

O papel das Bibliotecas Digitais de literatura científica cinzenta – os repositórios de *eprints* – na comunicação científica

Ana Maria Ramalho Correia
ISEGI - Instituto Superior de Estatística e Gestão de Informação
UNL – Universidade Nova de Lisboa
Campus de Campolide, 1070 – 124 Lisboa, Portugal
acorreia@isegi.unl.pt

Resumo. As bibliotecas digitais/repositórios digitais de literatura científica cinzenta – de *preprints* e *eprints* – interoperáveis, na *WWW (World Wide Web)* – criados com o objectivo de proporcionar novas formas de difusão do conhecimento original, produzido nas actividades de I&D e facilitar o debate científico entre pares, tem impulsionado transformações assinaláveis na comunicação científica, com significativo destaque nos domínios onde a rapidez e a eficiência na divulgação de resultados constitui factor crítico. Neste artigo são brevemente analisadas as transformações em curso na comunicação científica, nomeadamente as **decorrentes da** intensificação do uso das tecnologias de informação e das redes de comunicação, em particular, na publicação científica electrónica. São abordados **impactos** do repositórios de *eprints* na gestão de literatura científica cinzenta, em formato digital, dando conta de alguns dos mais importantes projectos em curso, neste domínio.

Palavras chave: comunicação científica, bibliotecas digitais, publicação periódica electrónica, arquivos, repositórios, preprints, eprints, LIZA, OAI.

1. Introdução

A publicação de documentos científicos – designadamente, artigos em publicações periódicas científicas e relatórios de investigação – tem como principal finalidade a difusão de resultados de I&D junto da comunidade de investigação, não visando os seus autores a obtenção de quaisquer lucros decorrentes do pagamento de *royalties* ou da comercialização de cópias/exemplares dos mesmos, objectivos prosseguidos na produção de publicações com fins lucrativos.

Diversos factores têm suscitado críticas por parte dos investigadores e outros actores, relativamente ao papel da publicação periódica científica como veículo privilegiado para a comunicação científica, designadamente:

- A considerável morosidade entre a submissão/apresentação de trabalhos para publicação e o seu efectivo aparecimento na literatura, o que constitui uma barreira à partilha eficiente de informação, revelando incapacidade do sistema em responder à aceleração no aumento do volume de novos conhecimentos científicos produzidos, em resultado da expansão, à escala global da actividade de I&D;
- O modelo tradicional de transferência dos direitos dos autores de artigos publicados para as editoras de periódicos científicos impede os autores de promover a difusão, tão alargada quanto possível, dos resultados de investigação alcançados e obter reconhecimento científico entre os respectivos pares [Bacharach 1998; Okerson 2001];
- O modelo actual de implementação de processo de peer-review – uma etapa fundamental na validação da informação científica – é demasiado rígido e muitas vezes funciona como “repressor” ou mesmo “supressor” do aparecimento de novas ideias, por favorecer a publicação de artigos de autores inseridos em instituições prestigiadas, preterindo os restantes, originando atrasos indesejáveis, na publicação de resultados de investigação [Harnard 1998; Harnard 2000; Odlyzko 1995];
- O agravamento do desequilíbrio entre os preços de assinaturas de publicações periódicas, estabelecidos pelos editores – e que desde há mais de uma década crescem a valores superiores às médias da inflação – e os orçamentos das bibliotecas universitárias e de investigação, mesmo das mais prestigiadas, os quais se mantêm estáveis ou mesmo encontram-se em regressão à escala mundial [Tenopir e King 2000: 32; Walker 1998].

É neste cenário que deve ser enquadrada preocupação em desenvolver novas formas para a comunicação formal e informal entre cientistas e o aparecimento de abordagens inovadoras para acesso a resultados da I&D.

Os repositórios de literatura científica cinzenta, em formato digital – eprints – o primeiro dos quais foi criado em 1991 por Ginsparg no Los Alamos National Laboratory [Luce 2001] – desempenham já um papel privilegiado no âmbito da comunicação científica em disciplinas como a Matemática, a Física, a Astronomia e a Economia.

Estes repositórios, cujo aparecimento foi suscitado pelo desenvolvimento da Internet e World Wide Web (WWW), facultam formas inovadoras de acesso à informação científica e técnica. Consistem em arquivos digitais incluindo tanto versões electrónicas preliminares de documentos científicos como também, nalguns casos, documentos já aceites para publicação na sequência de aprovação pelo processo de peer review e, ainda, anotações e actualizações dos mesmos, preparadas pelos autores e/ou comentários pelos pares. Tratam-se de serviços de Bibliotecas Digitais desenvolvidos na WWW por cientistas e para cientistas, que se encontram no cerne das alterações nos hábitos de comunicação científica, uma vez que se assumem crescentemente como forums privilegiados para a difusão de resultados e de debate científico situando-se, por isso, entre a comunicação formal e informal [Kling e McKim 1999]. No entanto, a difusão rápida de informação que os mesmos proporcionam não visa colidir com o sistema tradicional de peer review do sistema tradicional de publicação periódica científica, essencial à constituição da base universal de conhecimento validado.

