

O papel dos repositórios institucionais e temáticos na estrutura da produção científica¹

Simone da Rocha Weitzel

RESUMO

A proposta deste artigo é contribuir para a compreensão do novo cenário da comunicação científica no século XXI por meio de algumas reflexões sobre o papel dos repositórios institucionais e temáticos neste contexto, a partir das noções sobre a estrutura da literatura científica de Subramanyam (1980). O ponto de partida desta abordagem está focado nas características dos produtos (publicações, repositórios digitais e provedores de serviços), processos (editoração eletrônica e comunicação) e suas correspondências na estrutura na literatura científica. Conclui-se que o ciclo da produção científica online é composto pelas fontes primárias (publicações científicas online), secundárias (repositórios temáticos e institucionais) e terciárias (provedores de serviços) e que a implementação e uso destes três tipos de fontes são urgentes para promover o desenvolvimento científico na atualidade.

PALAVRAS-CHAVE: Repositórios institucionais. Repositórios temáticos. Comunicação científica eletrônica. Iniciativa de Arquivos Abertos. Movimento de Acesso Livre.

¹ Este trabalho traz resultados parciais do relatório de pesquisa apresentado no exame de qualificação do Programa de Pós-Graduação em Ciência da Informação da Universidade de São Paulo em 28 abril de 2005. Agradecimentos especiais à Prof.^a Dr.^a Sueli Mara Ferreira pelas sugestões, à Cláudia Figueiredo pela revisão gramatical e aos avaliadores da revista *Em Questão* pelas contribuições valiosas ao artigo.

1 Introdução

A produção científica pode ser entendida como um recurso imprescindível para promover o desenvolvimento da ciência. Sua origem remonta à constituição e consolidação de todo um sistema de comunicação científica no mundo e sua história está vinculada à história da própria ciência, de forma que sua estrutura foi acompanhada pela especialização dos saberes e pela autonomização do campo científico ao longo dos últimos quatro séculos.

Nesse processo, a revista científica, dentre os outros tipos de comunicação escrita, tornou-se o principal marco da constituição da estrutura da comunicação científica, pois surgiu da necessidade genuína de trocas de experiências científicas dos cientistas desde os tempos modernos da história universal.

No entanto, a partir da última década do século XX este cenário tem se modificado e está em pleno estágio de reorganização dos processos e produtos da comunicação científica, por meio da adoção das tecnologias da informação e comunicação, e da consolidação de algumas iniciativas, principalmente a Iniciativa de Arquivos Abertos e o Movimento de Acesso Livre. São iniciativas que vêm construindo as condições necessárias para permitir o acesso livre à produção científica de forma legítima, alterando não somente o processo de aquisição de informação científica, mas também a sua produção, disseminação e uso. Assim, as revistas científicas passaram a dividir o palco da ciência com novos protagonistas, especialmente os repositórios institucionais e temáticos, que também estão presentes no processo que move o ciclo produtivo da ciência.

É possível observar que a Internet, de um modo geral, oferece não só um novo sistema de publicação científica, mas também novas formas de acesso à ciência, as quais são mais flexíveis e promovem, sobretudo, a colaboração entre pesquisadores, a integração de comunidades dispersas e o compartilhamento de idéias. A introdução e adoção das tecnologias de informação e comunicação, particularmente, foram responsáveis por algumas alterações no ciclo da geração, disseminação e uso da informação científica, afe-

tando tanto o papel desempenhado pelos seus atores quanto a estrutura do sistema da comunicação científica.

Portanto, essa nova configuração trouxe mudanças de caráter estrutural para o fluxo da informação científica, cujas implicações necessitam de maiores reflexões no âmbito da Ciência da Informação.

Este trabalho tem por objetivo contribuir para a compreensão deste novo cenário por meio de algumas reflexões sobre o papel dos repositórios institucionais e temáticos no ciclo da comunicação científica.

A proposta de estudar repositórios digitais à luz da estrutura da comunicação científica está baseada em Subramanyam (1980, p. 398, tradução nossa). A estrutura total da literatura científica é constituída por três níveis hierárquicos de publicações: as publicações primárias, sob as quais novos conceitos são registrados e disseminados; as publicações secundárias, que são derivadas das publicações primárias pelo processo de representação (*surrogation*), reempacotamento e compactação; e as publicações terciárias, que são derivadas de um segundo processo de representação (*surrogation*) da literatura secundária.

Dessa forma, as publicações científicas online podem ser entendidas como fontes primárias; os repositórios temáticos e institucionais, como fontes secundárias; e os serviços associados a eles (provedores de serviços), como fontes terciárias. Para isso, é necessário antes entender o que é um produto (publicações), bem como seus processos de publicação (editoração eletrônica), de comunicação (produção, disseminação e uso) e como estão organizadas as suas estruturas (fontes primárias, secundárias e terciárias).

