

Ensaio

Do acesso livre à ciência aberta: conceitos e implicações na comunicação científica

From open access to open science: concepts and implications in scientific communication

El acceso libre a la ciencia abierta: conceptos e implicaciones en la comunicación científica

Lena Vania Ribeiro Pinheiroⁱ

RESUMO

Abordagem conceitual de acesso livre à informação científica e de ciência aberta, e análise das implicações desses dois fenômenos na comunicação científica. São enfatizados os benefícios tanto para cientistas e pesquisadores quanto para usuários de modo geral. No caso dos primeiros, expressos pela autonomia frente aos editores, mais visibilidade à sua produção científica e aumento de citações de suas pesquisas. Para a comunidade científica, na forma de amplo acesso à literatura da ciência e tecnologia, sem custos ou restrições. A terminologia estudada abrange, além dos dois termos centrais – via verde e via dourada – autoarquivamento, dados científicos, curadoria digital, publicações ampliadas, entre outros. No plano político, são mencionados alguns manifestos internacionais e o brasileiro, do Ibict, cujas ações de apoio e estímulo ao acesso livre à informação científica são descritas e configuradas como espaços de discussão e de produção de pesquisa sobre essas questões. No novo paradigma, os limites ou fronteiras entre o conhecimento gerado na área de ciência e tecnologia e aqueles que precisam desse saber construído em pesquisas são ultrapassados e se tornam ilimitados.

Palavras-chave: Acesso livre; Informação científica; Ciência aberta; Comunicação científica; Políticas de informação; Curadoria digital.

ABSTRACT

This essay deals with conceptual approach of open access to scientific information and of open science as well as an analysis of the implications of these phenomena for scientific communication. It shows their benefits to both scientists and researchers and to users. To the first ones insofar as they enjoy autonomy in relation to the editors, more visibility of their scientific production and an increase in the citation number of their researches. To the scientific community in general, the benefits should be seen as a large access to the science and technology literature without costs or restrictions. The terminology studied covers, besides the two central terms – the green road and the golden road – self-archiving, scientific data, digital curation, enhanced publications, among others. In the political setting some international and Brazilian manifestos as that from IBICT are mentioned. Their actions to support and encourage open access to scientific information are described and configured as a space to discussion and production of research about these issues. With the application of this new paradigm, the limits or boundaries between the knowledge generated by S&T and the people who need that knowledge based on the researches are overcome and become unlimited.

Keywords: Open access; Scientific information; Open science; Scientific communication; Information policies; Digital data curation.

ⁱDoutora em Comunicação e Cultura pela Universidade Federal do Rio de Janeiro (UFRJ). Pesquisadora e professora do Programa de Pós-Graduação em Ciência da Informação, criado pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) em parceria com a UFRJ. Rio de Janeiro, Brasil. lenavania@ibict.br

RESUMEN

Abordaje conceptual del acceso libre a la información científica y de ciencia abierta, y análisis de las implicaciones de esos dos fenómenos en la comunicación científica. Son enfatizados los beneficios tanto para los científicos e investigadores como para los usuarios de modo general. En los dos primeros casos, expresados por la autonomía frente a los editores, mas visibilidad para la producción científica y aumento de citaciones de sus investigaciones. Para la comunidad científica, en la forma de amplio acceso a la literatura de la ciencia y tecnología, sin costos o restricciones. La terminología estudiada abarca, además de los dos términos centrales – vía verde y vía dorada- auto-archivamiento, datos científicos, tutoría digital, publicaciones ampliadas, entre otros. En el plano político, son mencionados algunos manifiestos internacionales y el brasileño, del IBICT, cuyas acciones de apoyo y estímulo de acceso libre a la información científica son descritos y configurados como espacios de discusión y de producción de investigación sobre tales aspectos. En el nuevo paradigma, los límites o fronteras entre el conocimiento generado en el área de ciencia y tecnología y aquellos que necesitan de ese saber construido en investigaciones son ultrapasados y se tornan ilimitados.

Palabras clave: Acceso libre; Información científica; Ciencia abierta; Comunicación científica; Políticas de información; Tutoría digital.

Submetido: 20.maio.2014

Aceito: 26.maio.2014

Conflitos de interesse: Não há conflitos de interesse a declarar.

Fontes de financiamento: Não houve.

Introdução

Na história da humanidade, transformações têm marcado todas as atividades, das rotineiras às científicas e tecnológicas, entre as quais uma das mais marcantes e de maior impacto foi a Sociedade da Informação ou Sociedade do Conhecimento ou, ainda, da aprendizagem. Boaventura de Sousa Santos¹ expressa a dimensão dessas mudanças, a partir da revolução científica iniciada com Einstein e a mecânica quântica, e a “profunda e irreversível” crise da ciência, as mudanças de paradigmas, uma “nova ordem científica” emergente.

Por outro lado, os novos paradigmas, como o acesso livre à informação científica, afetaram profundamente a comunicação científica, tanto que têm estimulado inúmeras pesquisas, no Brasil e exterior, inclusive a publicação de fascículos temáticos de periódicos, em nosso país. Como exemplo, destacamos a revista *Ciência da informação*, do Ibiict, em 2006, com um número dedicado ao acesso livre à informação e, em 2012, a *Liinc*² em revista, sobre os “Novos paradigmas da comunicação científica: ampliando o debate”, ambas incluindo autores nacionais e estrangeiros. Em editorial elaborado para esse fascículo, do qual fui editora, juntamente com Helio Kuramoto, ressaltamos:

Os novos paradigmas que dão título a este fascículo especial da *Liinc* em Revista representam uma tentativa de refletir e discutir as principais questões que têm norteado os estudos e debates desse novo cenário, no exterior e no Brasil. O seu raio de ação atinge autores, indústria editorial científica, instituições de ensino e pesquisa, agências de fomento e demais atores.³

O sistema tradicional de comunicação científica foi profundamente afetado, juntamente com as questões que lhe são inerentes e decorrentes, como propriedade intelectual, autoria coletiva, direito autoral, produtividade científica, citação etc.