Nesta comunicação introduzem-se, sucintamente, as mutações em curso na comunicação científica. Analisa-se o papel e o impacto dos repositórios de literatura científica cinzenta, em formato digital, na transformação da comunicação científica referindo-se a aceitação dos mesmos, no contexto de diversas disciplinas científicas.

Abordam-se iniciativas em curso com vista a promover e facilitar o acesso a repositórios de eprints. Conclui-se tecendo algumas considerações sobre o respectivo impacto no futuro da comunicação científica, na Sociedade do Conhecimento.

2. A comunicação científica

As reuniões de investigadores (Invisible College) promovidas regularmente a partir dos anos 1640, sob os auspícios da Royal Society, com o intuito de divulgar experiências, apresentar artigos e anunciar resultados de investigação, constituem os primórdios da comunicação científica formal [Schauder 1994: 75; Meadows 1998]. A permuta de informação que tinha lugar no decurso destes encontros era complementada ainda, por troca de correspondência com vista a envolver no debate aqueles que não tinham tido oportunidade de participar presencialmente nas reuniões. Com o incremento do número de cartas trocadas, surgiu o “periódico científico” como meio eficaz de difundir informação, em escala mais ampla. Os primeiros periódicos científicos (ou journals) – o Journal des Sçavans (Paris) e as Philosophical Transactions of the Royal Society of London – ambos iniciados em 1665, foram publicados por sociedades científicas [Meadows 1998].

Na sua fase inicial, os periódicos científicos consistiam na compilação das cartas trocadas entre investigadores, constituindo também o arquivo dos resultados e das observações feitas no decurso das experiências realizadas. Naturalmente, os periódicos científicos passaram a ter associada a função de registo da “propriedade” da descoberta científica e de apresentação de informação na forma de um arquivo indexado [Schauder, 1994]. Actualmente, o periódico científico constitui parte integrante do sistema de reconhecimento científico [Cronin e Overfelt 1995]. A publicação de trabalhos científicos em periódicos concede visibilidade e contribui para o prestígio dos autores; os benefícios daí decorrentes traduzem-se na promoção na respectiva carreira e/ou na obtenção de financiamentos para prosseguir a actividade de investigação. Os artigos publicados em periódicos científicos são, geralmente, validados através de um processo de revisão anónimo por pares (especialistas), seleccionados pelo editor científico do jornal (*peer review*) [Harnard 2000]. Nesta conformidade, o periódico científico passou a constituir também, um meio de pre-selecção no processo da preservação e de produção do arquivo da literatura científica.

Actualmente, o periódico científico é um produto com valor comercial decorrendo este, fundamentalmente, do facto de o mesmo constituir o repositório de conhecimento validado, das várias disciplinas científicas.

A indústria da publicação científica

A expansão, sobretudo a partir de meados dos anos 1940, da actividade de científica – resultante, em larga medida, da resposta dos EUA ao lançamento do Sputnik, à Guerra Fria e ao crescimento industrial no mundo ocidental, no pós guerra – e o conseqüente aumento do número de artigos produzidos, passaram a exigir a profissionalização dos editores, por forma a assegurar que a selecção, processamento e organização da informação contribuíssem eficiente e eficazmente para o acréscimo efectivo de valor ao trabalho científico [Day; Rowland 1996]. Neste cenário, desponta a indústria de publicação científica, cuja finalidade consiste em facilitar a comunicação, à escala global, entre investigadores, cientistas e docentes. A produção de publicações periódicas científicas, que na sua fase inicial foi assegurada quase exclusivamente por sociedades científicas e por editoras universitárias começou, a partir de meados de 1940, a atrair empresas editoriais com fins lucrativos. O crescimento acentuado da investigação científica entretanto ocorrido a nível internacional, nas áreas de Ciência, Tecnologia e Medicina, criou condições adequadas para que alguns títulos de periódicos, naqueles domínios, até se tornassem numa actividade rentável e particularmente lucrativa. Paralelamente, o incremento da especialização científica determinou a proliferação de novos títulos, muitas vezes de tiragens reduzidas e com custos elevados [Walker 1998].

O mercado da publicação científica apresenta, no entanto, características muito específicas. Por um lado, os editores científicos encontram-se praticamente dependentes dos investigadores, no que se refere à autoria/produção de trabalhos para publicação e, por outro, são as bibliotecas que adquirem a maior parte do “produto final”.