2 A publicação e a comunicação científica

Alguns autores, como Odlyzko (2000),² criticam os rumos do debate atual sobre a comunicação científica, os quais enfatizam apenas uma parte

■
² Documento eletrônico.

desse sistema e concentram os estudos em apenas um dos modelos de publicação, a revista científica. Kling (2004, p. 593, tradução nossa) também chama a atenção para o uso correto dos termos publicação e comunicação científica: a publicação científica é uma parte formal da comunicação científica.

A publicação, de acordo com Kling (2004, p. 593, tradução nossa), pode ser descrita como um processo de mão única que se baseia na hipótese de que um artigo, ao estar disponível para o público, será lido ou então não atrairá atenção dos leitores, encerrando o processo. Já a comunicação científica é um processo de mão dupla que consiste em comunicadores e conteúdos.

Autores clássicos entendem a comunicação científica como um sistema composto de “[...] elementos (encontros científicos, relatórios técnicos, pré-publicações, artigos de periódicos, livros etc.), os quais cientistas usam para processar (originar, desenvolver, transmitir, transformar etc.) a informação” (GARVEY; GOTTFREDSON, 1976, p. 165, tradução nossa). Nesse sentido, o sistema de comunicação científica envolve práticas da comunicação em que o cientista dedica parte importante do seu tempo a atividades de comunicação, tais como leituras, redação e contatos; e também processos de comunicação, constituídos por um processo formal (escrito) e um processo informal (oral), conforme descreve Le Coadic (1996, p. 33). É a partir da seqüência do processo informal ao formal que a atividade científica é construída, comunicada e incorporada por outros cientistas, integrando o conhecimento privado ao conhecimento público, conforme explica Ziman (1979).

De acordo com Subramanyam (1980, p. 393-394, tradução nossa), o processo de produção, disseminação e uso da informação científica constitui a estrutura da literatura científica, que pode ser compreendida como uma linha do tempo que acompanha o progresso da produção desde a sua geração, como uma fonte primária, até o tratamento dado por serviços secundários e sua eventual interação com outros novos trabalhos e compactação em revisão de literatura. Dessa maneira, os produtos, ou “pacotes bibliográficos”, emanam

de cada atividade científica, podendo formar um ciclo da produção científica.

Na ciência da informação, muitos foram os autores que elaboraram representações desse ciclo de estrutura e fluxo da informação técnico-científica, tais como Redmond, Sinclair e Brown (1972), Garvey (1979), Subramanyam (1980), Vickery (1999), Hurd (2002) e Djenchuraev (2004). Essas representações são esquemas orientados para as revistas científicas impressas, alguns tipos de publicações online e representações mais abrangentes da estrutura da literatura.

As lições que essas representações podem trazer para a compreensão das novas configurações na comunicação científica são as noções de fontes primárias, secundárias e terciárias, presentes especialmente na representação de Subramanyam.

O esquema de Subramanyam se inicia com geração do conhecimento e seu registro como literatura primária. As fontes primárias são aqui entendidas como publicações em si, derivadas das pesquisas comunicadas em vários canais, tais como o artigo de uma revista científica, um trabalho apresentado em congresso, um relatório de pesquisa etc.

De acordo com Subramanyam (1980, p. 396-397, tradução nossa), a literatura secundária surge a partir da literatura primária, que passa pelos seguintes processos: a representação (*surrogation*), exemplificada tradicionalmente por bibliografias, catálogos, índices e resumos e serviços de alerta; o reempacotamento, exemplificado por dicionários, diretórios, tabelas, manuais, anuários e almanaques; e a compactação, exemplificada por revisões de literatura, monografias, livros-texto, tratados e enciclopédias. Em resumo, as fontes secundárias (bibliografias, dicionários, enciclopédias etc.) compreendem a massa de literatura primária dispersa em diversas fontes publicadas no mundo.

Como resultado da proliferação da literatura primária e da diversidade das fontes secundárias, surgiu a necessidade de facilitar a identificação desses tipos de fontes. A literatura terciária exerce essa função a partir de um processo de representação secundária (*secondary surrogation*). Nesse nível terciário da literatura não há reempacotamento ou compactação da informação. Os exemplos clássicos des-

se tipo de fonte são as bibliografias de bibliografias, os guias de literatura e os diretórios de diretórios (SUBRAMANYAM, 1980, p. 397-398, tradução nossa).

Da geração do conhecimento às fontes primárias, secundárias e terciárias, o ciclo se encerra com a disseminação e utilização da informação científica para fomentar então novo conhecimento.