Esse momento estaria sob o domínio de uma “filosofia aberta”, assim denominada por Costa e Moreira⁴, correspondendo a

um movimento observado nos últimos anos em direção ao uso de ferramentas, estratégias e metodologias que denotam um novo modelo de representar um igualmente novo processo de comunicação científica, ao mesmo tempo em que serve de base para interpretá-lo.

Alguns componentes são fundamentais nesse processo:

- software aberto (ou livre), para o desenvolvimento de aplicações em computador;
- arquivos abertos, para interoperabilidade em nível global; e
- acesso aberto – questão mais polêmica – para a disseminação ampla e irrestrita de resultados da pesquisa científica.⁴

A publicação de mais um fascículo dedicado a essas questões, pela Revista Eletrônica de Informação, Comunicação & Inovação em Saúde (Reciis), do Instituto de Comunicação e Informação Científica e Tecnológica da Fiocruz, responde tanto à necessidade, ainda, de mais pesquisas e aprofundamento de complexas e contemporâneas questões, quanto à sua amplitude, que hoje se estende até a chamada ciência aberta (“open science”), daí o título deste artigo.

Esta opção não significa, de modo algum, desconhecer ou não considerar os relevantes e múltiplos aspectos relacionados ao direito à informação, como a Lei de Acesso à informação, nº 12.527, de 18 de novembro de 2011⁵, que trata de um direito mais amplo e fundamental, de todo e qualquer cidadão brasileiro, bem como as iniciativas da Comissão Nacional da Verdade.

Este artigo é voltado à comunicação científica, aí incluídas as políticas de informação. É oportuno lembrar a afirmativa de Cubillo, autor chileno estudioso dessa área, citado por Moura⁶, de que “lamentavelmente, as políticas de informação pública são consideradas uma parente pobre das políticas públicas e que jamais figuraram como prioridade nos países da América Latina, sobretudo devido à mentalidade neoliberal, que açambarcou grande parte da gestão pública desses países”.ⁱⁱ

Neste contexto, é fundamental trazer à discussão o conceito de “regime de informação”, formulado por Frohman⁷ e descrito como “[...] um sistema ou rede mais ou menos estável na qual a informação flui através de canais determináveis - de produtores específicos, via estruturas organizacionais específicas, a consumidores ou usuários específicos”. Posteriormente, Gonzalez de Gómez⁸ desenvolve este conceito em alguns estudos, entre os quais o apresentado com Chicanel, no Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB), quando o sintetiza:

“[...] pelo modo de produção informacional dominante numa formação social, suas escolhas prescritivas dos que serão reconhecidos como sujeitos, instituições, regras e autoridades informacionais, assim como dos padrões de excelência e os critérios preferenciais de processamento seletivo de meios e recursos de informação.”⁸

Podemos afirmar que o regime de informação na comunicação científica foi alterado, a partir das políticas de acesso livre à informação e impactou todos os atores participantes do processo.

O presente artigo tem por objetivo discutir, tendo como fio condutor conceitos de acesso livre à informação científica até ciência aberta, as transformações na comunicação científica e suas implicações. Considerando as inúmeras questões imbricadas no tema, não é possível abranger a totalidade dos aspectos envolvidos, mas parte significativa, de forma a configurar o sistema atual de comunicação científica, expandindo o tema original, no caso, o acesso livre à informação científica.

ⁱⁱCubillo, J. apud ⁶.

É importante mencionar a Declaração de Berlim, de 2003, sobre a qual Sayão e Sales⁹ destacam a ampliação do conceito de acesso livre ao incluir “resultados de pesquisas científicas originais, dados não processados e metadados, fontes originais [...]”.

Este é ponto central do presente artigo, ao ser direcionado à expansão ou extensão de acesso livre à informação científica para ciência aberta, o que significa mais um patamar na conquista da democratização e transparência do conhecimento gerado na área de ciência e tecnologia. Pode ser concretizado pela abertura e acesso a dados científicos a cientistas, pesquisadores e todos aqueles que precisarem recorrer a essas informações.

Acesso livre à informação científica

O movimento de acesso livre à informação eclodiu como uma reação ao domínio das editoras no processo editorial de periódicos científicos, portanto, sobre os autores, e consequências como o alto custo das assinaturas, que atingiu bibliotecas e usuários em geral, limitados na possibilidade de serem informados sobre o conhecimento gerado e publicado em periódicos.

Segundo Costa,⁴ este movimento “representa uma das primeiras fortes reações da comunidade científica que ocorreu de forma planejada, organizada e envolvendo uma variedade de atores”.

As primeiras iniciativas são o ArXiv, de Los Alamos, nos EUA, e o EPrints, de Southampton, na Grã Bretanha, seguidos da Budapest Open Access Initiative (BOAI) em 2002, em torno do software e arquivos abertos. É importante lembrar que o objetivo da reunião de Budapeste foi “acelerar o progresso do esforço internacional de tornar artigos de pesquisa em todos os campos acadêmicos livremente disponíveis na Internet” e de forma irrestrita considerando, especificamente, o “acesso aberto à literatura periódica referendada”.⁴

No Brasil, desde meados dos anos 2000, o Ibict vem participando de reuniões, juntamente com a Universidade do Minho, que redundaram no Compromisso do Minho sobre o Acesso Livre à Informação em Países Lusófonos, assinado, entre outros, por representantes do Ibict. Da mesma forma tem participado, ativamente, das Conferências Luso-Brasileiras de Acesso Aberto (Confoa), seja na organização, na Comissão Científica ou na apresentação de trabalhos ou pôsteres, que alternadamente tem se realizado no Brasil ou em Portugal, onde será este ano a quinta Conferência, na cidade de Coimbra.

