletrônica) e a revisão pelo próprio autor e pelos pares, só possível pela aplicação das tecnologias da informação e comunicação.

**2ª Característica:** a capacidade de penetração dos efeitos das novas tecnologias pode ser percebida na comunicação científica, em virtude da adoção dos repositórios digitais, nos novos padrões e métodos da comunicação científica, seja em complementação ao que acaba de ser referenciado na primeira característica, seja na alteração da forma de divulgação, de comunicação e da aceitação internacional da proposição dos repositórios temáticos e institucionais. Há que sublinhar a diferença entre as aplicabilidades na área das ciências sociais que, pela natureza da investigação, precisam manter as pesquisas em algum sigilo, geralmente desenvolvidas de forma solitária, até à divulgação ao final do trabalho, diferentemente do que acontece com as pesquisas na área de ciências físicas e naturais em que há uma interdependência e necessidade da informação recente e atualizada, porque a pesquisa é cumulativa e coletiva, uma vez que utiliza, inexoravelmente, a informação de outros pesquisadores como patamar para o avanço do conhecimento e testagem do que está sendo investigado, pelos pares, o que resulta numa maior importância e uso para esta área do conhecimento.

O contexto da sociedade informacional exige uma crescente demanda por IC&T (informação, ciência e tecnologia). As tecnologias agem e sofrem os efeitos das concepções de espaço e tempo que propiciaram que a comunicação científica fosse alterada substancialmente, com o surgimento e fortalecimento das comunidades científicas virtuais

e com as alterações quanto à acessibilidade e fluência da comunicação que ganhou uma dimensão mundial.

É importante sublinhar, no entanto, que o uso de repositórios digitais é, cada vez mais freqüente na *web*, destinado ao arquivamento de objetos digitais para diferentes fins e usuários.

**3ª Característica:** a lógica das redes é incorporada ao conceito de repositório digital que se encaixa na morfologia da rede, integrada na estrutura de funcionamento, como na evolução dos colégios invisíveis numa nova concepção de avaliação pelos pares, da comunicação em grupos de discussão das comunidades científicas, da interação e da acessibilidade, democrática, aos conteúdos e aos bancos de dados disponíveis nos repositórios, com a incrementação do acesso livre a alguns conteúdos. A adoção do *software open source* que traz também a interoperabilidade tecnológica, e a desejável interoperabilidade humana, como se refere Kuramoto.

**4ª Característica:** a flexibilidade para adaptação e reconfiguração a uma sociedade sempre em mudança e em fluidez organizacional. A política de gestão dos repositórios digitais, o sistema de armazenamento garantido, de certa forma, pela adoção de padrões e protocolos com a solução de interoperabilidade e do acesso aberto pela reuplicação dos metadados, são soluções para a dinâmica dessa sociedade “líquida”<sup>18</sup>. Mas há também o aspecto de que os repositórios digitais são um serviço de informação passíveis de serem integrados na estrutura orgânico-funcional de uma instituição/organização e acompanharem as alterações dessa estrutura na dinâmica do tempo.

**5ª Característica:** que estabelece a convergência de tecnologias específicas para um sistema integrado que tornam invisíveis as trajetórias tecnológicas antigas. Em termos tecnológicos, fica cada vez mais difícil separar os componentes de um sistema de informação, sendo que as telecomunicações são apenas uma forma de processamento da informação, já que as tecnologias de transmissão e ligação são integradas na mesma rede operada por computadores. A Internet implica ser a transmissão de dados a forma predominante e universal de comunicação. E a transmissão de dados fundamenta-se nas instruções de codificação e decodificação de *software*, como acontece com os ideais e

---

<sup>18</sup> Metáfora, em paralelismo à proposta de Zygmunt Baumann para a “Modernidade Líquida” (o que para outros autores seria a pós-modernidade), no sentido de sublinhar a fluidez e a capacidade de assumir diversas formas.

princípios pleiteados pelo acesso aberto. Se os repositórios, como objeto de estudo, surgem no âmbito da pesquisa científica e da comunicação científica, o seu uso é disseminado na Net, para os mais variados fins, porque é uma tecnologia aplicada a diversas demandas.