Vale ressaltar que a comunicação científica sempre teve como fator determinante a geração de novo conhecimento, sua disseminação e uso para promover o desenvolvimento da ciência. De acordo com Chin (1999, p. 2), o processo de criação do novo conhecimento garante “[...] o crescimento continuado e a vitalidade de qualquer área”. Portanto, quanto maior o prestígio profissional de uma área, melhor é o tratamento dado ao novo conhecimento quanto ao armazenamento e fácil recuperação. Isto ocorre porque, segundo o autor, “a busca do novo conhecimento [...] é dependente da disseminação e armazenamento dos resultados de pesquisa e das atividades criativas, e pronto acesso ao novo conhecimento” (CHIN, 1999, p. 2, tradução nossa). Por isso, a produção de fontes secundárias e terciárias é fundamental para sustentar a estrutura que move o ciclo da produção científica, e, conseqüentemente, a própria ciência.

Os conceitos e propriedades revistos sobre a publicação e a comunicação científicas e também os processos que compõem a estrutura da literatura em fontes primárias, secundárias e terciárias são assumidos como pressupostos que apoiarão essa análise sobre o novo cenário da comunicação científica. A proposta deste artigo é identificar as fontes primárias, secundárias e terciárias entre as publicações científicas online, repositórios temáticos e institucionais e provedores de serviços, respectivamente.

2.1 As publicações científicas online como literatura primária

A possibilidade de publicar tudo e qualquer coisa na internet entrou em conflito com questões sobre direito legal, preservação digital, segurança e qualidade das informações que concorrem com a instauração das novas possibilida-

des de publicação científica online. Era necessário fomentar uma infra-estrutura básica para sustentar e organizar um modelo de publicação que apresentasse as garantias e os benefícios praticados pelo sistema tradicional baseado em papel.

Neste sentido, muitas iniciativas foram desenvolvidas nos últimos quinze anos, a fim de garantir a evolução das publicações para o meio digital alterando não somente o produto final, mas também seus processos de editoração e comunicação. Dessa forma, os vários tipos de publicações científicas online possuem tecnologias próprias, desenvolvidas pela própria comunidade para apoiar e estimular ampla adoção e uso.

O caso das revistas científicas é um exemplo de que essas iniciativas possibilitaram gerenciar o fluxo editorial online e também empreender a editoração e publicação das revistas online por meio de softwares específicos. Transformações também ocorreram com as conferências e simpósios e suas publicações decorrentes, os anais. Todo o processo de inscrições, submissão de trabalhos, editoração e publicação de anais está integrado por softwares livres. O mesmo foi feito com o fluxo de trabalho das teses e dissertações, que pode ser iniciado com o depósito do projeto, passar pela qualificação, e chegar até a defesa, culminando com o relatório final, de modo que todo o processo possa ser executado online em conjunto com a Secretaria de Pós-Graduação do curso.

Existem vários softwares (proprietários e livres) que podem ser adotados para cada tipo de processo e publicação online. Para exemplificar alguns deles é citado o caso do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), que desenvolveu versões em português para os softwares de três tipos de publicações: revistas, anais de eventos e teses e dissertações. Suas ações têm por finalidade apoiar iniciativas internacionais e promover no país a transição da comunicação científica baseada em papel para o formato digital, descrito a seguir:

- **Controle das submissões e publicação de revistas científicas:** o software *Open Journal System* foi traduzido com o nome de SEER – Sistema

Eletrônico de Editoração de Revistas. Este software foi originalmente desenvolvido pela Universidade de British Columbia, no Canadá.

- **Controle de conferências e simpósios e publicação dos anais:** o IBICT, em parceria com o Centro de Informações Nucleares da Comissão Nacional de Energia Nuclear, investiu na versão brasileira do *Open Conference System*, o Sistema Online de Acompanhamento de Conferências (SOAC), também desenvolvido pela Universidade de British Columbia (KURAMOTO, 2005)³.
- **Controle e publicação de teses e dissertações:** no caso das teses e dissertações, o IBICT desenvolveu o TEDE, Sistema de Publicações de Teses e Dissertações, baseado na Iniciativa do ETD (*Electronic Theses and Dissertation*) da Virginia Tech (KURAMOTO, 2005)⁴.

Esses recursos tecnológicos potencializaram a literatura primária, possibilitando melhores desempenhos dos pesquisadores no processo de identificação, seleção e uso da informação. Devido à natureza das publicações online, a recuperação de dados e do texto completo tornou-se muito mais fácil. De fato, essa particularidade é uma inovação importante, pois, anteriormente, a literatura primária dependia das fontes secundárias, especialmente do processo de representação (*surrogation*), segundo Subramanyam (1980), para ser identificada e localizada. Hoje, as próprias publicações primárias cumprem uma parte deste papel.

2.2 Repositórios temáticos e institucionais como fonte secundária

As publicações científicas online, que adotam os softwares recomendados pela OAI e o Movimento de Acesso Livre, também estão presentes em espaços integrativos online que possibilitam sua identificação, seleção e uso pelos pes-

■
³ Documento eletrônico.