Conceitos básicos do acesso livre à informação científica

Se pensarmos que a expressão acesso à informação significava, inicialmente, conseguir o documento cuja referência foi recuperada num catálogo de biblioteca convencional ou em uma base de dados, no acervo dessa biblioteca ou por meio de cópia, podemos dimensionar os avanços que as tecnologias de informação e comunicação (TIC) proporcionaram às comunidades científicas e aos usuários em geral. Naquela fase, as bases de dados de texto completo eram raras, o que hoje se tornou natural com as bibliotecas digitais e repositórios. Integra essa cadeia terminológica a acessibilidade, relativa tanto à facilidade de acesso físico de deficientes em geral, especificamente a unidades de informação, quanto à disponibilidade de um sistema amigável homem-máquina, com base nos princípios ergonômicos.

Segundo o decreto-lei 5.296, de 02/12/2004, artigo 8º, acessibilidade é a “condição para utilização, com segurança e autonomia, total ou assistida, dos espaços, mobiliários e equipamentos urbanos, das edificações, dos serviços de transporte e dos dispositivos, sistemas e meios de comunicação e informação, por pessoa portadora de deficiência ou com mobilidade reduzida”.¹⁰

Embora expressões próximas semanticamente, acesso livre e acesso aberto não têm o mesmo significado, ainda que alguns autores os considerem sinônimos e uns adotem acesso livre e, outros, acesso aberto para a tradução do inglês “open access”. Outras expressões integram uma rede conceitual relativa à questão, como via verde (“green road”) e via dourada (“golden road”), definidas por Harnad e colaboradores. A primeira corresponde ao autoarqui-

vamento, pelos próprios autores, de seus artigos publicados ou aceitos em periódicos científicos que adotem a avaliação pelos pares. Neste processo, os repositórios representam o mecanismo para sua concretização, tanto os temáticos quanto os institucionais e, mais recentemente, os repositórios de dados científicos, os quais serão abordados no próximo tópico deste artigo. Na via dourada o livre acesso é proporcionado pelos editores de periódicos cujo conteúdo já é de acesso livre.ⁱⁱⁱ⁴

Pinheiro e Kuramoto³, já citados, reconhecem que:

O acesso livre está no centro das questões porque atinge diretamente o sistema tradicional de comunicação científica e, como um desdobramento, a propriedade intelectual, criando impasses contemporâneos ainda não compreendidos no seu âmbito e para cuja superação são necessários muitos estudos e pesquisas.

Sobre a denominação acesso livre ou acesso aberto, Hélène Bosc e Hans Dillaerts^{iv} ilustram as diferenças conceituais com o supermercado, cujo acesso é aberto, mas os seus produtos não são livres, têm preço para venda e compra. Pinheiro e Kuramoto³ concordam com essa opção pelas mesmas razões e a reforçam, exemplificando com o conceito original formulado pelos que lideraram o acesso livre, como Stevan Harnad e Peter Suber. Para esses autores, a produção científica deveria ser livre de custos, portanto, não basta acessar a informação, é necessário que esse acesso não represente custos e esteja livre das restrições de direitos de autoria e licenciamentos.

Sobre os benefícios do acesso livre para a pesquisa e pesquisadores, constatados por inúmeros estudos, Bosc e Dillaerts¹¹ destacam o quanto o mandato aumenta o índice de autoarquivamento, chegando a 80%.

No entanto, ainda é pequeno o número de instituições que aderiram à obrigatoriedade do arquivamento dos artigos científicos, pelos próprios pesquisadores/autores, a partir do reconhecimento de que pesquisas financiadas com recursos públicos devem estar disponíveis a toda a sociedade. Um bom exemplo desta prática é a Universidade do Minho, em Portugal, onde o autoarquivamento é mandatário.

Bosc e Dillaerts¹¹ enfocam o acesso livre na França. Ressaltamos diferenças entre as áreas de conhecimento, o que é natural na comunicação científica e foi sempre explicitado em pesquisas dessa natureza. Não se trata apenas de distinções no comportamento de cientistas na sua produção e geração de conhecimento na ciência e tecnologia, ou entre as ciências exatas e naturais e as humanidades mas, mesmo entre os campos do conhecimento no âmbito de uma grande área como as ciências sociais. A ciência da informação, por exemplo, enquadrada como ciências sociais aplicadas, apresenta características e especificidades que a distinguem na sua produção e na forma de gerá-la, por exemplo, dos campos do conhecimento do direito ou da economia.

Não apenas cientistas e pesquisadores são beneficiados pela maior visibilidade de suas pesquisas, pelo prestígio, reconhecimento pelos pares e pelas vantagens, como bolsas de estudos e prêmios, como leitores e usuários têm ampliada a possibilidade de acessar literatura do mundo todo, o que foi pontuado na reunião do BOAI.⁴

Entre os estudos citados por Bosc e Dillaerts¹¹ estão os de Gargouri, que identificam, em nível mundial, a porcentagem de autoarquivamento espontâneo aumentando de 15% a 21%, de 2005 a 2010, com crescimento de 1% ao ano. Por outro lado, inúmeras pesquisas realizadas depois de 2005 atestam que somente por meio de mandato, portanto, de obrigatoriedade, é possível chegar ao arquivamento de 80%, num período de dois ou três anos, como comprovam Sale^v e os repositórios da Universidade de Liège.¹¹ Sobre os resultados de pesquisa nessa universidade, são apresentadas estatísticas que comprovam suas afirmativas, ilustradas por gráficos.

ⁱⁱⁱHarnad S. et al. The access/impact problem and the green and gold roads to open access. 2001. Disponível em: <<http://www.ecs.soton.ac.uk/~harnad/Temp/impact.html>> apud⁴. p. 41

^{iv}Segundo informações da autora principal, Bosc, a primeira parte do artigo de Bosc e Dillaerts¹¹, relativa a acesso livre, foi escrita por ela, e a referente aos periódicos eletrônicos de livre acesso, por Dillaerts.

^vSALE, A. Comparison of content policies for institutional repositories in Australia. *First Monday*, v. 11, n. 4, 3 Apr. 2006. Disponível em: <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1324/1244> apud¹¹

Na Conferência mencionada no início deste artigo, a Confoa, têm sido muitas as comunicações apresentadas sobre acesso livre à informação, tema do próprio evento, o que inclui algumas pesquisas específicas sobre repositórios e bibliotecas digitais, instrumentos viabilizadores desse processo.

