



Reflexões sobre os repositórios institucionais¹

Simone da Rocha Weitzel²

Universidade Federal do Estado do Rio de Janeiro (UNIRIO)

Universidade de São Paulo (USP)

Resumo

Apresenta reflexões sobre temas que emergem do contexto da comunicação científica no século XXI. Um deles se refere à distinção entre publicações eletrônicas, enquanto produção científica, e as novas formas de organização desta produção em repositórios institucionais ou temáticos. Outro tema está relacionado com a infra-estrutura fomentada pela OAI (Open Archives Initiative) que permitiu a re-organização da produção científica online e se tornou fundamental para o ciclo da comunicação científica no contexto digital. Conclui-se que esta infra-estrutura, que permite a publicação online, a constituição de repositórios e os provedores de serviços, é vital para todas as áreas do conhecimento uma vez que se tornou um recurso imprescindível para promover o desenvolvimento científico.

Palavras-chave

Repositórios institucionais; comunicação científica eletrônica; Iniciativa de Arquivos Abertos.

Introdução

Autores como Ziman, (1979), Meadows (1999) e Miranda (2003) consideram a produção científica como um recurso secular e imprescindível para a promoção do desenvolvimento científico. A publicação de artigos em revistas científicas tem sido o meio mais tradicional para alimentar este ciclo produtivo bem como a conseqüente citação em novos trabalhos. Este processo obedece a critérios igualmente seculares sustentado pelos pilares da comunicação científica – acessibilidade, fidedignidade e disseminação (KLING, 2004).

Em outras palavras, o artigo científico deve estar acessível por um longo período de tempo de modo que outros pesquisadores tenham oportunidade de tomar conhecimento da obra (acessibilidade). **O conteúdo do texto deve também satisfazer as normas de qualidade por meio da prática da revisão de pares** (fidedignidade); e,

¹ Trabalho apresentado no XVI Endocom – Encontro de Informação em Ciências da Comunicação

² Professora Assistente lotada no Departamento de Estudos e Processos Biblioteconômicos do Centro de Ciências Humanas e Sociais da UNIRIO, onde leciona disciplinas para a Escola de Biblioteconomia; e doutoranda do Programa de Pós-Graduação em Ciência da Informação da Escola de Comunicações e Artes da USP, onde integra o Núcleo de Pesquisa “Design de Sistemas Virtuais Centrado no Usuário” coordenado pela Prof.^a Dr.^a Sueli Mara Ferreira do Departamento de Biblioteconomia e Documentação da ECA/USP.



por fim, obra deve estar disponível para um amplo público para alcançar seu objetivo de fomentar novos conhecimentos a partir deste (disseminação).

Este modelo de publicação foi praticado ao longo de três séculos e fomentou todo um sistema de comunicação científica no mundo. No entanto, no último quartel do século XX, este modelo chegou ao um ponto crítico denominado de “crise das revistas científicas” afetando sobremaneira dois dos três pilares da comunicação científica: a disseminação e a acessibilidade. A crise, na verdade, se refere ao fato de que as bibliotecas não conseguiam mais acompanhar os altos preços das assinaturas dos melhores títulos de revistas científicas comprometendo tanto o acesso da comunidade científica à produção científica que ela própria fomentara quanto o próprio desenvolvimento da ciência. Esta lógica da produção científica foi muito questionada pela comunidade especialmente porque muitos acreditam que o modelo de publicação científica se tornou um negócio mais lucrativo para as editoras que para a comunidade científica, uma vez que os pesquisadores são dependentes do acesso à produção científica para promoção do desenvolvimento científico ou para cumprir os requisitos da atividade científica. A dispersão da produção e a explosão informacional também foram fatores agravantes na eficiência do processo de comunicação científica.

A adoção das tecnologias da informação e comunicação pela Ciência revolucionou as possibilidades de comunicação entre cientistas alterando vários aspectos do fluxo da comunicação científica descrito acima. O início do século XXI foi marcado pela re-organização dos processos e produtos da comunicação científica por meio de iniciativas de cientistas que vêm construindo as condições para permitir o acesso livre à sua própria produção de forma legítima a fim de superar o gargalo de todo o processo.

