



Repositórios Digitais e unidades de informação tradicionais e suas interfaces na Sociedade Informacional

Sílvia Mendes Masson

masson.idt@terra.com.br

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

Resumo

As instituições que desde a Antigüidade custodiam a informação, os Arquivos e Bibliotecas, armazenam a memória do conhecimento humano e 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, inclusive pelas ambigüidades terminológicas e conceituais.

Com o advento das novas tecnologias e a informação em suporte digital, a partir da década de 70 do século passado, no âmbito da sociedade informacional, como a designou Castells, e na conjuntura do paradigma tecnológico, profundas alterações, ainda em curso, aprofundam e intensificam as questões relativas à discussão do campo e objeto de estudo da emergente Ciência da Informação e de suas relações inter e transdisciplinares com a Biblioteconomia e Arquivologia. É nesse contexto que surgem os repositórios digitais, freqüentemente conceituados por suas funções de reunir, preservar, dar acesso e disseminar o conhecimento de uma instituição científica, ou de uma área do conhecimento, dando-lhe visibilidade e se constituindo numa ferramenta de gestão do conhecimento científico.

Na literatura, do ponto de vista epistemológico e teórico-metodológico, repositórios digitais não têm sido objeto de estudo da Ciência da Informação, embora estejam presentes em artigos que o tratam como uma inovadora mídia da comunicação científica, relacionada com a OAI. O objetivo deste trabalho é refletir sobre as proximidades e distanciamentos de tratar o objeto “informação” nos arquivos, nas bibliotecas e nos repositórios digitais.

Palavras-chave: Arquivo. Biblioteca. Repositório Digital. Sistema de Informação. Ciência da Informação. Iniciativa dos Arquivos Abertos -OAI.



1 Sociedade, tecnologia e informação: um novo paradigma

Não é possível pensar em sociedade humana sem conhecimento e informação, sem tecnologia e sem comunicação. Informação é um fenômeno que, necessariamente, envolve comunicação e, no tratamento da informação, importa saber o contexto de produção, avaliar e conhecer os critérios e condições de armazenamento, recuperação, acessibilidade e, ainda, a potencialidade de gerar nova informação/conhecimento, o que inclui a inserção de sociedade, cultura e tecnologia.

Castells, ao analisar a sociedade contemporânea, na obra *Sociedade em Rede*, 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”*. (Castells, 2002:6)

Ao tratarmos de sociedade, informação e tecnologia, estamos também a tratar de cultura. Informação, na concepção de 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 pelos diversos meios tecnológicos, atravessando o tempo e o espaço, desempenham papel capital na evolução da humanidade, substituindo a tradição oral, e criando uma memória coletiva fora das mentes dos membros individuais do grupo. A história da construção da memória exossomática mostra a engenhosidade técnica do homem e sua atividade de criador de símbolos, ou seja, de representar e codificar por signos ou símbolos, para comunicar e, se registradas as representações em suporte material, de produzir informação transmissível em tempos e espaços diferentes daqueles em que foi produzida.

A “memória”, entendida aqui como sinônimo de documento, ou seja, informação codificada e inscrita em suporte, foi de início armazenada em templos, depois em bibliotecas e arquivos e estes repositórios da memória, inseridos na sociedade, agem, reagem e interagem às evoluções tecnológicas e à cultura.



A partir da Segunda Guerra Mundial, ocorre a explosão de informação e a ampliação do conhecimento, com o surgimento de novas áreas de pesquisa. A ciência e a tecnologia tornam-se conceitos muito próximos e a interdisciplinaridade aproxima os conhecimentos, em intercâmbios mútuos e recíprocos, rompendo, pela primeira vez, com a estrutura rigidamente definida das ciências desde o século XVIII.

A revolução tecnológica que se sucedeu a partir da revolução industrial, propulsora de um processo de tecnificação do mundo e de cientificação das atividades sociais é caracterizada pelo papel da informação/conhecimento. Segundo Castells (2002:36),

“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.”

“Sociedade em rede” e “sociedade da informação” são duas das várias designações para esta sociedade que tem na informação o diferencial que a distingue da sociedade industrial. Castells (2002), ao analisar essa sociedade surgida após o industrialismo, destaca a interação dialética entre descobertas científicas, inovação tecnológica e as práticas sociais, que constituíram, na década de 1970, um novo paradigma tecnológico, fundamentado nas novas tecnologias de informação em interação com a economia global, o que resultou na organização de um novo estilo de produção, de comunicação, de gestão e de viver.

A partir de 1980, a revolução das tecnologias de informação foi um fator fundamental para que houvesse uma reestruturação do sistema capitalista, dando origem a uma sociedade capitalista e informacional a que Castells denominou de sociedade informacional, pois comporta uma organização social em que as fontes principais da produtividade e do poder são a produção, o processamento e a transmissão da informação, ao contrário do que aconteceu nas revoluções tecnológicas anteriores em que a informação é que atuou sobre as tecnologias.