Política de acesso livre à informação científica do IBICT

No Brasil, o Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), além do Manifesto do Acesso Livre à Informação Científica, assinado em 2005, tem disponibilizado ferramentas e proporcionado atividades que o estimulam e apoiam. Entre estas, o Sistema Eletrônico de Editoração de Revistas (SEER), a Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), o Diadorim e a incubadora INSEER.^{12,13,14,15}

O SEER “é um software desenvolvido para a construção e gestão de uma publicação periódica eletrônica”, originado do Open Journal Systems, software desenvolvido pela Universidade British Columbia, traduzido e customizado pelo Ibict. Abrange as atividades básicas que constituem esse processo e tem, como um dos objetivos, contribuir para agilização e aperfeiçoamento da qualidade dos periódicos, inclusive pela adoção de padrões internacionais. Complementam as vantagens do uso do SEER a disseminação e preservação dos conteúdos de revistas.¹² A ampla ação política do Ibict não se restringiu à formulação do manifesto e fornecimento de instrumentos e se completa pela promoção de treinamentos frequentes, gratuitos, na forma presencial e a distância, cujo resultado é comprovado pela adoção do SEER.

A BDTD “tem por objetivo integrar, em um único portal, os sistemas de informação de teses e dissertações existentes no país e disponibilizar para os usuários um catálogo nacional de teses e dissertações em texto integral, possibilitando uma forma única de busca e acesso a esses documentos”. Adota as tecnologias OAI (Open Archives Initiative) e tem como modelo os padrões de interoperabilidade. É oportuno esclarecer que o documento, na sua íntegra, bem como os metadados, são da responsabilidade da instituição onde são produzidos, portanto, a coleta é descentralizada, com operações em rede. Complementando o suporte ao processo, o Ibict desenvolve, disponibiliza (*download* em <<http://tedesite.ibict.br>>) e distribui o Sistema TEDE, gerador de bibliotecas digitais de teses e dissertações institucionais, a metodologia de implantação, manuais operacionais e de usuário e documentação e treinamento.¹³

Outra ação é a do Diadorim, “diretório/serviço de informações da política editorial das Revistas científicas brasileiras sobre o armazenamento dos seus artigos em Repositórios Institucionais de Acesso Aberto (RI)”. A sua função principal é auxiliar autores, editores de revistas e gestores de repositórios institucionais na localização de informações sobre o assunto, especificamente sobre a permissão para o armazenamento de artigos em RI.¹⁴

Finalmente, a Incubadora de Revistas INSEER, criada pelo Ibict com apoio da Financiadora de Estudos e Projetos (Finep), cujo objetivo principal é “dar suporte e estimular a construção e manutenção de revistas científicas de acesso livre na Internet”. Com essa iniciativa, o Instituto procura estimular a criação não somente de novas publicações seriadas, mas também abrir novas “oportunidades para a sustentabilidade de revistas científicas já existentes, especialmente aquelas que não estejam ainda disponíveis em meio eletrônico”. Assim, o processo editorial eletrônico das revistas hospedadas é otimizado e estas alcançam maior visibilidade, num período até que adquiram infraestrutura tecnológica que as sustentem.¹⁵

As ações políticas do Ibict são fundamentais como estímulo e sustentação do acesso livre à informação científica no Brasil, bem como as das universidades brasileiras, como a Universidade de São Paulo (USP), Universidade de Brasília (UNB), Universidade Federal de Minas Gerais (UFMG), Universidade Federal da Bahia (UFBA), Universidade Federal da Paraíba (UFPB) etc.

Ciência aberta (“open science”)

Quando a internet foi lançada, muitas indagações cercaram essa nova tecnologia. Uma das mais inquietantes era pensar que todo um processo de pesquisa, que inclui inúmeras etapas – dados primários, cálculos, estatísticas, anotações – em geral não incluídos nas publicações, como artigos, até por sua extensão, poderiam ser perdidos definitivamente, se não registrados, preservados eletronicamente e armazenados para futura recuperação.

Antes da internet e do acesso livre à informação científica, pouco ou quase nada se discutia ou pesquisava sobre a hoje denominada ciência aberta. De maneira rara e dispersa, algum autor abordava os dados científicos e raras também eram as bases de dados com esse tipo de conteúdo. No Brasil, uma das exceções e das primeiras manifestações foi a base de dados científicos sobre primatas, Primatam, do Museu Paraense Emílio Goeldi, em Belém do Pará. Esta base serviu de exemplo para o Ibict, no início dos anos 1980, como base de dados científicos, em definição elaborada pela autora deste artigo, sob a orientação de Michel Menou, na ocasião consultor do Ibict.

A ciência aberta e a eScience coexistem com um termo próximo que é Big Data, mais amplo, não diz respeito à ciência e pode ser definida como “um conjunto de problemas e suas soluções tecnológicas com características que tornam seus dados difíceis de tratar”^{vi}, marcados por volume, velocidade e variedade.¹⁶

Não foram somente as TIC e o acesso livre à informação que propiciaram a hoje chamada eScience, mas também a complexidade das ciências, com equipes mais numerosas trabalhando cada vez mais em colaboração, inclusive internacionalmente. Portanto, não se trata apenas de um fenômeno tecnológico, mas científico e político, entre outras instâncias.

A incomensurável quantidade de dados disponibilizados pelas mais diversas tecnologias configuram o novo paradigma, o quarto, definido pela ciência orientada por dados ou eScience. Em outras palavras, “o fazer científico é reordenado pela intensificação do uso de redes e de computadores e pelo uso sem precedentes de conjuntos de dados distribuídos”^{vii}.⁹ Neste também novo cenário, cientistas da computação, da informação e das mais distintas áreas travam novos diálogos e representam outros papéis. Não pode ser esquecido que a validade e a importância dos dados somente afloram quando são analisados e reutilizados, gerando novos conhecimentos. No nosso entendimento, a ciência não é diferente em si mesma, continua adotando seus princípios, metodologias, fiel à sua ética, mas os recursos e o instrumental tecnológicos disponíveis é que potencializam os seus resultados e perspectivas.