Assim, as revistas científicas passaram a dividir o palco da ciência com novos protagonistas que também são responsáveis pelo processo que move o ciclo produtivo da ciência. Portanto, é possível observar que a internet oferece mais que um novo sistema de publicação científica, mas também novas formas de organização da produção científica que são mais flexíveis e promovem, sobretudo, a colaboração entre cientistas, integração de comunidades científicas dispersas e o compartilhamento de idéias.

No entanto, é comum verificar atribuições inadequadas aos mais recentes processos e produtos da comunicação científica prejudicando a compreensão de ações e políticas empreendidas pelo Estado e pela própria comunidade científica voltadas especialmente para promover o acesso livre à produção científica.



Este trabalho tem por objetivo contribuir para a compreensão deste novo cenário por meio de algumas reflexões focadas nas diferenciações entre produtos (publicações, repositórios institucionais e temáticos e serviços associados a eles, denominados de provedores de serviços) e processos (editoração eletrônica e comunicação).

Como ponto de partida para a compreensão desta perspectiva da comunicação científica eletrônica dois pontos de vista são apresentados – um mais restrito, voltado para produtos e processos, e outro mais amplo, voltado para as novas possibilidades de organização destes produtos – que são respectivamente: a) um novo sistema para publicações científicas eletrônicas; e b) um novo sistema para acesso integrado às publicações científicas.

Um novo sistema de publicações científicas eletrônicas

A publicação eletrônica certamente possui inúmeras vantagens sobre qualquer publicação impressa: seja pelos recursos de multimídia, de pesquisa e *links* para referências e citações, seja pelos serviços, fóruns, debates, contato com autor, entre outros recursos possíveis. A outra grande vantagem está na possibilidade de qualquer um “publicar” qualquer coisa na internet. Para o sistema de comunicação científica tradicional essa possibilidade entrou em conflito com questões sobre direito legal, preservação digital, segurança e qualidade das informações veiculadas na internet concorrendo com a instauração das novas possibilidades de publicação eletrônica científica. Naquela ocasião a ausência de infra-estrutura básica para sustentar e organizar um modelo de publicação assim tão audacioso também reforçou uma descrença sobre as reais condições da internet em abrigar um sistema de comunicação científica secular, isto é, que apresentasse as garantias e os benefícios do sistema tradicional, em especial aquele baseado em papel.

Neste sentido, muitas iniciativas foram desenvolvidas nestes 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 produção, disseminação e uso em geral, e processo de editoração eletrônica em particular. Toda essa infra-estrutura tecnológica possibilitou o acesso livre e o controle da produção científica. Em outras palavras, tornou mais eficiente a acessibilidade e a disseminação no sistema da comunicação científica.



Como exemplo desta iniciativa podemos citar o caso das revistas científicas que possibilitou não somente gerenciar o fluxo editorial online, mas também publicar revistas de acesso livre por meio do software *Open Journal System*, traduzido no Brasil pelo IBICT com o nome de SEER – Sistema Eletrônico de Editoração de Revistas. Este software foi originalmente desenvolvido pela Universidade de British Columbia, no Canadá.

Estas transformações também ocorreram com as conferências e simpósios e suas publicações decorrentes, os anais. Todo o processo de inscrições, submissão de trabalhos e publicação de anais estão integrados por software livres. 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.

O mesmo foi feito com o fluxo de trabalho das teses e dissertações que podem iniciar com o depósito do projeto, passar pela qualificação e chegar até a defesa culminando com o relatório final, isto é, todo o processo é executado online em conjunto com a Secretaria de Pós-Graduação. No caso das teses, 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).

Neste sentido, muitas iniciativas foram desenvolvidas nestes 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.