No âmbito da sociedade informacional, com os novos contextos dos quais resultou o crescimento excessivo e acelerado da produção de documentos, nas administrações, a Arquivística e os arquivos entraram numa fase de debates a respeito de seu próprio objeto, confrontando-se com a existência de novos suportes e do uso de novas tecnologias, além de uma alteração de demanda, a que Silva et alii (1998:156) denominam de “*uma maior apetência informativa por parte das administrações e da*



própria pesquisa histórica”. Entre os autores que revisam o papel da Arquivística neste novo contexto, destacam-se Gagnon-Arguin, Menne-Haritz, Couture e Rousseau, Luciana Duranti e Eastwood, entre outros. E, entre as diversas proposições, destacamos a defendida por Silva et alii (1998) de que a Arquivística, no paradigma pós-custodial em que os arquivos emergem como sistemas de informação, submetidos à normalização do acesso à informação, é uma ciência epistemologicamente redimensionada, no âmbito transdisciplinar da Ciência da Informação.

A Biblioteconomia e as bibliotecas também passam por revisão de paradigma, emergido no século XIX, insuficiente para atender aos efeitos da revolução tecnológica da sociedade informacional, que exige a adequação às atividades relacionadas com a forma digital da informação, seja a produção, a edição, o armazenamento, a distribuição, a recuperação e o uso da informação, ou todas as formas de trabalho que lidem diretamente com dados textuais, simbólicos, visuais e auditivos, tudo isto atravessado por uma alteração radical: a informação cada vez menos ligada ao suporte, e usuário com outras necessidades e expectativas, em razão da automação e das práticas culturais que levam a anseios de acessibilidade em novas dimensões e em novos conceitos.

Bauwens (apud Levacov, 1997:3) afirma que a profissão de bibliotecário era tradicionalmente ligada a um produto, o livro, e a um local, a biblioteca e, diante das tecnologias de informação e comunicação e das noções de espaço e tempo, que ganham dimensões inéditas peculiares da sociedade informacional, está sendo modelado o que Browning (apud Levacov, 1997:3) chamou de “bibliotecas sem paredes para livros sem páginas”. Uma hipérbole metafórica, é óbvio, mas que tem procedência, desde que se sublinhe que em comunicação, uma nova tecnologia não elimina radicalmente as tecnologias que a antecederam.

Biblioteca, serviço ou sistema de informação, destina-se a incorporar e tornar acessível informação editada e posta a circular, através da facilitação e difusão desses conteúdos, portanto inteiramente suscetível às novas tecnologias da informação e da comunicação, do novo paradigma tecnológico.

O novo paradigma tecnológico, a partir dos anos 70, do século XX, até a esta data, fundamenta-se na informática que integra vários modos de comunicação, em rede interativa a partir de múltiplos pontos, em tempo escolhido (real ou passado), em rede global, em condições de acesso livre e a preço módico, com o recurso do hipertexto e a

metalinguagem, reunindo no mesmo sistema as modalidades escrita, oral e audiovisual da comunicação humana, alterando de forma fundamental o caráter da comunicação e moldando a cultura. (Castells, 2002:432)

É um período de profundas transformações da sociedade e também do conhecimento, gestadas a partir da revolução industrial e durante o século XIX, mas que somente na segunda metade do século XX se acelera, em razão de uma sociedade baseada na informação, em que a ciência e a tecnologia atingem uma relação de interdependência tal que faz surgir as tecno-ciências que alimentam e se retroalimentam do universo da informação.

Para Morin, “*vivemos uma era histórica em que os desenvolvimentos científicos, técnicos e sociológicos estão cada vez mais em inter-retroações estreitas e múltiplas*”, em que a técnica, produzida pelas ciências, transforma a sociedade, e em que a sociedade tecnologizada transforma a própria ciência. (Morin, 2002:19-20).

O conceito de cultura apresentado por Morin enfatiza a perspectiva semiótica de cultura, ou seja, a concebe como um conjunto de informações, destinado às organizações sociais :

“A 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.” (Morin, 1969:17)

Silva, afinado com o conceito de cultura de Morin, assim define informação:

“Conjunto estruturado de representações mentais e emocionais codificadas (signos e símbolos) e modeladas com/pela interação social, passíveis de serem registadas num qualquer suporte material (papel, filme, banda magnética, disco compacto etc) e, portanto, comunicadas de forma assíncrona e multi-direccionada.” (Silva: 2006: 150/151),

A disciplina que emerge para o estudo da informação é a Ciência da Informação, cuja definição surgida nas conferências do *Georgia Institute of Technology*, dos EUA, no início da década de 1960, foi retomada por Harold Borko em 1968. A partir dessa definição de Borko, Silva entende que a 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.” (Silva, 2006:150-151)

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 trata, ou deveria tratar, de arquivos, de bibliotecas, de sistemas e serviços de informação e de repositórios digitais enquanto instituições que se destinam a guardar e tornar acessível informação.

Não nos deteremos na revisão das definições e campos de estudo da ciência da Informação e não evocaremos a diversidade de autores e suas concepções relativas ao caráter disciplinar da Ciência da Informação, mas optamos pela concepção de Silva e Ribeiro (2002) e Silva (2006) e é sob estas referências que realizamos as reflexões sobre unidades de informação, uma vez que todas têm em comum o objeto “informação social”. E o fazemos, porque é necessário ter uma coerência teórico-metodológica para analisar a questão, além de que, ao estabelecer a fundamentação em conceitos e definições, claramente delimitados, criamos uma plataforma unificada que permite evitar discrepâncias e conflitos conceituais, além de aumentar a garantia de uma descodificação mais unissêmica .