A construção do conceito de repositórios digitais, como objeto de estudo no âmbito da Ciência da Informação, deve contemplar uma contextualização obrigatória e, ao mesmo tempo, um recorte de especificidade, em razão da extrema complexidade que envolvem a produção de informação e seus fluxos para atingir a comunicação.

## Referências bibliográficas

---

- Alves, Ivone et al. Dicionário de terminologia arquivística. Lisboa: Instituto da Biblioteca Nacional e do Livro; Organismo de Normalização Sectorial para a Informação e Documentação, 1993.
- Baptista, Ana Alice; Costa, Sely Maria de souza; Kuramoto, Hélio; Rodrigues, Eloy. Comunicação científica: o papel da Open Archives Initiative no contexto do acesso livre. Enc. Bibli: Ver. Eletr. Bibliotecon. Ci. Inf., Florianópolis, nº esp.m 1º semestre de 2007. Disponível em <http://www.periodicos.ufsc.br/index.php/eb/article/view/377/435> - Acessado em 20/01/2008.
- Battles, Matthew. A conturbada história das bibliotecas. São Paulo: Editora Planeta do Brasil, 2003.
- Boeres, Sonia A, de Assis; Arellano, Miguel A. Márdero. Políticas e estratégias de preservação de documentos digitais. Anais do VI CINFORM – Encontro de Ciência da Informação, Salvador, 2005. Disponível em [www.cinform.ufba.br](http://www.cinform.ufba.br) Acesso em 30/07/2007.
- Buarque de Holanda Ferreira, Aurélio. Novo Dicionário da Língua Portuguesa, 1ª edição, 15ª impressão. São Paulo: Editora Nova Fronteira.
- Cacaly, Serge. Dictionnaire encyclopédique de l'information et de la documentation. Paris: Nathan, 1997.
- Café, Lígia; Lage, Márcia Basílio. Auto-arquivamento: uma opção inovadora para a produção científica. In DataGramZero. Revista de Ciência da Informação, v 3, n 3, jun/02, Artigo04. Disponível em: [www.dgz.org.br/jun02/Ind\\_art.htm](http://www.dgz.org.br/jun02/Ind_art.htm) - Acessado em 12/02/2007.

Camargo, Ana Maria de Almeida; Bellotto, Heloísa Liberalli. Dicionário de Terminologia Arquivística. São Paulo: Imprensa Oficial e AARQ-SP, 1996.

Carvalho, Maria Carmen Romcy de. Oliveira, Leonardo Silva. Oliveira, Luciana Lima de. Silva, Laura Patrícia da. Mendes, Cristiani Alencar. O repositório institucional como recurso de apoio à gestão do conhecimento da Universidade Católica de Brasília. In 1ª Conferência Iberoamericana de Publicações Eletrônicas no Contexto da Comunicação Científica, Universidade de Brasília, 25 a 28 de abril de 2006. Disponível em: [http://dspace.ibict.br/dmdocuments/Maria\\_Carmen\\_Romcy\\_de-Carvalho.pdf](http://dspace.ibict.br/dmdocuments/Maria_Carmen_Romcy_de-Carvalho.pdf) Acessado em 12/02/2007.

Castells, Manuel. A era da informação: economia sociedade e cultura. Volume I: A sociedade em rede. Trad. Alexandra Lemos, Catarina Lorga e Tânia Soares. Lisboa: Fundação Calouste Gulbenkian, 2002.

Ejarque, Luis Garcia. Diccionario del archivero bibliotecario: terminologia de la elaboración, tratamiento y utilización de los materiales propios de los centros documentales. Gijon: Ediciones Trea, 2000.

Faria, Maria Isabel; Pericão, Maria da Graça. Dicionário do Livro: terminologia relativa ao suporte, ao texto, à edição e encadernação, ao tratamento técnico, etc. Lisboa: Guimarães Editores Lda., 1988.