No caso brasileiro, o IBICT tem apoiado fortemente as ações para divulgação e uso desses recursos para promover sua ampla adoção pela comunidade científica. É necessário esclarecer que existem outros softwares para cada tipo de produção e publicação online, mas foram mencionados somente aqueles adotados pelo governo brasileiro, representado aqui pelo IBICT.

Certamente que pela natureza das publicações online a recuperação de dados é muito mais facilitada. Mas, os recursos possibilitados pela rede potencializam a busca integrada, quase universal, dispensando o acesso individual a cada título de revista, por exemplo, para uma espécie de portal universal. Por um lado, esta facilidade é fruto da



infra-estrutura da OAI, na qual as publicações eletrônicas se tornaram provedores de dados. Por outro lado, essas publicações também podem ser agrupadas sob a forma de repositórios digitais integrando vários tipos de publicações em um só local. Os repositórios digitais são o tema do próximo tópico.

Um novo sistema para acesso integrado às publicações científicas

As publicações científicas eletrônicas podem ser categorizadas como revistas científicas, anais de eventos, relatórios de pesquisa etc conforme foi visto acima. Mas, também podem estar presentes e disponíveis em outros espaços, ou veículos além do seu próprio. Ou seja, existe um outro nível de organização da produção científica que possibilita o acesso integrado às publicações científicas, independente de serem artigos já publicados ou trabalhos apresentados em eventos, etc. tal como ocorre em uma biblioteca, ou bibliografia. Estes espaços são denominados de repositórios digitais.

Um repositório digital é um arquivo digital que reúne uma coleção de documentos digitais. Os repositórios 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, conforme foi visto acima, podem ser coletados por sistemas “virtuais” globais (provedores de serviços) que 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 (EPRINTS, 2006).

Os repositórios digitais podem ser categorizados como institucionais ou temáticos. Os institucionais se referem à organização e acesso à produção científica de uma instituição, e os temáticos se referem a uma área.

Existe um número relevante de softwares para implantar repositórios digitais dentre os quais destacamos o software DSPACE e Eprints. O IBICT também desenvolveu versões em português para ambos. DSPACE, desenvolvido pelo MIT (Massachusetts Institute of Technology) e pela HP (Hewlett-Packard), foi traduzido 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). O IBICT também



desenvolveu versões em português para ambos. O Eprints teve seu nome alterado para DICI – Diálogo Científico na versão brasileira e foi originalmente desenvolvido pela University of Southampton.

O Reposcom (<http://reposcom.portcom.intercom.org.br>) é um exemplo deste tipo de organização da produção científica. Trata-se de um repositório institucional que reúne a coleção da produção científica entre livros, capítulos de livros, trabalhos apresentados em eventos, teses e dissertações, etc., produzidos por associações, núcleos de pesquisas e programas de pós-graduação em ciências da comunicação dos países de língua portuguesa. O Reposcom é mantido pela PORTCOM e utiliza o software DSPACE.

O E-LIS (<http://eprints.rclis.org>), é um exemplo de repositório temático e reúne a produção da área de Biblioteconomia, Ciência da Informação e afins entre artigos de revistas, trabalhos apresentados em eventos, teses e dissertações, etc. O E-LIS é mantido pelo trabalho voluntário de bibliotecários do mundo todo e utiliza o software Eprints.

Para detalhar melhor como essas publicações podem ser agrupadas em outro nível de organização é necessário explicar antes a Iniciativa dos Arquivos Abertos – OAI (Open Archives Initiative), um movimento instituído em 1999 por um grupo de pesquisadores europeus e norte-americanos. Esta iniciativa foi pioneira na reflexão sobre o processo de re-organização das publicações científicas e suas influências na comunicação científica eletrônica. 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). Foi desenvolvido um conjunto de padrões técnicos e tecnológicos tanto para viabilizar e estimular a publicação na *web* a partir de modelos de negócios próprios³, conforme visto acima, quanto para organizar e implementar repositórios digitais. Estes repositórios são auto-sustentáveis baseados, sobretudo, no auto-arquivamento da produção científica (que compreende a descrição padronizada dos metadados⁴ e o *upload* do arquivo correspondente no sistema), interoperabilidade entre os diferentes

³ Conforme foi visto existem outros modelos próprios de publicação online, tais como o de teses, revistas, etc. que são implementados com softwares específicos. Mas, todos podem integrar-se a partir da infra-estrutura da OAI.