O compartilhamento do objeto de estudo, a informação, por diversas disciplinas, leva também ao compartilhar de vocábulos em distintas terminologias, o que torna fundamental a elaboração de dicionários e glossários, constantemente revistos para a apreensão da dinâmica da evolução científica, e buscar a eficiência da comunicação.

“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.” (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 freqüentemente presentes em artigos que o tratam como uma importante mídia da comunicação científica, relacionado com a OAI – *Open Archives Initiative*. 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 de repositórios digitais são, em sua maioria, escritos por bibliotecários que estão focados com os estudos de mediação e acesso, incluindo os repositórios digitais em suas investigações, apenas como uma referência para abordar casos específicos ou para divulgar instituições e serviços pelos quais são responsáveis.

2 Repositórios Digitais: a comunicação científica na sociedade informacional

Os repositórios digitais emergem no contexto da assim designada, por Castells (2002), sociedade informacional e do informacionalismo, 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 tecnológicas e sociais em ritmo que se acelera à medida que se afastam das sociedades industriais.

Os repositórios digitais são uma alternativa para o tradicional sistema de comunicação científica, uma vez que, na área científica, a publicação dos resultados é a forma pela qual ocorre o confronto com os pares e do qual resulta, ou não, a certificação de validade do conhecimento científico. No entanto, o modelo de comunicação científica, através de revistas científicas, nas duas últimas décadas do século XX, principiou a gerar insatisfação na comunidade científica, em razão do grande aumento da quantidade de títulos; da elevação do preço dos periódicos, muitas vezes inviáveis para manter o acervo em instituições de menor porte ou dependentes de verbas de governos; das dificuldades na distribuição e no crescimento de custos em razão das distâncias geográficas, impedindo que centros de pesquisa de menor expressão tivessem acesso à informação atualizada; a não remuneração dos direitos autorais dos pesquisadores, mas a cobrança, por alguns editores comerciais, despesa, às vezes, assumida pelos autores ou pelas instituições, uma vez que almejam a validação e a visibilidade que a publicação certifica e acarreta prestígio junto à comunidade científica; o prolongado tempo entre a entrega dos artigos e sua publicação, numa sociedade em que o conhecimento e as tecnologias são rapidamente substituídos e alterados; a dificuldade de consultar e recuperar a informação disponível, mesmo que esteja disponível em bancos de dados, mas muitas das vezes, não funcionando em programas interoperáveis de recuperação etc.



A disseminação dos repositórios digitais e o advento da *Internet* envolvem a cúpula científica, representada pelos cientistas das universidades e dos centros de pesquisa, os recursos das tecnologias de informação e comunicação disponíveis e a inovação contra-cultural. Tanto a *Internet*, quanto os repositórios digitais tornam mais próximo o alcance do antigo sonho do acesso livre à informação/conhecimento. Retorna-se aos tempos da ação contra-cultural, dos anos 60, que tem a sua continuidade, em outro contexto, através dos movimentos do *Open Archives Initiative* e do *Open Access Movement*.

O movimento *Open Archives Initiative* resulta do objetivo de ampliar o acesso a repositórios de artigos científicos como uma alternativa para aumentar a transmissibilidade da comunicação científica, corrigindo as dificuldades detectadas com os repositórios digitais pioneiros, através da 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.

O Movimento *Open Access*, ou Movimento Acesso Livre, advém da preocupação com o compartilhamento do conhecimento, da concepção de que a inclusão científica leva a uma maior inclusão social e da disponibilidade de maior acesso à informação.

Ao final da década de 1990, já existiam repositórios de artigos científicos e foram criados com o objetivo de agilizar a comunicação científica, multiplicando assim a idéia inicial do ArXiv, criado em 1991, por Paul Ginsparg. O primeiro computador em rede foi ativado em 1º de setembro de 1969, com os primeiros quatro nós de rede localizados na Universidade da Califórnia (Los Angeles), no Instituto de Investigação de Stanford, na Universidade da Califórnia, em Santa Bárbara, e na Universidade de Utah. Esta experiência surgiu para atender aos institutos de investigação que cooperavam com o Departamento de Defesa norte-americano, mas passou a ser utilizada pelos cientistas, para os seus próprios propósitos de comunicação científica. Daí o uso extrapolou para as conversas pessoais, até que, em 1983, foi feita a separação entre a ARPANET, exclusiva para fins científicos, e a MILNET, para fins militares.