Nos Estados Unidos, onde desde sempre o governo esteve atento ao poder estratégico da informação, o famoso relatório Weimberg pode ser considerado um exemplo dessa preocupação, no início dos anos 1960.

Nos dias atuais, o relatório do National Science Board¹⁷, em ação conjunta e com apoio da National Science Foundation, representa manifestação semelhante, por parte do governo dos EUA. Trata-se de um documento político, de reconhecimento de que as coleções digitais de dados reúnem “poder catalizador do progresso e democratização da ciência e educação”. O foco desse documento está voltado à “criação e manutenção de coleções e sua rápida multiplicação ou potencial de décadas de curadoria”. Por outro lado, as coleções digitais também concentram “poder de Inclusão, remoção de barreiras para participação de todas as idades e níveis de educação”.¹⁷

No relatório, são traçados alguns conceitos e questões relacionadas à eScience; entre estes, dados, coleções, coleções digitais e “*long lived*”.

O primeiro, dados, é compreendido como “qualquer informação que pode ser armazenada na forma digital, incluindo texto, números, imagens, vídeo ou filmes, áudio, software, algoritmos, equações, animações, modelos, simulações etc. Tais dados podem ser gerados por vários meios, incluindo observação, computação e experimento”.¹⁷

^{vi}Xexéo, G. Big data computação para uma sociedade conectada e digitalizada. Ciência Hoje, n.306, p.18--23, ago. 2013. Disponível em: cienciahoje.uol.com.br/revista-ch/2013/306/pdf.../bigdata306.pdf/.../file apud¹⁶

^{vii}Cesar Júnior RM. Do mundo aos dados e dos dados ao conhecimento. In: Hey T, Tansley S, Tolle K (Org.). O quarto paradigma: descobertas científicas na era da eScience. São Paulo: Oficina do Texto; 2011 apud⁹.

O outro termo definido foi coleção, “usado para se referir não somente ao estoque de dados, mas também à infraestrutura, organizações e indivíduos, necessários para preservar o acesso aos dados”. Esta definição serve de base à subseqüente, coleção digital, aquela “que pode ser acessada eletronicamente, via Internet, por exemplo,” objeto do relatório. Está relacionada, ainda, à longa duração (“*long-lived*”), por sua vez “fornecida pelos padrões do Open Archival Information System (OAIS), ou seja, a um período suficientemente longo para que haja preocupação com os impactos das mudanças de tecnologia”.¹⁷

Note-se que o enfoque é de coleções digitais de dados, não especificamente os científicos.

No relatório da National Science Board (NSB) são enfatizados determinados aspectos e componentes das coleções digitais de dados, entre os quais merecem destaque:

- elementos do universo de coleções de dados digitais;
- papéis e responsabilidades de indivíduos e instituições (autores, gestores, cientistas, agências de fomento); e
- perspectivas políticas de coleções de dados digitais.¹⁷

Outro documento político foi emanado da Organização para Cooperação e Desenvolvimento da Economia (OCDE^{viii}), elaborado em 2007 e citado por Sayão e Sales.¹⁶ Nesse importante relatório, o uso efetivo de dados é considerado “condição crítica para políticas de ciência e tecnologia”. Nele, são apontadas inúmeras circunstâncias, entre as quais a cadeia de inovação, a cooperação internacional, a promoção de novas pesquisas e testes de hipóteses novas e alternativas, no desenvolvimento de novas pesquisas, entre outras e, sobretudo, “na promoção de uma atividade científica mais aberta e mais transparente, que tenha como princípio produzir conhecimento publicamente acessível”.¹⁶

No Brasil, Sayão e Sales^{9,16ix} e Sales e Sayão¹⁸, já citados neste artigo, aparecem entre os poucos autores dedicados a essa questão contemporânea, tendo produzido pesquisas sobre dados científicos e os diferentes elementos envolvidos nesse processo, sob o enfoque da ciência da informação; daí a opção por analisar os seus artigos. Entre os poucos documentos brasileiros citados por ambos estão os de Didati em 2003, sem necessariamente direcionar o foco à comunicação científica, privilegiado em suas pesquisas.

Para ilustrar o quanto dados científicos são importantes, Sayão e Sales recorrem a um exemplo eloquente, extraído de Didati^x.⁹ Trata-se de um relato sobre a gripe espanhola de 1918/1919, que matou entre 20 e 80 milhões de pessoas e se extinguiu sem intervenção de vacinas ou tratamento específicos. Somente muitos anos depois foram encontrados vestígios em tecido humano, localizados em um hospital militar da Inglaterra, onde também foram preservados arquivos científicos de 1916. Graças a essa descoberta foi possível desenvolver vacinas e meios para tratamento e prevenção à gripe espanhola.

Uma das constatações que motivaram Sayão e Sales foi o reconhecimento de que dados científicos são únicos e que, atualmente, com a proliferação de documentos digitais, estão muito vulneráveis à perda e à impossibilidade de recuperação. Ao mesmo tempo, a literatura que estudaram apontou que cada vez mais cientistas confiam nos dados de pesquisa de seus colegas.^{xix} No entanto, é oportuno lembrar que na comunicação científica, desde os seus primórdios, cientistas estabeleciam um elo de confiança com seus pares e diálogo sobre suas pesquisas e descobertas, o que

^{viii}OCDE. Principles and guidelines for access to research data from public data2007.[Acesso em: 17 fev. 2012]. Disponível em: <<http://www.oecd.org/dataoecd/9/61/38500813.pdf>> apud ¹⁶.

^{ix}Estes artigos têm relação direta com a tese de Luana Sales, em fase de finalização no Programa de Pós-Graduação em Ciência da Informação, IbiCT-UFRJ, sobre essa temática e tendo como orientadores Rosali Fernandez de Souza e Luís Fernando Sayão. Por outro lado, o lócus da pesquisa é a Comissão Nacional de Energia Nuclear (CNEN), instituição onde exercem atividades os dois autores.