⁴ Documento eletrônico.

quisadores em um só local, como se este fosse uma espécie de portal, dispensando o acesso individual a cada título de revista ou tese, por exemplo. Esses espaços integrativos são na verdade repositórios temáticos ou institucionais que funcionam como fontes secundárias na estrutura da literatura científica.

Um repositório digital é um arquivo digital que reúne uma coleção de documentos digitais. Os repositórios digitais que adotam o modelo OAI, isto é, que adotam o protocolo OAI-PMH (*Open Archive Initiative – Protocol for Metadata Harvesting*), compartilham os mesmos metadados, tornando seus conteúdos interoperáveis entre si. Seus metadados podem ser coletados por sistemas “virtuais” globais (provedores de serviços), que funcionam como fontes terciárias e serão detalhados no item 2.3. Esse esquema permite a navegação de forma integrada por qualquer usuário, conforme é feito em uma base de dados, mas com a vantagem do acesso ao texto completo (SELF-ARCHIVING..., 2006)⁵.

De um modo geral, os termos “repositórios institucionais” ou “temáticos” são adotados para caracterizar os repositórios digitais que reúnem respectivamente a produção científica de uma instituição e de uma área. O Reposcom (<http://reposcom.portcom.intercom.org.br>) e o E-LIS (<http://eprints.rclis.org>) são exemplos de repositórios temático e institucional, respectivamente.

2.2.1 Estrutura do modelo OAI

Para detalhar melhor a relação entre as publicações, os repositórios temáticos e institucionais, e os provedores de serviços, é necessário explicar antes a Iniciativa dos Arquivos Abertos – *Open Archives Initiative* (OAI), um movimento instituído em 1999 por um grupo de pesquisadores europeus e norte-americanos. Essa iniciativa foi pioneira na reflexão sobre o processo de publicação e reorganização das publicações científicas, bem como de suas influências na comunicação científica eletrônica.

■
⁵ Documento eletrônico.

A OAI foi responsável pela constituição das bases necessárias para implementar um novo modelo de acesso à publicação científica a partir da integração de soluções tecnológicas e também pela busca de mecanismos para garantir sua legitimação (VAN DE SOMPEL; LAGOZE, 2000).

Para isso foi desenvolvido um conjunto de padrões técnicos e tecnológicos (incluindo a OAI-PMH, padrão de metadados etc.) que sustentou uma infraestrutura para publicação na *web* a partir de modelos de próprios, conforme o de teses, revistas etc. que, como vistos anteriormente, são implementados com softwares específicos e no seu agrupamento em repositórios digitais. Segundo a OAI, esses repositórios digitais funcionam como provedores de dados, possibilitando a coleta integrada dos metadados pelas ferramentas de busca comuns, tais como o *Google* e o *Google Scholar*, e também por provedores de serviços próprios segundo os moldes da OAI, tais como a *OAIster*. Essa estrutura aumenta a oportunidade de outros pesquisadores conhecerem novos trabalhos e promoverem o desenvolvimento do ciclo produtivo da ciência. (OAISTER...)⁶

Em resumo, a OAI fomentou a infra-estrutura necessária para apoiar publicações científicas online, repositórios temáticos e institucionais e também provedores de serviços que possibilitam a integração total de buscas e textos completos, tais como as fontes primárias, secundárias e terciárias, respectivamente.

2.2.2 Softwares e características dos repositórios temáticos e institucionais

Tanto as publicações científicas online como os repositórios temáticos e institucionais também necessitam de softwares próprios para serem implementados. Dentre os vários existentes, atualmente destacam-se os

■
⁶ Documento eletrônico.

softwares Eprints e DSPACE. O IBICT também desenvolveu versões em português para ambos. Na versão brasileira, o Eprints, desenvolvido originalmente pela University of Southampton, teve seu nome alterado para DICI – Diálogo Científico. Já o software DSPACE, desenvolvido pelo MIT (*Massachusetts Institute of Technology*) e pela HP (*Hewlett-Packard*), não teve seu nome alterado. Tanto o Eprints quanto o DSPACE foram traduzidos pelo IBICT em parceria com a equipe da PORTCOM (Rede de Informação em Comunicação dos Países de Língua Portuguesa) da INTERCOM (Sociedade Brasileira de Estudos Interdisciplinares da Comunicação) e do Núcleo de Pesquisa *Design de Sistemas Virtuais Centrado no Usuário* da USP (Universidade de São Paulo).

Os repositórios temáticos e institucionais apresentam algumas características comuns: são auto-sustentáveis, baseados sobretudo no auto-arquivamento da produção científica (que compreende a descrição padronizada dos metadados e o *upload* do arquivo em PDF ou outro formato de texto) e fornecem interoperabilidade entre os diferentes sistemas e o acesso livre para todos os interessados em pesquisar e baixar arquivos da produção científica. Em outras palavras, é possível depositar artigos já publicados ou quaisquer outras publicações online em repositórios digitais, a fim de possibilitar o acesso aos textos completos dos trabalhos já publicados bem como os seus dados descritivos (metadados) de autoria, título, palavras-chave etc. A interoperabilidade é um fator primordial, pois possibilita a busca em um só local e a reunião de conteúdos dos diferentes tipos de publicações online produzidas por diferentes softwares sem conflitos.