Nos anos 80, algumas instituições científicas criaram outras redes e surgiu a rede das redes: a Internet. Diante destas novas tecnologias e, em paralelo, da necessidade de agilizar os fluxos da comunicação científica, os próprios cientistas pesquisadores criam



uma solução pragmática, diante dos recursos disponíveis: arquivos digitais depositados virtualmente para acesso por outros computadores. Denominados inicialmente “arquivos públicos digitais”, os repositórios digitais nascem, na década de 1990, da inventividade de pesquisadores, envolvidos com iniciativas isoladas para publicação na *web* que, ao experienciarem a criação de um ambiente digital para a comunicação científica de alta velocidade, com tecnologia de baixo custo, com ampla participação da própria comunidade científica, num só local, tanto na gestão desse ambiente quanto na geração, disseminação e no uso da informação, deram origem a uma “revolução” nos paradigmas tradicionais da comunicação científica. Esse foi um movimento inspirador para outras iniciativas, liderado pelos EUA e pesquisadores europeus. (Weitzel, 2005:174-175)

Steven Harnard, a partir de 1995, lança a proposta para que o cientistas passassem a publicar, por meio da *web*, em forma de *preprints*, o que permitiria inclusive a vantagem de recebimento de comentários, revisão e, ao mesmo tempo, ficar disponível para acesso livre. Estas duas características serão embrionárias para a iniciativa de “arquivo aberto” e movimento de “acesso livre”.

Entre 21 e 22 de outubro de 1999, reunidos em Santa Fé (Novo México), num fórum que reuniu bibliotecários, profissionais de informática e os principais fundadores dos chamados “arquivos públicos” resultou na formação e organização do movimento *Open Archives Initiative*, com a proposição do arquivamento público e universal da produção acadêmica e científica, denominando esse modelo de arquivamento de “*open archives*”. (Ginsparg, Luce e Van de Sompel (1999), apud Weitzel, 2005: 176)

Na proposta, nascida em Santa Fé, os “arquivos públicos” destinam-se a designar um repositório de *papers* acadêmicos, para o armazenamento de informação. Lagoze e Van de Sompel (2001, apud Weitzel, 2005:178) revelam que o significado, em que deve ser denotado o termo arquivo, na denominação *Open Archives Initiative*, não é o da terminologia Arquivística. E o adjetivo “aberto”, denota o sentido que tem em arquitetura de sistema informático, ou seja, que permite interfaces que facilitam a disponibilidade de conteúdos, tomando como ponto de partida a variedade de provedores. Trata-se, portanto, de “arquivo”, no sentido que tem em Processamento de Dados, ou seja, arquivo como um conjunto de dados relacionados, tratados como uma totalidade, e o adjetivo “aberto” (*open*), referindo-se à interoperabilidade.



Em artigo publicado em 2007, Weitzel (2007:2) esclarece que o modelo estrutural dos “*open archives*” fornece as bases da publicação científica na *Internet* e atende aos padrões e protocolos que certificam sua origem e permanência. Esta infraestrutura é composta por provedores de dados que armazenam os objetos digitais, e por provedores de serviços que coletam os metadados contidos nos provedores de dados, resultando na possibilidade de disseminação ampla da produção científica, com a permissibilidade de realizar a busca nos provedores de dados e também a busca integrada por meio dos provedores de serviços. Com essa infra-estrutura é possível reagrupar as publicações *on-line* em repositórios digitais, cujos metadados são igualmente coletados pelos provedores de serviços, funcionando com uma segunda instância de disseminação, reunindo a produção científica publicada. Podem ser repositórios digitais institucionais e temáticos.

Consta que a primeira apresentação, no Brasil, sobre *open archives*, foi por Nathália Kneipp Sena, jornalista e analista em ciência e tecnologia da Gerência de Novas Tecnologias, do IBICT, em matéria escrita para a Revista Ciência da Informação, no final do ano de 2000, em que a Autora define “*como arquivos que congregam e-prints das diversas áreas do saber e que são abertos à consulta pública bem como à publicação automatizada dos trabalhos por parte dos pesquisadores.*” (Sena, 2000: 2)

Com essa matéria/artigo, temos o marco inicial, no Brasil, de uma sucessão, não tão numerosa, de outros trabalhos, sejam artigos publicados em revistas ou apresentados em eventos, como congressos, encontros e seminários na área da Ciência da Informação e das Ciências da Comunicação, sejam teses e dissertações sobre esse tema, relativamente recente, que impactou o sistema de comunicação científica a partir da criação do primeiro arquivo eletrônico de *e-prints*, em 1991, nos EUA, por Ginsparg, destinado à Área da Física, no Laboratório Nacional de Los Alamos, no Novo México, depois ampliado para as áreas de Matemática e Ciências da Computação. (Ginspard, Paul, apud Café e Lage, 2002)

Sena se referia aos “*open archives*” e reproduzia a proposta “subversiva” de Harnard para que a atividade editorial acadêmica passasse a ser sinônimo de transferência da pesquisa acadêmica para a *Internet*. Essa aspiração teria sua materialização na forma de arquivos abertos (*open archives*) de *e-prints*, “*através da padronização das tecnologias de informação e comunicação empregadas nesses repositórios, em bibliotecas digitais e outras aplicações conexas*”. Destacava também



que a implementação dessa proposta envolvia um esforço inter e multidisciplinar, que reunia informáticos, profissionais da informação e de diversas áreas do saber que deveriam trabalhar juntos para encontrar um caminho alternativo para a comunicação científica, que atravessara décadas de avanços tecnológicos até encontrar o “surgimento e a consolidação dos *open archives*, arquivos *on-line* de acesso público, também definidos como ‘diretórios existentes em um computador que estão abertos para o acesso via *ftp* ou *http*, armazenando uma coleção de séries de artigos ou uma coleção de dados sobre artigos armazenados em outro local”.