^xDidati C. Preservação de documentos eletrônicos. Rio de Janeiro: Arquivo Nacional/ CTDE; 2003 apud ⁹.

^{xi}About Daisy. What is digital curation?Edinburgh, UK: Digital Curation Centre; 2008. [Acesso em 20 dez 2011]. Disponível em: <http://www.era.lib.ed.ac.uk/bitstream/1842/3362/3/Abbott%20What%20is%20digital%20curation_%20_%20Digital%20Curation%20Centre.doc> apud ⁹.

pode ser verificado na História da Ciência, com alguns exemplos como o da carta de Galileu para Kepler, ao olhar pela primeira vez os anéis de Saturno.²¹ Além disso, existe a frase atribuída a Isaac Newton, comprovando que cientistas tomam por base de suas pesquisas a produção científica de seus antecessores, reconhecendo que se apoiavam “nos ombros de gigantes”. No entanto, na sua tese de doutorado, Moura²² revela que, na realidade, a frase é muito mais antiga e seu autor foi Clarvaux, em 1100: “somos anões apoiados nos ombros de gigantes, capazes de ver mais além não por virtude de brilho individual, mas pela mestria dos clássicos”.^{xii22}

A partir de seus estudos, Sayão e Sales⁶ fazem duas afirmativas contundentes: a primeira, que “os objetos digitais nunca sobrevivem inercialmente como os seus equivalentes impressos” e, a segunda, que “dados de pesquisa são únicos e não podem ser substituídos se forem destruídos ou perdidos”, o que demonstra o papel único e relevante que representam.

No segundo trabalho, Sales e Sayão¹⁸ têm por objetivo estudar a curadoria digital de dados, para cuja concretização é fundamental o protocolo Open Archives for Metadata Harvesting. Este, por sua vez, é considerado um novo patamar de comunicação científica, oriundo das próprias comunidades de cientistas e padrão de compartilhamento, interoperabilidade e reuso de metadados, como a memória científica virtual. Os autores, então, se perguntam: “se compartilhamos os resultados, por que não compartilhamos os dados que levaram a tais resultados?” Esta é a questão.

Algumas iniciativas nesse sentido merecem destaque, como o Genbank, do International Nucleotide Sequence Database Collaboration (INSDC), base de dados de informações sobre o sequenciamento genômico de diferentes espécies e o Digital Curation Centre (DCC) exemplo de curadoria digital.¹⁸

Consideramos oportuno observar que o termo curadoria é próprio da Museologia:

[...] ao falarmos de ação curatorial, não tratamos apenas do estudo de coleções, mas também e principalmente do sentido atribuído a elas, ou seja, é a partir da ação curatorial, que se estabelece uma identidade para o acervo e quando esta identidade é bem trabalhada, as exposições passam a ser “agentes de informação para o público visitante”.^{xiii 23}

No terceiro trabalho, Sayão e Sales¹⁶ têm por objetivo contribuir para o estabelecimento de um modelo de curadoria digital para o Brasil e partem da constatação de que em nosso país “ainda são poucas e fragmentadas as ações em torno do tema, agravadas, ainda, pela incompreensão de suas potencialidades e pela falta de visão do futuro”. Assim, “há necessidade do desenvolvimento e implantação de modelos teóricos e práticos de avaliação e de desenvolvimento de coleções de dados de pesquisa, processo no qual são importantes estudos em Ciência da Informação e em Ciência da Computação”.

Na proposta de modelo, Sayão e Sales¹⁶ abordam algumas questões fundamentais como formação e desenvolvimento de coleções de dados digitais de pesquisa, princípios e diretrizes para o acesso a dados de pesquisa, e as instâncias políticas, organizacionais, sociais, legais e éticas, infraestrutura tecnológica e de padronização, formação de recursos humanos, sustentabilidade econômica e serviços. A cada um desses aspectos, os autores apontam e descrevem os componentes necessários que constituem partes essenciais do modelo de coleção de dados digitais de pesquisa.

^{xii}Lindberg apud ²².

^{xiii}Barbosa Barbosa CR. As diversas faces do curador de exposições científicas e tecnológicas. In: Julião L, Bittencourt JN (Orgs.). Cadernos de Diretrizes Museológicas, 2. Mediação em museus: curadoria, exposições, ação educativa. Belo Horizonte: Secretaria de Estado de Minas Gerais. Superintendência de Museus; 2008. 78-89 apud ²³.

Conceitos básicos relativos à ciência aberta

Este novo panorama e circunstâncias do fazer ciência, com recursos tecnológicos antes não existentes, engendram terminologias para representar atores, canais, processos, instrumentos etc., entre estes dados científicos, gestão de dados de pesquisa, curadoria digital, publicações ampliadas.¹⁸

Os autores citam a definição de dados científicos ou dados de pesquisa, de acordo com a OCDE: “registros de fatos usados como fontes primárias na investigação científica e que geralmente são aceitos na comunidade científica como necessários para a validade dos dados de pesquisa”.¹⁸

Os mesmos autores estabelecem uma distinção entre dados intermediários e dados finais de pesquisa. Os primeiros são os obtidos do processo preliminar de pesquisa e, “com muita frequência, não são arquivados ou permanecem inacessíveis para outros pesquisadores”, e os segundos são “os selecionados para serem incluídos em bases de dados”.¹⁶

Baseados em Higgins¹⁹ e Conway²⁰, o conceito de curadoria digital foi formulado por Sales e Sayão¹⁶ e diz respeito à gestão e preservação de recursos digitais durante o ciclo de vida de interesse para comunidades científicas, com perspectiva de atender a gerações atuais e futuras, o que inclui arquivamento e preservação digitais.