⁴ Metadados são dados que descrevem uma informação, um item. Exemplo: autor, título, data de publicação, etc.



sistemas de publicação e o acesso livre para todos os interessados em pesquisar e baixar arquivos da produção científica.

Esta infra-estrutura permite tanto a publicação digital de itens individuais (revistas, teses, etc) quanto a constituição de arquivos digitais (os repositórios digitais) contendo preferencialmente os textos completos dos trabalhos já publicados bem como os seus dados descritivos (metadados). Os provedores de serviços complementam esta estrutura, pois agrupam dados dos repositórios funcionando como se fosse uma bibliografia de bibliografia, com a diferença de permitir o acesso ao texto completo.

Estes repositórios, funcionam como provedores de dados possibilitando a coleta integrada dos metadados pelas ferramentas de busca de ampla audiência, tal como o Google e o Google Scholar e provedores de serviços do porte da OAIster, aumentando a oportunidade de outros pesquisadores de conhecer novos trabalhos e promover o desenvolvimento do ciclo produtivo da ciência.

Os repositórios digitais de um modo geral não substituem as publicações genuínas, tais como teses e dissertações, revistas científicas, anais de congressos, etc. Em outras palavras, os repositórios digitais não são publicações, são como se fossem bibliotecas digitais ou bibliografias constituídas pela própria comunidade. Sua função precípua é permitir o acesso organizado e livre àquelas publicações em especial e a toda a produção científica no geral por meio dos repositórios digitais. Isto é feito de forma descentralizada e depende da iniciativa de cada autor. No entanto, em repositórios institucionais este processo deve ser instituído por meio de políticas institucionais, incentivando o maior número de depósitos da produção científica não somente para estimular a disseminação e uso desta produção, mas também como forma de espelhar a memória institucional.

É fato que apenas este foco da questão foi responsável pelo desencadeamento de mudanças no processo de produção científica propriamente dito redefinindo os papéis do autor (que também assume uma parte do papel de editor) e do editor (que perde a exclusividade de distribuição da produção científica no contexto digital). O debate central que está ocorrendo sob o ponto de vista do sistema de publicações eletrônicas na atualidade está em combater a lógica do acesso pago onde editores se beneficiam do produto do trabalho de pesquisadores e cientistas, sobretudo das revistas científicas e permitir o acesso de forma controlada, organizada e legítima.

Portanto, a infra-estrutura estabelecida pela OAI viabilizou não somente um sistema de publicações eletrônicas (produtos e processos), mas também uma outra



forma de organizar a produção científica por meio dos repositórios digitais, um nível secundário de organização da produção científica⁵, tal como uma bibliografia tradicional, mas que têm a função de promover o acesso livre à produção científica, isto é, prover acesso ao texto completo. Este recurso procura resolver um gargalo no processo da comunicação científica que desencadeou a chamada crise das revistas científicas ao focar suas estratégias em um dos pilares da comunicação científica: a acessibilidade como condição fundamental para o desenvolvimento científico. Também proporciona eficiente controle bibliográfico da produção científica em larga escala, minimizando os problemas de dispersão da informação científica.

Implicações para o sistema de comunicação científica eletrônica

Visto os dois pontos de vistas deste novo contexto da comunicação científica eletrônica em relação às publicações eletrônicas propriamente ditas e das novas formas de organização destas publicações eletrônicas em repositórios e em provedores de serviços é possível entender que esta infra-estrutura tecnológica proveniente da OAI possibilitou não somente o acesso livre, mas também uma nova forma de lidar com a comunicação científica influenciando, sobretudo, o modo como os cientistas publicam os resultados de suas pesquisas.