A partir dessa primeira divulgação dos repositórios digitais, o IBICT desenvolve projetos, como a biblioteca digital de teses e dissertações e enceta uma série de atividades para a divulgação e implementação de modelos para uma interoperabilidade nacional da comunicação científica. Novos vocábulos entram para a terminologia de quem escreve e/ou pensa sobre comunicação científica e repositórios digitais: “arquivos abertos de *e-prints*”, “*open archives*”, interoperabilidade, arquivamento de longo prazo, auto-arquivamento e uma série de outros, inclusive de informática e das TIC, como metadados, “*harvester*”, provedores de dados, provedores de serviços etc., e se acumulam tentativas de definir ou conceituar, num processo de conceitualização ainda em construção, inclusive em que a própria denominação esteve “instável”: arquivos públicos digitais, arquivos de *e-prints*, *open archives*, arquivos abertos e repositórios digitais, que parece ser a que deva se manter.

A construção de conceitos obedece à evolução cognitiva, que parte da exploração do que algo é, por suas semelhanças e diferenças com o que já é conhecido, ressaltando as suas características e funções. Um objeto de estudo em construção exige “caracterizações” para ser conhecido e reconhecido e esta é uma fase que pode levar a muitas tentativas, todas contendo faces desse objeto, mas não conseguindo captá-lo inteiramente. Não é diferente com os repositórios digitais, em razão de sua criação recente, há menos de dezoito anos, e, no Brasil, há menos de oito, o conceito está em construção, como dissemos, e apresenta dificuldades de clareza e/ou tautologismos, nas tentativas de definir para alcançar o conceito: repositórios são coleções digitais que armazenam, ou um arquivo digital que reúne coleção de documentos digitais, ou uma forma de armazenamento de objetos digitais, ou são sistemas de informação que servem para armazenar.

Destacamos alguns desses conceitos:

Quadro I: Conceitos de Repositórios Digitais

Autor/Citação	Conceito
Crow (2002), apud Leite (2006:88)	“Repositórios digitais são “coleções digitais que capturam e preservam a produção intelectual da comunidade de uma única universidade ou de uma comunidade multiuniversitária”.
Weitzel (2006:5-6)	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.
Viana, C. L.M. et alii.	“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.” http://bibliotecas-cruesp.usp.br/3sibd/docs/viana358.pdf
Universidade do Minho	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.”

Elaborado por Masson-2008: Fontes: Vide Referência)

A palavra repositório, no sentido vernacular, denota algo capaz de armazenar, assim como arquivos, bibliotecas e centros de documentação, mas com particularidades que resultam da conjuntura histórico-econômico-técnico-científico-cultural em que são criados, e que na origem tinham o objetivo de atender às necessidades de comunicação da comunidade científica, mas extremamente fundamentados em ideais de acessibilidade a esse conhecimento, a partir de segurança de interoperabilidade técnica, mas também e, principalmente, de interoperabilidade humana. (Kuramoto, 2007: [http://kuramoto.wordpress.com/.](http://kuramoto.wordpress.com/))

Os repositórios digitais surgem e passam a ser disseminados, no contexto da chamada sociedade informacional e do informacionalismo (Castells, 2002), como serviço de informação em busca de fluxos de comunicação mais eficientes, porém impregnados da cultura do “ambiente” de contexto. As unidades de informação - bibliotecas, arquivos e centros de documentação - que tradicionalmente armazenavam documentos e os reuniam num lugar físico, um prédio, com a virtualização dos documentos, privilegiam a informação e surgem os arquivos, bibliotecas, centros de documentação e até museus digitais, em razão da contribuição e mudanças causadas pelo uso das TIC, que no caso do sistema de comunicação científica, propiciaram as condições para a criação de repositórios digitais, um “sítio” ou espaço virtual, a que Castells denomina o espaço de fluxos. Nesse espaço de fluxos, a publicação eletrônica da informação científica acontece, com operacionalidades peculiares do paradigma tecnológico da sociedade informacional: acesso público, independente do espaço e tempo, passível de ser conectado de qualquer ponto do planeta; o auto-arquivamento; lista de discussão que propicia a revisão e a validação da informação científica, pelos



pares, em processo dinâmico e eficiente; o armazenamento e garantia de acessibilidade, independente das mudanças tecnológicas e diversidades de *hardware* e *software* – interoperabilidade estabelecida por protocolos; as alterações no fluxo informacional, com independência de editoria, de disseminação, de mediação para o acesso, visibilidade institucional e projeção e reconhecimento de pesquisadores e pesquisas, etc.

3 Sistemas de Informação: as interfaces emergem da natureza do objeto “informação social”

As instituições que desde a Antiguidade custodiavam a informação, os Arquivos e Bibliotecas, armazenavam a memória do conhecimento humano ou de suas atividades, e se constituíram nos repositórios de 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.