Os repositórios temáticos ou institucionais, de um modo geral, não substituem as publicações genuínas, tais como teses e dissertações, revistas científicas, anais de eventos, etc. Em outras palavras, os repositórios digitais não são publicações, são como bibliografias especializadas, ou melhor, são serviços de indexação e resumo constituídos pelas próprias comunidades científicas. Sua

função precípua é permitir o acesso organizado e livre às publicações e a toda a produção científica.

Isto é feito de forma descentralizada e dependente da iniciativa de cada autor. Por isso, é fundamental que sejam instituídas políticas para cada repositório, a fim de incentivar o maior número de depósitos da produção científica para estimular a disseminação e uso desta produção, mas também como forma de espelhar a memória institucional ou de uma área.

Na verdade, o contexto do surgimento dos repositórios digitais está relacionado ao combate da lógica do acesso pago em que editores se beneficiam do produto do trabalho de pesquisadores, que é cedido gratuitamente. São várias as iniciativas para promover o acesso livre à produção científica, mas todas desejam isso de forma organizada e legítima. Dessa forma, em vez de simplesmente “postar” seus trabalhos em um site pessoal, o pesquisador pode fazê-lo nos repositórios digitais. Isto também não quer dizer que os autores devam desistir de publicar. As publicações científicas online têm seu papel de relevância na comunicação científica, mas podem e devem ser também depositadas em repositórios digitais, a fim de dar continuidade ao ciclo da produção científica de modo a promover a identificação, seleção e uso da informação. Ou seja, os repositórios digitais, de um modo geral, estão exercendo o mesmo papel cumprido pelos serviços de indexação e resumo na atualidade, mas com a diferença de permitir o acesso livre.

2.3 Provedores de serviços como fontes terciárias

Conforme visto anteriormente, os provedores de serviços agrupam conteúdos de vários repositórios digitais, facilitando a busca e otimizando o acesso ao texto completo. De acordo com Ferreira e Souto (2006), essas facilidades existem por causa do sistema de coleta de metadados/haverting, estrutura originária da OAI, já mencionada:

O “harvester ou “agregador” visita cada um dos repositórios parceiros e autônomos coletando os metadados disponíveis, abastecendo um repositório global com eles e oferecendo uma interface única de busca. A partir daí, os usuários são direcionados diretamente ao registro original e/ou documento completo localizado nas diversas coleções locais dos parceiros. (FERREIRA; SOUTO, 2006)

Os provedores de serviços, assim como os provedores de dados, também precisam de softwares para sua implementação e organização. Ferreira e Souto (2006) apresentam três exemplos de softwares, bem como os respectivos provedores de serviços:

- *ARC* – A Cross Archive Search Service: projeto desenvolvido pela Old Dominion University Digital Library Research Group (ODU) que utiliza softwares livres e o modelo OAI.
- *Metalis*: provedor de serviço para a área de Biblioteconomia e Ciência da Informação baseado em software livre e protocolo OAI.
- *OAIster*: projeto da University of Michigan Digital Library Production Service que faz a coleta de metadados de repositórios temáticos ou institucionais que adotam o modelo OAI.

Os provedores de serviços como fontes terciárias reúnem tanto as publicações científicas online (fontes primárias) quanto os repositórios institucionais e temáticos (fontes secundárias). No meio digital isto é possível por causa da infra-estrutura fomentada pela OAI, que é composta por provedores de dados e de serviços. Neste caso, tanto as publicações científicas online quanto os repositórios digitais são provedores de dados.

3 Algumas implicações na comunicação científica

Visto que a publicação científica online pode estar presente em repositórios digitais e em provedores de serviços, é possível afirmar que o contexto da comunicação científica eletrônica manteve a estrutura da literatura científica.

Por outro lado, a infra-estrutura tecnológica proveniente da OAI possibilitou o acesso livre e também uma nova forma de lidar com a comunicação

científica, influenciando, sobretudo, o modo como os cientistas publicam, disseminam e utilizam os resultados de suas pesquisas.

Dessa maneira, a Internet potencializou o sistema de comunicação científica estruturado pelas relações sociais sem hierarquias, caracterizando-se pela autogestão, simultaneidade do processo, heterogeneidade e colaboração. Assim, o autor/pesquisador interage diretamente no processo de produção científica e no fluxo das informações sem intermediários. Também é possível incorporar outras atividades no processo de comunicação científica eletrônica, tal como a comunicação informal, o que não era possível no sistema baseado em publicações impressas (BARRETO, 1998).