Alguns autores como Odlyzko (2000) criticam os rumos do debate atual na área da Ciência da Informação sobre a comunicação científica que enfatiza apenas uma parte do sistema de comunicação concentrando os estudos em apenas um dos modelos de publicação, a revista científica. O fato é que a área Ciência da Informação prescinde de estudos mais abrangentes sobre o este novo panorama da comunicação científica que integre as publicações eletrônicas, incluindo aí as revistas científicas, e as novas formas de organização destas publicações: os repositórios e os provedores de serviços.

Até o presente, a literatura da área da Ciência da Informação já comprovou que a internet potencializou o sistema de comunicação científica como um todo estruturado pelas relações sociais sem hierarquias, características próprias do contexto digital, assim como pela auto-gestão, simultaneidade do processo, heterogeneidade e a transdisciplinaridade. Assim, o autor/pesquisador interage diretamente no processo de

⁵ É possível pensar sobre a existência de níveis de organização da produção científica uma vez que a estrutura do modelo OAI é constituída de provedores de dados (nível secundário que se refere à reunião de publicações) e de serviços (nível terciário que se refere à reunião e replicação de metadados de dois ou mais repositórios).



produção científica e no fluxo das informações científicas sem intermediários além da possibilidade de incorporar outras atividades científicas incluindo a comunicação informal, o que antes não ocorria no sistema baseado em publicações impressas (BARRETO, 1998).

Todas essas mudanças parecem anunciar a re-estruturação do sistema de comunicação científica, no entanto se mantêm os pilares que sustentam a comunicação científica (fidedignidade, acessibilidade e disseminação) valorizando ainda mais o triplo papel do pesquisador enquanto produtor-disseminador-consumidor.

Algumas considerações sobre repositórios institucionais

Apesar das propostas de mudança da OAI terem sido implementadas ao longo da década de 1990, somente uma parcela pequena da comunidade científica está presente nestes repositórios. 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]).

No mundo todo ocorre movimentos da comunidade científica para que as novas regras do jogo do sistema de publicações científicas 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. O Movimento de Acesso Livre ao Conhecimento em 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.

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 tanto nas áreas das ciências quanto sociais, 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 estas iniciativas ocupem seu espaço.

São muitos os desafios que emergem neste cenário tão mutante. Alguns 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 serão mais rápidos em buscar soluções para implementação de repositórios institucionais ou temáticos. Essas soluções estão relacionadas tanto com a questão do financiamento quanto de organização 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 neste processo.

No entanto, como se trata de uma estrutura que atua diretamente na organização, controle, disseminação e acesso da produção científica, a implementação de repositórios digitais é urgente em todas as áreas do conhecimento, já que é possível que o ciclo da comunicação científica fique comprometido.

Considerações finais

Sob o pretexto de apresentar breves reflexões sobre os repositórios institucionais foram vistos afinal dois temas que emergem do contexto atual da comunicação científica: o desenvolvimento de iniciativas que apóiam a produção e organização de publicações científicas eletrônicas para promover o acesso livre e a necessidade de 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. 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 também representando institutos de pesquisa e universidades debater e propor soluções e estratégias de ação para alcançar este novo estágio de desenvolvimento científico. Até o momento, a melhor estratégia de ação recomendada pelas iniciativas de acesso livre é implementar os repositórios institucionais.

Referências bibliográficas



EPRINTS. *Self-archiving FAQ*. Disponível em: <<http://www.eprints.org/openaccess/self-faq/>>. Acesso em 11 de abril. 2006.

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.

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.

MEADOWS, A. J. *A comunicação científica*. Brasília: Briquet de Lemos/Livros, 1999. 268 p.

MIRANDA, Antonio. *Ciência da Informação*. Brasília: Thesaurus, 2003.

MOVIMENTO de Acesso Livre ao Conhecimento em Ciências e Humanidades -
<https://repositorium.sdum.uminho.pt/about/DeclaracaoBerlim.htm>

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.

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

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