Com a redução e/ou estagnação dos orçamentos das bibliotecas – universitárias e especializadas – para aquisição de informação, os editores de publicações científicas passaram a ter que competir por recursos cada vez mais escassos em orçamentos reduzidos [Odlyzko 1995/1996]. O mercado da publicação científica continua, no entanto, a ser fundamentalmente, impulsionado pelos autores, tendo em conta o papel que aquela desempenha na progressão da carreira científica e/ou docente dos autores.

Tensões entre os vários actores, como alavancas da transformação

O aparecimento de tensões entre os vários actores neste mercado – os investigadores e docentes enquanto autores, os editores, as bibliotecas, os investigadores e docentes na qualidade de leitores/consumidores da publicação científica e as entidades financiadoras de investigação – tem contribuído para transformações muito profundas no sistema de comunicação científica [Odlyzko 1997/1998; Oppenheim 2000: 363].

Enquanto os autores se encontram fortemente pressionados para publicar, com vista a obter reconhecimento para efeitos de progressão científica e académica – *publish or perish* – as bibliotecas – universitárias e especializadas – pelas razões anteriormente aduzidas, não dispõem de capacidade financeira para adquirir “tudo” o que é produzido. Gera - se, deste modo, uma situação onde as forças que motivam e impulsionam a produção de publicações científicas são largamente suplantadas pela capacidade dos potenciais compradores [Morton 1997].

Harnard [1995] corrobora esta constatação ao afirmar que os interesses dos cientistas enquanto autores, por um lado, e os dos editores, por outro, são de sinal contrário dado os cientistas pretenderem a difusão livre dos trabalhos que produzem, já que é por esta via que será produzido conhecimento novo e obtém reconhecimento. No entanto, esta posição dos investigadores é oposta à dos editores cujo objectivo consiste em restringir o fluxo de informação apenas àqueles que pagaram para a adquirir. Os editores, contrariamente aos investigadores, não estão preocupados com a difusão do conhecimento na sociedade, investindo naquela disseminação exclusivamente para obter benefícios financeiros, tratando, deste modo, as publicações periódicas como um produto comercializável (*commodities*).

Harnard [1995] introduz a designação “*faustian bargain*” (“pechincha faustiana”) ao referir-se à situação que obriga os autores a depender dos editores como a única forma de serem lidos por colegas e pares e alcançarem a almejada difusão dos resultados obtidos na respectiva actividade de investigação.

Outros ainda argumentam que a informação produzida no contexto da actividade de I&D constitui um “bem público”, já que a mesma foi criada, na maioria dos casos, à custa de financiamentos de entidades governamentais, os quais são obtidos através dos impostos suportados pelos cidadãos, pelo que deverão ser exploradas formas de reduzir as barreiras de acesso à mesma, por todos os que a procuram para desenvolver a sua actividade de I&D, na Educação ou, ainda, na Sociedade em geral. Nesta conformidade, os custos de acesso à mesma deverão ser minimizados [Odlyzko 1997/1998].

Crescimento dos preços das publicações periódicas científicas

King e Tenopir [1998] dão conta dos vários componentes dos custos da produção das publicações periódicas científicas, os quais incluem:

- i) processamento dos artigos – engloba, designadamente, a recepção do artigo, a decisão inicial em enviar aos *referees*, a selecção destes, o processo de revisão, a preparação dos gráficos, a formatação, a edição da cópia, o envio ao autor para aprovação e a preparação da primeira cópia (*master*);
- ii) processamento de materiais não sujeitos ao processo de revisão científica (*refereeing*);
- iii) produção dos fascículos - onde se abrange a impressão e a colocação de capas;
- iv) distribuição;
- v) apoio à edição abarcando, nomeadamente, os serviços administrativos e financeiros, de marketing e promoção e a gestão de direitos de autor.

É neste contexto que deverá ser enquadrada a dinamização de iniciativas e programas, a partir da década de 80, tanto da parte de editoras comerciais, como da *British Library* (Reino Unido), e da *National Science Foundation* (EUA), com vista a adquirir experiência e a testar a viabilidade da publicação científica periódica, em versão electrónica [Tenopir e King 2000: 21]. Mais recentemente, programas vocacionados para o desenvolvimento da Biblioteca Digital e de seus componentes, como o *e-Lib* (*Electronic Library*), financiado pelo *Joint Information Systems Committee* [JISC], no Reino Unido e

IST - Information Society Technologies (KAIII – Multimedia Content and Tools) [IST] devotaram recursos à dinamização e consolidação, da Publicação Electrónica Científica facilitando a criação de documentos com características inovadoras e de assinalado “valor acrescentado”, face à versão impressa [Hitchcock *et al.* 2000].