Flamino, Adriana Nascimento. MARCXML: um padrão de descrição para recursos informacionais em Open Archives. Marília: UNESP, 2006. Dissertação de Mestrado em Ciência da Informação.

Foucault, Michel. A arqueologia do saber. Trad. Luiz Felipe Baeta Neves. Rio de Janeiro: Forense-Universitária, 1972. (Edição Francesa, 1969)

Foucault, Michel. As palavras e as coisas. Trad. Antônio Ramos Rosa. São Paulo: Livraria Martins Fontes Editora Ltda, s.d. [Éditions Gallimard, 1966]



Galvão, Maria Cristiane Barbosa . Construção de conceitos no campo da ciência da informação. Ci. Inf. vol. 27 no. 1 Brasília 1998. Disponível em [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0100-19651998000100006&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19651998000100006&lng=en&nrm=iso) – Acessado em 08/08/2008.

Giddens, Anthony. As conseqüências da modernidade. Trad. Raul Fiker. São Paulo: Editora UNESP, 1991.

Goulemot, Jean-Marie. Bibliotecas, enciclopedismo e angústias da perda: a exaustividade ambígua das luzes. In Baratin, Marc; Jacob, Christian. O poder das bibliotecas: a memória dos livros no Ocidente. Rio de Janeiro: UFRJ, 2006, 2ª ed.; p 257-272

Kuhn, Thomas S. A estrutura das revoluções científicas. São Paulo: Perspectiva, 2006.

Kuramoto, Hélio; Marcondes, Carlos H.; Toutain, Lídia Brandão; Sayão, Luís. Bibliotecas digitais: saberes e práticas. Salvador, BA: EDUFBA; Brasília: Instituto Brasileiro de Informação em Ciência e Tecnologia: 2006; 2ª. Ed.

Kuramoto, Hélio. Blog do Kuramoto. Disponível em (<http://blogdokura.blogspot.com/> - Acessado em 15/07/2007.

Lara, Marilda. Glossário. Población, Dinah Aguiar et alii. Comunicação e produção científica: contexto, indicadores, avaliação. São Paulo: Angellara, 2006.

Leite, Fernando César Lima. Gestão do conhecimento científico acadêmico: proposta de um modelo conceitual. Brasília: 2006. 240 p. Dissertação: Mestrado Ciência da Informação – Programa de Pós Graduação, em Ciência da Informação, Universidade de Brasília. Disponível em: <http://eprints.rclis.org/archive/00006259/> - Acessado em 05 de maio de 2007.

Leite, Fernando César Lima. Cibercultura. São Paulo: Ed. 34, 1999.

López Yepes, José (Editor). Diccionario enciclopédico de Ciencias de la Documentación, 2 vol. Madrid: Editorial Síntesis, 2004.

Meadows, Arthur Jack. A comunicação científica. Brasília, DF: Briquet de Lemos/Livros, 1999.

Ministerio da Cultura: Subdirección General de los Archivos Estatales. Diccionario de Terminología Archivística. Madrid: Gráficas Varona (Salamanca), 1995, 2ª ed.).

Morin, Edgar. O método V: a humanidade da humanidade - a identidade humana. Mem Martins: Publicações Europa-América, 2003.

Mueller, Suzana Pinheiro Machado. A comunicação científica e o movimento de acesso livre ao conhecimento. Ci. Inf., Brasília, v 35, n 2, p 27-38, maio/ago 2006. Disponível em: <http://revista.ibict.br/index.php/ciinf/article/viewArticle/826> – Acessado em 14/04/2007

Pinheiro, Lena Vânia Ribeiro (org). Ciência da informação, ciências sociais e interdisciplinaridade. Brasília/Rio de Janeiro: IBICT/DDI/DEP, 1999. Disponível em: - Acessado em 14/04/2007

Población, Dinah Aguiar; Witter, Geraldina Porto; Silva, José Fernando Modesto da. Comunicação e produção científica: contexto, indicadores, avaliação. São Paulo: Angellara, 2006.