Todas essas mudanças parecem anunciar a reestruturação do sistema de comunicação científica, a qual valoriza o tríplice papel do pesquisador enquanto produtor-disseminador-consumidor. Por outro lado, o autor assume uma parte do papel do bibliotecário ao fazer a descrição dos metadados e o auto-arquivamento de sua produção, e, conseqüentemente, isto fez com que o editor perdesse a exclusividade de distribuição da produção científica no contexto digital. O bibliotecário, por sua vez, encontrou parceiros no processo de tratamento e controle da produção científica. São implicações importantes a serem consideradas no âmbito da Biblioteconomia e da Ciência da Informação.

A OCLC (2004)⁷ procurou descrever o atual processo de comunicação científica eletrônica identificando os papéis de antigos atores presentes no fluxo de produtos científicos. Nessa representação, os elementos do fluxo são compostos por:

- a) Revistas científicas;
- b) Repositórios: de caráter institucional, pessoal ou de comunidades que podem ser especializadas ou gerais;

■
⁷ Documento eletrônico.

- c) Provedores de Serviços ou Serviços Agregadores: constituídos de serviços comerciais ou oferecidos pela comunidade, como o OAIster ou financiados por agências de fomento.

Em resumo, os pesquisadores, docentes e discentes em processo de produção e descoberta, em suas atividades de ensino e/ou pesquisa, utilizam os provedores de serviços, repositórios digitais e revistas, entre outras publicações primárias, para cobrir a literatura. Publicam os resultados das pesquisas em revistas, entre outros e por meio do auto-arquivamento em repositórios digitais, cujos metadados são coletados pelos provedores de serviços, e assim é reiniciado o ciclo.

4 Tendências futuras

É preciso esclarecer que a proposta da OAI para a produção científica tem por objetivo eliminar um gargalo no processo da comunicação científica que desencadeou a chamada crise das revistas científicas – prejudicando a acessibilidade à produção científica. Para isso, a OAI focou suas ações para promover o acesso livre como condição fundamental para o desenvolvimento científico. Dentre as principais ações, foram destacados a infra-estrutura tecnológica e o combate da lógica do acesso pago, conforme visto no item 2.2, desencadeando as mudanças mencionadas no fluxo da comunicação científica.

No entanto, apesar das propostas de mudança da OAI terem sido implementadas ao longo da década de 1990, somente uma pequena parcela da comunidade científica está presente nestes repositórios digitais. Estima-se que apenas 15% da produção científica no mundo estejam em repositórios digitais para acesso livre, sem incluir as revistas científicas de acesso livre (HAJJEM; HARNAD; GINGRAS, 2005, [p. 8])⁸.

■
⁸ Documento eletrônico.

No mundo todo ocorrem movimentos da comunidade científica para que as novas regras do jogo do sistema de publicações sejam aceitas no sistema de comunicação científica, sobretudo em órgãos de fomento e de avaliação institucional. Os repositórios institucionais foram considerados como uma estratégia de ação para aumentar esses índices. *A Declaração de Berlim sobre o Acesso Livre ao Conhecimento nas Ciências e Humanidades* apóia a implementação dos repositórios institucionais e defende o auto-arquivamento de todos os trabalhos de pesquisa financiados pelo governo como atitude política de todos os pesquisadores do mundo. (DECLARAÇÃO...)⁹

É uma estratégia importante para países que já avançaram neste debate e implementaram um número relevante de repositórios institucionais e temáticos. No caso do Brasil, a partir da década de 2000 já se tem notícia das primeiras iniciativas, e o País figura entre os cinco primeiros no ranking de número de itens depositados em repositórios digitais (KURAMOTO, 2005).

No entanto, a comunidade científica ainda não está suficientemente esclarecida de suas vantagens, sua função e da importância do seu engajamento para que essas iniciativas ocupem seu espaço. Ainda são muitos os desafios que emergem nesse cenário tão mutante. Alguns se referem às competências e habilidades necessárias aos pesquisadores para atuar neste novo fluxo de informação científica como agentes que produzem, disseminam e utilizam a produção científica sem intermediários, o que é um grande desafio para as diferentes culturas existentes no mundo. Há ainda o mito de que o autor não tem garantido os direitos autorais. Seus defensores acreditam que os repositórios digitais zelam pelos direitos autorais e que, quanto mais visível estiver o trabalho de um autor, menos chance existe de ele ter seus direitos lesados. Dessa forma, o debate deve ter o foco sobre os direitos de cópia, e não do autor. No modelo praticado pelas revistas científicas, os direitos do autor são cedidos ao editor como uma contrapartida pela publicação dos trabalhos.

Outros desafios que também são relevantes se referem às características

de cada área do conhecimento e de sua comunidade científica. Certamente que as comunidades mais engajadas e conscientes sobre a importância da produção científica (ou aquelas que possuem condições de levantar mais verbas) serão mais rápidas em buscar soluções e estratégias para implementação de repositórios institucionais ou temáticos. Essas soluções estão relacionadas tanto com a questão do financiamento quanto de implementação e gerenciamento de repositórios digitais, buscando parcerias com o Estado ou iniciativa privada, conforme as alianças natas de cada área. Ou seja, é o modo como cada área lida com suas pesquisas que definirá o papel de cada um nesse processo.