A polissemia das palavras “biblioteca” e “arquivo”, leva-nos à necessidade de definirmos, ou caracterizarmos, 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 objetivos 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”, inteligível, 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, Sá, Furtado (2004:3 apud Silva et al, 2006a: 171), chegaram à seguinte síntese conceitual:

“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)

O conceito de biblioteca digital, em ambas as citações, cabe perfeitamente no conceito de repositório digital; na segunda citação, ajusta-se, quase em perfeito, à caracterização geral, e há quem sublinhe que toda biblioteca digital é um repositório digital, mas nem todo repositório digital é uma biblioteca digital, em razão do uso que dele faz a comunidade científica, como um elemento do fluxo informacional, porém sem eliminar as outras formas de publicação científica. Neste sentido, no caso do conceito de biblioteca digital, retrocitado, se nos determos ao uso dos repositórios digitais exclusivamente na comunicação da informação científica, teríamos apenas que delimitar a informação, exclusivamente a informação científica. Aliás, como é o entendimento dos que consideram que a Ciência da Informação tem como objeto de estudo apenas a informação científica, entendendo por informação, apenas a informação/conhecimento científico.

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.” (Camargo e Belotto, 1996:5)

Silva, na “Terminologia Essencial”, da sua proposta teórico-metodológica de Ciência da Informação, arquivo pode ser entendido 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)

O conceito de arquivo, na terminologia Arquivística tradicional, ou na Arquivística como uma transdisciplina no campo da Ciência da Informação, é perfeitamente ajustável ao que se conceitua como repositório digital institucional - acumula informação produzida como atividade-fim e atividade-meio de uma instituição, independente de ser laboratório de pesquisa ou universidade. Ora, acontece que um laboratório de pesquisa ou uma universidade se constituem em organização com objetivos a serem alcançados, a partir de uma estrutura orgânico-funcional, instituída para atingir a esses objetivos, portanto há um sistema de informação nessa organização que resulta de suas atividades e produz informação. A informação científica é um dos fluxos que circula nesse sistema informacional e os repositórios digitais seriam um sub-sistema para prestar uma modalidade de serviço específico, na estrutura orgânico-funcional do sistema de informação constituído pelo Laboratório, ou pela Universidade, que geram também informação/documento de outras naturezas e que estão relacionados com a informação científica, porque a organização é sistêmica e a sua produção informacional passível de tratamento arquivístico.

O repositório digital temático, que acumula informação produzida por uma entidade criada para colecionar informação, com o objetivo de atender a um tema, ou qualquer outro critério de seleção para constituir a coleção, também é, ou está incluído em uma organização, ou seja, é um sub-sistema dentro de uma instituição, um serviço de informação, ou uma instituição, ela própria criada para esse fim: acumular e disponibilizar informação. Neste caso, também dotada de estrutura orgânico-funcional instituída para essa finalidade, e estamos a falar de um serviço prestado por bibliotecas: acumular, organizar e disponibilizar informação.

Trabalhando com o conceito de Ciência da Informação, proposto por Silva (2006:140), o objeto informação social, sem outras adjetivações restritivas, se científica,

técnica, etc., é o objeto de estudo da CI, que analisa e trata das propriedades inerentes à gênese do fluxo, organização e comportamento informacionais, o que inclui a origem, coleta, organização, armazenamento, recuperação, interpretação, transmissão, transformação e utilização da informação, portanto também capaz de estudar, epistemologicamente, os repositórios digitais.

Silva (2006: 162-163) alerta, no verbete “sistema de informação”, que se trata de uma expressão sujeita a ambigüidades semânticas e que o conceito operatório, fundamental em Ciência da Informação, é comumente confundido com sistema tecnológico de informação ou informático. O conceito, que aqui estamos utilizando, é o que Silva nos fornece:

“um Sistema de Informação é constituído pelos diferentes tipos de informação registada ou não externamente ao sujeito (o que cada pessoa possui em sua memória é informação do sistema), não importa qual o suporte (material e tecnológico), de acordo com uma estrutura (entidade produtora/receptora) prolongada pela acção na linha do tempo”. (Silva, 2006:162)

Abaixo, a partir de um quadro comparativo apresentado por Bellotto, com o objetivo de demarcar características de diferentes unidades de informação, acrescentamos, nas colunas em cinza, repositório digital e sistema de informação, para dar visibilidade de que unidades de informação são sistemas de informação, incluindo entre eles, o repositório digital que pode ser, em si, um sistema de informação, no caso dos repositórios temáticos, ou um sistema de informação que presta um serviço de informação, no contexto orgânico funcional de uma instituição.

Quadro II:

Fonte: *Bellotto, (1991, 18);

** A partir de Silva & Ribeiro (2002) e Silva (2006), elaborada por Masson (2008), para enfatizar a caracterização de SI e RD em relação às outras unidades de informação.