Outro conceito é o de publicações ampliadas (*enhanced publication*), assim definidas: “são instâncias de objetos digitais complexos que combinam vários recursos heterogêneos relacionados para uma mesma finalidade”. Como exemplo os autores citam uma tese, registrada em um repositório institucional ou temático, “que agregue os dados de pesquisa utilizados para que ela pudesse ser gerada”.¹⁸

Como decorrência desse novo panorama, são elaborados, além dos repositórios, institucionais e temáticos, os repositórios de dados científicos que, juntamente com a biblioteca digital, têm por finalidade garantir a permanência e durabilidade dessas informações. Para tal, a preservação digital é uma ação essencial.

Em 2013, tese sobre acesso livre e direito autoral, tomando para pesquisa a própria comunidade científica da ciência da informação, no Brasil em Portugal, foi desenvolvida por Jaury Nepomuceno de Oliveira, no Programa de Pós-Graduação em Ciência da Informação, convênio IbiCT-UFRJ, com doutorado sanduíche na Universidade de Coimbra. A temática, sua abrangência e período (2001-2011) tornam a pesquisa de interesse particular para o presente artigo, especialmente o capítulo “O quadro atual do acesso livre na Ciência da Informação no Brasil e em Portugal”. O autor ressalta os avanços do acesso livre em Portugal, liderado pela Universidade do Minho, com participação das universidades de Coimbra, do Porto e de Lisboa, entre outras. Destacam-se como principais instrumentos a Declaração OA do Conselho de Reitores das Universidades Portuguesas (CRUP), de 2006, e o Repositório Científico de Acesso Aberto de Portugal (RCAAP), implantado em 2008, constituído por um portal único de acesso à produção científica portuguesa. Para o autor, “o pressuposto geral e irrestrito do acesso livre é que a obra científica depositada em repositórios de *pré-print* ou *pós-print*, com a revisão dos pares, possa e deva ser acessada livremente, descarregada, copiada e utilizada de maneira correta na sua íntegra. Com essas ações, em Portugal houve maior crescimento de repositórios digitais do que no Brasil, país onde a Biblioteca Digital de Teses e Dissertações (BDTD), já mencionada neste artigo, tem obtido resultados consideráveis, hoje alcançando mais de 200 mil registros.”²¹

Na ocasião do levantamento para a pesquisa empírica de sua tese, Oliveira²¹ identificou um número significativo de revistas científicas brasileiras na área da ciência da informação, em um total de vinte (20), bem classificadas no Qualis da CAPES (A e B), portanto, nos padrões da via dourada, enquanto em Portugal existe somente uma de acesso livre à informação. Embora o objetivo de sua pesquisa não fosse um estudo comparativo, esse pesquisador observou algumas distinções, entre Brasil e Portugal como, por exemplo, “que as questões de direitos autorais não são suficientemente conhecidas pela comunidade científica da área no Brasil, enquanto em Portugal há maior conscientização dos seus pesquisadores sobre as necessidades e obrigações relativas aos seus direitos autorais”. Por outro lado, a maioria dos periódicos científicos no campo, nasceu ou migrou para o formato eletrônico, e os artigos com esse formato atingem 75%, em Portugal, e 97% em nosso país.²¹

Com esses resultados, podemos afirmar que tanto o Brasil quanto Portugal estão perfeitamente engajados nas políticas e ações de acesso livre, o que abre perspectivas para uma verdadeira “ciência de multidão”.

Reflexões finais

Estudar conceitos relacionados ao acesso livre à informação científica e ciência aberta, no âmbito da comunicação científica, por sua vez no escopo da ciência da informação, sobretudo nas suas injunções políticas, demonstrou que este processo é irreversível em nível mundial.

Apesar dos evidentes benefícios do acesso livre para pesquisadores e cientistas em geral, com autonomia em relação aos editores e trazendo mais visibilidade à sua produção científica e, conseqüentemente, maior índice de citação, a literatura estudada demonstra que nem todos os campos científicos reagem igualmente às novas circunstâncias da ciência e tecnologia. Por outro lado, este é um processo em que todos os setores e agentes envolvidos – autores, editores, universidades e institutos de pesquisa, órgãos de fomento – devem estar engajados e, nas suas respectivas funções e atividades, aderir ao movimento e criar mecanismos para a sua concretização.

O autoarquivamento, por exemplo, deve ser efetivado por meio de ações mandatórias emanadas de instituições de ensino e pesquisa, como forma de alcançar maior índice.

No Brasil, sobretudo o Ibict, com o seu manifesto e inúmeras iniciativas de estímulo e apoio, além de universidades brasileiras e, em Portugal, a Universidade do Minho, que lidera o processo e outras são responsáveis pelos grandes avanços do movimento de acesso livre à informação científica, cujos resultados são apresentados na Conferência Luso-Brasileira de Acesso Livre, alternadamente realizada no Brasil e em Portugal.

A ciência aberta representa um alargamento do acesso livre, tornando acessíveis dados científicos, únicos e insubstituíveis, dos mais diversos tipos, básicos para pesquisas, mas em geral não publicados, a outros e futuros pesquisadores, para a sua reutilização. Assim, são abertas perspectivas para novos conhecimentos, quem sabe, quem mandando etapas e agilizando o processo de fazer ciência. No entanto, esta ação exige esforços maiores para registro e recuperação da informação, o que torna a sua gestão, denominada curadoria digital, mais complexa.

No exterior, estudos e relatórios de órgãos importantes, especialmente dos Estados Unidos, têm tratado o assunto com a atenção que merece, estabelecendo parâmetros e normas para a sua implantação. No Brasil, a questão está em fase inicial e aparece em poucas pesquisas, inclusive na ciência da informação, enfocando a terminologia, a metodologia e modelos, entre outros aspectos inerentes à sua concretização. Neste processo a preservação digital é um ponto central, que manterá os dados científicos e assegurará o sucesso da nova e complexa empreitada, sob a forma de documentos ampliados.