5 Considerações finais

As breves reflexões apresentadas tiveram como objetivo esclarecer que as publicações científicas online, os repositórios temáticos e institucionais e os provedores de serviços são produtos diferentes, com características próprias e funções distintas. Sob o enfoque da estrutura da literatura científica, essas três esferas do ciclo da comunicação científica podem ser entendidas respectivamente como fontes primárias, secundárias e terciárias. Portanto, no contexto digital atual, os meios para registrar, disseminar e usar a informação científica são efetuados pelas publicações online, repositórios temáticos e institucionais e provedores de serviços.

No entanto, a consolidação das iniciativas que apóiam a produção, disseminação e o acesso livre às publicações científicas online, especialmente no Brasil, ainda prescinde do engajamento das comunidades científicas para instituir esses modelos no âmbito da produção científica.

São questões de importância visceral para a promoção do desenvolvimento científico neste século, pois, conforme foi visto, o prestígio de uma área está diretamente relacionado ao tratamento dado ao novo conhecimento: isto é, quanto ao seu armazenamento e recuperação.

As bases anteriores que apoiaram o desenvolvimento científico até então

não apresentam mais condições de lidar com as características próprias de modelos abertos, compartilhados, flexíveis e autônomos.

Trata-se de um momento de decisão e de extrema relevância para o futuro da ciência no mundo. Caberá às instituições científicas (ou seja, ao Estado, à iniciativa privada e à comunidade científica – representada por institutos de pesquisa e universidades) debater e propor soluções e estratégias de ação para alcançar este novo estágio de desenvolvimento científico.

The role of institutional and discipline repositories in the structure of scientific production

ABSTRACT

The aim of this article is to contribute for the understanding of the new landscape of the scientific communication in the 21st century through some considerations about the role of institutional and discipline repositories in this context, based on Subramanyam's (1980) statements on the structure of scientific literature. The starting point of this approach focuses the characteristics of products (publications, digital repositories and services providers), processes (electronic publishing and communication) and their counterparts in the structure of scientific literature. It concludes that the cycle of online scientific production is constituted by primary sources (online scientific publications), secondary sources (institutional and discipline repositories) and tertiary sources (services providers) and that the implementation and use of these three types of sources are imperative to promote scientific development in present time.

KEYWORDS: Institutional repositories. Discipline repositories. Electronic scientific communication. Open Files Initiative. Free Access Movement.

El papel de los repositorios institucionales y temáticos en la estructura de la producción científica

RESUMEN

La propuesta de este artículo es contribuir para la comprensión del nuevo escenario de la comunicación científica en el siglo XXI por medio de algunas reflexiones sobre el papel de los repositorios institucionales y temáticos en este contexto, a partir de las nociones sobre la cultura de la literatura científica de Subramanyam (1980). El punto de partida de este abordaje está focado en las características de los productos (publicaciones, repositorios digitales y proveedores de servicios), procesos (editoración electrónica y comunicación) y sus correspondencias en la estructura de la literatura científica. Se concluye que el ciclo de la producción científica on line se compone por las fuentes primarias (publicaciones científicas on line), secundarias (repositorios temáticos e institucionales) y terciarias (proveedores de servicios) y que la implementación y uso de estos tres tipos de fuentes son urgentes para promover el desarrollo científico en la actualidad.

PALABRAS-CLAVE: Repositorios institucionales. Repositorios temáticos. Comunicación científica electrónica. Iniciativa de Archivos Abiertos. Movimiento de Acceso Libre.

Referências

BARRETO, Aldo de Albuquerque. Mudança estrutural no fluxo do conhecimento: a comunicação eletrônica. *Ciência da Informação*, Brasília, v. 27, n. 2, p. 122-127, maio/ago. 1998.

CHIN, Robert A. Disseminating, archiving, and retrieving new knowledge in industrial technology: implications for the discipline and NAIT. *Journal of Industrial Technology*, v. 15, n. 2, Feb./Apr. 1999. Disponível em: <<http://www.nait.org/jit/current.html>>.

DECLARAÇÃO de Berlim sobre o acesso livre ao conhecimento nas ciências e humanidades. Disponível em: <<https://repositorium.sdum.uminho.pt/about/DeclaracaoBerlim.htm>>. Acesso em: 25 maio 2006.

DJENCHURAEV, Nurlan. **Toward a new policy for scientific and technical communication: the case of Kyrgyz Republic.** 2004. Disponível em: <http://www.policy.hu/djenchuraev/frp.html#_Toc72643094>. Acesso em: 1 fev. 2006.