Quadro Comparativo Entre Unidades de Informação a Partir da Caracterização de Bellotto

	*Arquivo	*Biblioteca	*Museu	*Centro de Documentação	Sistema de Informação **	Repositório Digital **
Tipo de Suporte	Manuscritos, impressos, audiovisuais: exemplar único	Impressos, Manuscritos, audiovisuais, exemplares múltiplos	Objetos bi-tridimensionais; Exemplar único	Audiovisuais (reproduções) exemplar único ou múltiplo	Analógico e eletrônico	Digital (inclui informação codificada em código textual, imagético, sonoro, numérico etc.)
Tipo de Conjunto	Fundos: documentos unidos pela origem	Coleção: documentos unidos pelo conteúdo	Coleção; documentos unidos pelo conteúdo ou função	Coleção: documentos unidos pelo conteúdo	Informação reunida pela origem orgânico-funcional, agregada a toda informação que a contextualiza, instrumentalize ou fundamente a ação para se atingir ao objetivo da organização.	Coleção: informação reunida por instituição ou por tema

Produtor	A máquina administrativa	Atividade humana e coletiva	Atividade humana, a natureza	Atividade humana	Organização em todas as atividades humanas, seja com fim científico, técnico, artístico, educativo, administrativo que produza informação social.	Atividade Científica
Fins de Produção	Administrativos, Jurídicos, funcionais, legais	Culturais, científicos, técnicos, artísticos, educativos	Culturais, artísticos, funcionais	Científicos	Administrativo, Jurídico, Funcional, legal, cultural, científico, técnico, artístico, educativo.	Info-comunicacional da atividade científica
Objetivo	Provar, testemunhar	Instruir, informar	Informar, entreter	Informar	Informar, permitir a disponibilização da informação, contextualizada, para a tomada de decisão eficaz e eficiente, provar, testemunhar, permitir a transparência administrativa, exercício da cidadania e preservar a memória institucional.	Informar, instruir, validar, opinar, provar, testemunhar etc.
Entrada dos documentos	Passagem natural de fonte geradora única	Compra, doação, permuta de fontes múltiplas	Compra, doação, permuta de fontes múltiplas	Compra, doação, pesquisa	Informação avaliada a partir da natureza orgânico-funcional sistêmica, desde o momento de sua concepção, mantida a ordenação em que é produzida, integrada a toda a informação que com ela tenha vínculo ou pertinência, seja no sentido de complementar ou fundamentar àquela produzida pela Organização.	Auto-arquivamento e captura por "harvesters"
Processamento técnico	Registro arranjo descrição: guias, inventários, catálogos etc.	Tombamento, classificação, catalogação: fichários	Tombamento, classificação, catalogação: inventários, catálogos	Tombamento, classificação, catalogação: fichário, computador	Informação mantida segundo a estrutura que a produziu: quadro sistêmico (orgânico-funcional), relacionada com toda a fonte de informação que com ela se integre, devidamente indexada de maneira a permitir o acesso de acordo com as necessidades do usuário → mecanismos de buscas e de indexação automatizados	Informatizado (metadados, robôs, mecanismos de buscas e de indexação automatizados, software aberto etc.)
Público	Administrador e Pesquisador	Grande público e pesquisador	Grande público e pesquisador	Pesquisador	Administrador, cidadãos, pesquisador e quem necessitar da informação, exceto as restrições previstas em Lei.	Usuário / pesquisador

Bibliotecas e arquivos, inclusive os digitais, 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 função dar tratamento a essa informação. O que determina o tipo de tratamento e o tipo de unidade de informação a que se destina, não é o suporte, mas o uso que se fará dela e objetivo pelo qual a informação é produzida. A função essencial das bibliotecas



não se altera, substancialmente, ao longo das evoluções tecnológicas, mas no caso do suporte digital ganha outras dimensões porque há um contexto técnico-cultural que exige novos procedimentos. O mesmo acontece com o arquivo, centro de documentação, museu etc. A informação em suporte digital integra, em seu tratamento, 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 social é produto decorrente de uma organização (pessoa jurídica ou física), que visa a um objetivo estabelecido por contrato social, lei ou interesses de um indivíduo, alcançado através de atividades que resultam em um fluxo de informação que se constitui no sistema de informação dessa organização, destinar-se-á ao arquivo, como unidade de informação. Se a informação é colecionada em razão de um tema escolhido, de uma área do conhecimento ou de uma atividade humana, com o destino de armazenar e disponibilizar essa informação, em razão de que a organização se propõe a ser o repositório da informação referente ao critério de coleção selecionado, haverá uma estrutura orgânico-funcional da unidade de informação para alcançar o objetivo de sua criação, que gerará informação da atividade-meio, e o sistema de informação, resultante da coleção realizada, constituído pela informação coletada em razão da atividade-fim. Para o usuário, interessa apenas o sistema de informação constituído pela informação da atividade-fim, porque a aquisição desta informação é realizada de forma múltipla, por compra, doação etc., e agora, através da interoperabilidade de repositórios digitais..

A razão para uma informação destinar-se ao arquivo ou biblioteca não está no suporte, mas no objetivo da instituição e na forma como a informação é produzida/acumulada. O tratamento ao suporte da informação acontece, desde sempre, com a preocupação do armazenamento, da conservação e preservação de acordo com as características físico-químicas do suporte, e é assim desde as tabuinhas de argila, passando pela evolução tecnológica dos suportes, sejam rolos de papiro, manuscritos, livros, plantas, vídeos, gravações, etc. Por que é então que surgem os repositórios digitais, nas modalidades institucional e temático, em vez de serem serviços ou sistemas de informação, como arquivos ou bibliotecas digitais?