Rodrigues, Eloy. Almeida, Matilde. Miranda, Ângelo. Guimarães, Augusta. Castro, Daniela. RepositóriUM: criação e desenvolvimento do Repositório Institucional da Universidade do Minho. Disponível em: [https://repositorium.sdum.uminho.pt/bistream/1822/422/1/BAD\\_artigo+-+Final.pdf](https://repositorium.sdum.uminho.pt/bistream/1822/422/1/BAD_artigo+-+Final.pdf) Acesso em 20abril de 2007.

Sabbattini, Marcelo. As publicações eletrônicas dentro da comunicação científica. São Paulo: Universidade Metodista de São Paulo, 1999. Disponível em:

[http://bocc/ubi.pt/pag/\\_texto.php3?html2=sabattini-marcelo-publicacoes-electronicas](http://bocc/ubi.pt/pag/_texto.php3?html2=sabattini-marcelo-publicacoes-electronicas). Acesso em 26/11/2007.

Silva, Armando Malheiro da; et al. Arquivística – teoria e prática de uma ciência da informação. Porto: Edições Afrontamento, 1999.

Silva, Armando Malheiro da; Ribeiro, Fernanda. Das “ciências” documentais à ciência da informação: ensaio epistemológico para um novo modelo curricular. Porto: Edições Afrontamento, 2002.

Silva, Armando Malheiro da. A informação: da compreensão do fenómeno e construção do objecto científico. Porto: Edições Afrontamento, 2006.

Silva, Neusa C.; Sá, Nysia O.; Furtado, Sandra R. S. Bibliotecas digitais: do conceito às práticas. Disponível em:  
[http://www.scielo.br/sicelo.php?script=sci\\_arttex&pid=S0100-19651998000200003&Ing=pt&nrm-iso&tlng=pt](http://www.scielo.br/sicelo.php?script=sci_arttex&pid=S0100-19651998000200003&Ing=pt&nrm-iso&tlng=pt) - Acessado em 28/10/07

Silva, Helena Pereira da; Jambeiro, Othon; Barreto, Ângela Maria. Bibliotecas digitais: uma nova cultura, um novo conceito, um novo profissional. In Marcondes et al, Bibliotecas digitais - saberes e práticas. Brasília: IBICT, 2006 (2006a)

Thomaz, Kátia; Soares, Antonio José. A preservação digital e o modelo de referência Open Archival Information System (OAIS). DataGramaZero – Revista de Ciência da Informação, v 5, n 1, fev/04 Artigo 01. Disponível em:  
[http://www.dgz.org.br/fev04/Art\\_01.htm](http://www.dgz.org.br/fev04/Art_01.htm) - Acessado em 14/04/2007

Thomaz, Kátia. Repositórios digitais confiáveis e certificação. Arquivística.net. Rio de Janeiro, v. 3, n. 1, p 80-89, jan.jun. 2007. Disponível em:  
[www.arquivistica.net/ojs/include/getdoc.php?id=372&article=118&mode=pdf](http://www.arquivistica.net/ojs/include/getdoc.php?id=372&article=118&mode=pdf)  
Acessado em 20/02/08.

Weitzel, Simone da Rocha. Reflexões sobre os repositórios institucionais. In Intercom – Sociedade Brasileira de Estudos Interdisciplinares da Comunicação, XXIX Congresso Brasileiro de Ciências da Comunicação. UnB, 6 a 9 de setembro de 2006. Disponível em: [http://www.dgz.org.br/fev04/Art\\_01.htm](http://www.dgz.org.br/fev04/Art_01.htm) - Acessado em 14/04/ 2007.

Weitzel, Simone da Rocha. Os repositórios de e-prints como nova forma de organização da produção científica: o caso da área das ciências da comunicação no Brasil. (Tese Doutorado em CI) São Paulo: USP/ECA, 2006.