FERREIRA, Sueli Mara S. P. Ferreira; SOUTO, Leonardo Fernandes. Dos sistemas de informação federados à federação de bibliotecas digitais. *Revista Brasileira de Biblioteconomia e Documentação Nova Série*, São Paulo, v. 2, n. 1, p. 23-40, jan./jun. 2006.

GARVEY, W. D. **Communication: the essence of science.** Elmsford, NY: Pergamon Press, 1979.

GARVEY, W. D.; GOTTFREDSON, S. D. Changing the system: innovations in the interactive social system of scientific communication. **Information Processing and Management**, New York, v. 12, n. 3, p. 165-176, 1976.

GOOGLE. Disponível em: <<http://www.google.com.br>>. Acesso em: 11 abr. 2006.

GOOGLE Scholar. Disponível em: <<http://scholar.google.com>>. Acesso em: 11 abr. 2006.

HAJJEM, C.; HARNAD, S.; GINGRAS, Y. Ten-Year Cross-Disciplinary Comparison of the Growth of Open Access and how it Increases Research Citation Impact. **IEEE Data Engineering Bulletin**, v. 28, n. 4, p. 39-47, 2005. Disponível em: <<http://eprints.ecs.soton.ac.uk/11688/01/ArticleIEEE.pdf>>. Acesso em: 11 abr. 2006.

HURD, Julie M. **Information technology**: catalyst for change in scientific communication. Disponível em: <<http://www.iatul.org/conference/proceedings/vol06/papers/full/hurd.html>>. Acesso em: 1 fev. 2006.

KLING, Rob. The internet and unrefereed scholarly publishing. In: CRONIN, Blaise (Ed.). **Annual Review of Information Science and Technology**, Medford, NJ, v. 38, p. 591-631, 2004.

KURAMOTO, Helio. Os *open archives* e as políticas públicas para a informação científica. In: SIMPÓSIO INTERNACIONAL DE BIBLIOTECAS DIGITAIS, 3., 2005, São Paulo. [Anais eletrônicos...]. São Paulo: CRUESP, 2005. Disponível em: <http://bibliotecas-cruesp.usp.br/bibliotecas/APRESENT/Helio_Kuramoto.ppt>. Acesso em: 25 maio 2006.

LE COADIC, Yves-François. **A Ciência da Informação**. Brasília: Briquet de Lemos/Livros, 1996. 119 p.

OAISTER. Disponível em: <<http://oaister.umdl.umich.edu/o/oaister/>>. Acesso em: 11 abr. 2006.

OCLC ONLINE COMPUTER LIBRARY CENTER. **Análise do cenário da OCLC em 2003**: reconhecimento de padrões: resumo executivo do reporte *The 2003 Environmental Scan: pattern recognition*. Dublin, c2004. 17 p. Disponível em: <www.oclc.org/reports/escan/downloads/escansummary_po.pdf>. Acesso em 1 fev. 2006.

ODLYZKO, Andrew. The future of scientific communication. In: WOUTERS, P.; SCHROEDER, P. (Eds.). **Access to publicly financed research: the global research village III**. Amsterdam: NIWI, 2000. p. 273-278. Disponível em: <<http://www.dtc.umn.edu/~odlyzko/doc/future.scientific.comm.pdf>>. Acesso em: 10 fev. 2006.

REDMOND, Donald A.; SINCLAIR, Michael P.; BROWN, Elinore. University libraries and university research. **College & Research Libraries**, Chicago, Ill., v. 33, n. 6, p. 447- 453, Nov. 1972.

SELF-ARCHIVING FAQ. Disponível em: <<http://www.eprints.org/openaccess/self-faq/>>. Acesso em 11 de abr. 2006.

SUBRAMANYAM, K. Scientific Literature. In: KENT, A.; LANCOUR, H.; DAILY, J. E. (Eds.). **The Encyclopedia of Library and Information Science**. New York: M. Dekker, 1980. v. 26.

VAN DE SOMPEL, Herbert; LAGOZE, Carl. The Santa Fe Convention of the

Open Archives Initiative. **D-Lib Magazine**, Reston, Virg., v. 6, n. 3, Feb. 2000.
Disponível em: <<http://www.dlib.org/dlib/february00/vandesompel-oai/02vandesompel-oai.html>>. Acesso em: 5 feb. 2002.

VICKERY, Brian. A century of scientific and technical information. **Journal of Documentation**, London, v. 55, n. 5, p. 476-527, Dec. 1999.

ZIMAN, John. **Conhecimento público**. Belo Horizonte: Itatiaia, 1979.

Simone da Rocha Weitzel

*Professora Adjunta do Departamento de Estudos e
Processos Biblioteconômicos do
Centro de Ciências Humanas e Sociais da UniRio
Doutoranda do Programa de Pós-Graduação em
Ciência da Informação da ECA/USP
E-mail: simone.weitzel@gmail.com*