Tanto para o acesso livre à informação científica, quanto para a ciência aberta, foram as tecnologias de informação e comunicação (TIC) que criaram condições de sua concretização. No entanto, é fundamental reconhecer que esta questão transcende as ferramentas tecnológicas, depende e exige um árduo esforço da ciência, do ensino, da pesquisa, da política, da economia e da cultura, entre os muitos fatores e componentes envolvidos neste complexo fenômeno. Da sua convergência talvez seja possível a utopia da “ciência de multidão” tornar-se uma realidade no mundo contemporâneo.

Referências

1. Santos BS. Um discurso sobre as ciências. 13thed. Porto: Edições Afrontamento; 2002.
2. Novos paradigmas da comunicação científica: ampliando o debate. Liinc rev.2012; 8(2):307-508. Disponível em: <http://revista.ibict.br/liinc/index.php/liinc/issue/view/40>.
3. PinheiroLVR, Kuramoto H. Novos paradigmas da comunicação científica: ampliando o debate (editorial). LIINC rev. 2012;8(2):307-10. Disponível em: <http://revista.ibict.br/liinc/index.php/liinc/article/viewFile/501/371>.

4. Costa SMS. Filosofia aberta, modelos de negócios e agências de fomento: elementos essenciais a uma discussão sobre o acesso aberto à informação científica. *Ci Info*. 2006; 35(2):39-50.
5. Brasil. Lei de Acesso à Informação nº12.527, de 18 nov 2011. Diário Oficial da União, Brasília (DF); 18 nov 2011; Seção 1:1.
6. Moura MA. Introdução. In: Moura MA, organizador. *A construção social do acesso público à informação no Brasil: contexto, historicidade e repercussões*. Belo Horizonte: Editora UFMG; 2014. P.19-23.
7. Frohmann B. Taking information policy beyond information science: applying the actor network theory. In: *Proceedings of the 23rd Annual Conference of the Canadian Association for Information Science*; 1995 June 7-10; Edmonton, Alberta. Olson HA, Ward DB, editors. Edmonton, Alberta: School of Library and Information Studies, University of Alberta; 1995. p. 19-95.
8. Gonzalez de Gomez MN, Chicanel M. A mudança de regimes de informação e as variações tecnológicas. In: *XI ENANCIB Encontro Nacional de Pesquisa em Ciência da Informação*; 2008, São Paulo. Anais do XI ENANCIB, 2008. São Paulo: Universidade de São Paulo; 2008. p. 45-59.
9. Sayão LF, Sales LF. Curadoria digital: um novo patamar para preservação de dados digitais de pesquisa. *Inf e Soc*. 2012;22 (3):179-91.
10. Brasil. Decreto-lei 5.296, de 2 dec 2004. Diário Oficial da União, Brasília (DF); 2 dec 2004; Seção 1:5.
11. Bosc H, Dillaerts H. Le libre accès en France en 2012: entre immobilisme et innovation. *LIINC rev*. 2012;8(2):323-47. Disponível em: <http://revista.ibict.br/liinc/index.php/liinc/article/view/502>
12. Instituto Brasileiro de Informação em Ciência e Tecnologia. SEER. [Internet]. Brasília: IBICT [acesso em 2014 may5]. Disponível em: <http://www.ibict.br/pesquisa-desenvolvimento-tecnologico-e-inovacao/sistema-eletronico-de-editoracao-de-revistas-seer>
13. Instituto Brasileiro de Informação em Ciência e Tecnologia. BDTD [Internet]. Brasília: IBICT [acesso em 2014 apr2]. Disponível em: <http://bdttd.ibict.br/a-bdttd>
14. Instituto Brasileiro de Informação em Ciência e Tecnologia. Diretório de Políticas de Acesso Aberto das Revistas Científicas Brasileiras (Diadorim) [Internet]. Brasília: IBICT [acesso em 2014 may3]. Disponível em: <http://www.ibict.br/informacao-para-ciencia-tecnologia-e-inovacao%20/diadorim-diretorio-de-politicas-de-acesso-aberto-das-revistas-cientificas-brasileiras>
15. Instituto Brasileiro de Informação em Ciência e Tecnologia. Incubadora de revista INSEER [Internet]. Brasília: IBICT [acesso em 2014 may 2]. Disponível em: <http://www.ibict.br/informacao-para-ciencia-tecnologia-e-inovacao%20/inseer-incubadora-de-revistas-seer-%20%20%28inseer%29>
16. Sayão LF, Sales LF. Dados de pesquisa: contribuição para o estabelecimento de um modelo de curadoria digital para o país. *TPBCI*. 2013; 6(1):1-26.
17. National Science Board. Long-lived digital data collections: enabling research and education in the 21st century. National Science Foundation, sept. 2005. Disponível em: <http://www.nsf.gov/pubs/2005/nsbo540/nsbo540.pdf>
18. Sales LF, Sayão LF. O impacto da curadoria digital dos dados de pesquisa na Comunicação Científica. *Enc Bibli*. 2012;17:118-35. ong-lived digital data coll Fo
19. Higgins, S. Digital curation: the emergence of a new discipline. *The International Journal of Digital Curation*, v.6, n. 2, 2011. Disponível em: <http://www.ijdc.net/index.php/ijdc/article/view/184>.
20. Conway, E. et al. Curating scientific research data for the long term: a preservation analysis method in context. *The International Journal of Digital Curation*, n. 2, v.6, 2011.
21. Oliveira JN. Acesso livre e direito do autor: a comunicação científica eletrônica na Ciência da Informação no Brasil e em Portugal [tese]. Rio de Janeiro: IBICT-UFRJ; 2013.
22. Moura LT. Práticas de comunicação científica: estudo exploratório a partir da Escola de Tradutores de Toledo, séculos XII e XIII [tese]. Rio de Janeiro: IBICT-UFRJ; 2013.

23. Moraes JN. Curadoria e ação interdisciplinar em museus: a dimensão comunicacional e informacional de exposições. In: 12th ENANCIB Encontro Nacional de Pesquisa em Ciência da Informação; 2011, Brasília (DF).

Disponível em:

<http://repositorios.questoesemrede.uff.br/repositorios/bitstream/handle/123456789/2073/Curadoria%20-%20Moraes.pdf?sequence=1>