Talvez porque emergem da necessidade prática, em que o próprio pesquisador, usou os recursos tecnológicos e informáticos de que dispunha, para resolver questões



que passariam a independer de mediadores, atalhando caminhos e dificuldades, criando outras, solucionadas pelo OAI e adotando procedimentos de acesso livre, imbuídos de uma ideologia que emergiu nos movimentos de contra-cultura da década de 60, e resultando em alternativa para o modelo clássico da comunicação científica, passível de conviver com ela, mas trazendo alterações irreversíveis. No entanto, estas alterações estão mais relacionadas com o paradigma tecnológico em que vive a sociedade contemporânea, do que com a forma e o suporte em que a informação é produzida.

Repositórios digitais armazenam informação digital e os procedimentos adotados, para o seu funcionamento, têm menos a ver com o sistema de informação que são, do que com a natureza do serviço que prestam e, funcionalmente, estão integrados a uma organização que produz ou coleciona informação. Constituem uma experiência que assume as alterações nas práticas sociais e culturais, advindas, do paradigma tecnológico da sociedade informacional e do informacionalismo, como a que acontece em outras atividades da vida social, por exemplo a desintermediação de serviços bancários e de outros atendimentos e serviços hoje feitos de maneira direta pelo usuário, em novas noções de tempo e espaço, numa dinâmica peculiar da sociedade informatizada.

No caso dos repositórios digitais, tanto o pesquisador, como produtor ou usuário de informação científica, ganha autonomia e a mediação, que antes era feita por pessoas, num determinado lugar, em horários de expediente, submetido a “ritos” e normas, é agora realizada mediante outras práticas e com versatilidades, tornadas possíveis pelas TIC e pelo paradigma do informacionalismo.

Na sociedade informacional, os recursos que permitem automação de serviços, proporcionam autonomia e independência, eficácia e eficiência, no armazenamento e no acesso à informação. O pesquisador que antes dependia das editoras para publicar suas pesquisas, hoje o faz através dos repositórios, de uma forma mais ágil e eficiente, a baixo custo, maior poder de disseminação, possibilitando que o cientista tenha ampla divulgação de seus trabalhos, e o usuário, um estoque diversificado e acessível com mais conforto e livre de condicionalismos de tempo, espaço e de normas rígidas de acesso e/ou consulta, cobrindo melhor índice de aproveitamento, impensável, sem a automação.

O fato é que as funcionalidades dos repositórios digitais extrapolam o nicho das comunidades científicas e universitárias e, se o OAI teve sua origem do objetivo de



ampliar o acesso aos repositórios de artigos científicos, os padrões e ferramentas tecnológicas desenvolvidos são independentes do conteúdo.

A Ciência da Informação, como aqui foi compreendida, se constitui a área de estudo em que o repositório digital poderá integrar abordagens epistemológicas que possibilitarão compreender a sua natureza peculiar e as interfaces com as unidades de informação percebidas como sistemas de informação.

Referências

BELLOTTO, Heloísa Liberalli. Arquivos permanentes: tratamento documental. São Paulo: T.A. Queiroz, 1991.

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.

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, Artigo 04.

GALVÃO, Maria Cristiane Barbosa. Construção de conceitos no campo da ciência da informação. Ci. Inf., Brasília, v. 27, n. 1, pp. 46/52, jan./abr. 1998.

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. Acessado em 05/05/2007.

LEVACOV, Marília. Bibliotecas virtuais: (r) evolução? Ci. Inf., v. 26, n. 2, Brasília Mat/Aug, 1997.

MORIN, Edgar. Ciência com consciência. Trad. de Maria D. Alexandre e Maria Alice Sampaio Dório. Edição revista e modificada pelo autor. Rio de Janeiro: Bertrand , Brasil, 2002.

SENA, Nathália Kneipp. Open Archives: caminho alternativo para a comunicação científica. Ci. Inf. V. 29, n.3, Brasília, set./dez. 2000. Acessado em 08/10/07. http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19652000000300007&Ing=en&nrm=iso

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

_____. A informação: da compreensão do fenômeno e construção do objecto científico. Porto: Edições Afrontamento, 2006.



VIANA, Cassandra Lúcia de Maya; Arellano, Miguel Angel Márdero. Repositórios institucionais baseados em DSPACE e EPRINTS e sua viabilidade nas instituições acadêmico-científicas. Acessado em 20/04/08. <http://eprints.rclis.org/archive/00008488/>

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 de Doutorado. Universidade de São Paulo. São Paulo, 2006.

_____. a) 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. Disponibilizado na net, acessado em 14 de abril de 2007.

_____. E-prints: Modelo da comunicação científica em transição. Ferreira, Sueli Mara Soares Pinto; Targino, Maria das Graças. (Orgs) Preparação de revistas científicas. São Paulo: Reichmann & Autores, 2005. Capítulo 6, pp. 161-193