No entanto, apesar da crescente utilização de tecnologias de publicação electrónica pelas editoras científicas, os preços dos periódicos científicos continuam a aumentar a ritmos muito superiores aos da média da inflação [Tenopir e King 2000; Walker 1998]. Dados coligidos pela *Association of Research Libraries* apontam para crescimentos daqueles da ordem dos 207%, no período de 1986 a 1999 [ARL 2000].

A literatura da especialidade oferece diversas justificações para este facto e respectivas consequências na comunicação do conhecimento científico [Tenopir e King 2000: 32]. Alguns reclamam que os custos de produção da primeira cópia (*master*) obrigam os editores a manter os preços elevados [Oppenheim *et al.* 2000], já que a introdução da tecnologia de publicação electrónica não reduzirá o contributo intelectual exigido nesta fase. Outros [Odlyzko 1997/1998] argumentam que os preços aumentaram apenas em resultado do incremento da actividade de I&D e do aumento do volume de artigos publicados. Outros ainda, que os editores necessitam de amplas margens de lucro para reinvestirem na criação de novos produtos electrónicos [Oppenheim *et al.* 2000: 364]. Por outro lado, alguns atribuem exclusivamente o incremento tão acentuado dos preços dos periódicos ao domínio quase exclusivo do mercado da publicação periódica científica por editoras comerciais com fins lucrativos e ao funcionamento ineficiente do mesmo [Newsletter on Serials Pricing Issues].

Esta situação, que cria barreiras ao fluxo da comunicação científica, tem sido crescentemente criticada pelos investigadores, instituições universitárias e utilizadores em geral, tanto mais que, os autores – investigadores e docentes das universidades – transferem para as editoras os direitos de propriedade intelectual dos respectivos trabalhos. Por seu turno, são estes que também desempenham a função de *peer review* e que integram os corpos editoriais, na maioria das situações, a título gratuito. Posteriormente, os editores comercializam os resultados da investigação científica às bibliotecas universitárias a preços elevados.

Outro aspecto que também preocupa os investigadores, já referido anteriormente, reside nos prazos extensos que decorrem entre a entrega (submissão) de artigos e a publicação dos mesmos.

Este cenário leva-os a questionarem-se quanto à eficiência da publicação periódica científica, como veículo privilegiado de comunicação, designadamente, nas áreas científica, tecnológica e da medicina, onde a procura de informação, sempre actualizada, é enorme.

Transformação da comunicação científica

Garvey e Griffith [1972] desenvolveram um modelo de comunicação científica que reflecte os comportamentos de comunicação formal e informal da comunidade de psicólogos [Hurd 1996]. Segundo aqueles autores, o modelo era válido não só para as Ciências Sociais como também para as Ciências Exactas. O modelo ilustrava o processo pelo qual a informação é comunicada, detalhando as várias etapas que medeiam entre a criação de um conceito inicial até que o mesmo seja inserido na base de conhecimento global. Assim, a informação científica uma vez criada, é comunicada de várias maneiras e em diversos formatos; em particular, as publicações periódicas (jornais científicos) referendadas (através de *peer review*), desempenham um papel fundamental, por acrescentarem valor ao processo da comunicação científica, através da validação dos resultados (controlo de qualidade), para além de se constituírem em repositórios do conhecimento acumulado e de promoverem a comunicação [Schauder 1994].

Desde que Garvey e Griffith apresentaram este modelo, ocorreram, como se referiu anteriormente, transformações muito significativas no processo de comunicação científica, em particular, as suscitadas pela rápida expansão das tecnologias de informação e das redes de computadores, a consolidação da publicação electrónica, o desenvolvimento de bibliotecas digitais, as quais, de uma forma geral, contribuíram, para a reestruturação da publicação científica, das bibliotecas académicas e de investigação, promoveram mudanças nos métodos de acesso a recursos de informação, suscitaram o debate sobre as práticas para a gestão dos “direitos de autor” [Okerson 2001], conduzindo em suma ao aparecimento de tensões no relacionamento entre autores, bibliotecas (intermediários), editores e utilizadores.

O primeiro modelo de comunicação científica por via electrónica foi proposto por Lancaster [1978]; mais recentemente [Hurd 1996; Hurd 2000] actualizou-o, considerando que o acesso por via electrónica à informação científica irá desempenhar, no futuro, um papel central. Nesta conformidade, a abordagem linear ao acesso à informação, representada no modelo de Garvey e Griffith, modifica-se pela introdução de novas ferramentas electrónicas que suportam a produção, o armazenamento e a utilização de informação, a diversos níveis, por diversas comunidades.

Por outro lado, a Internet e a *World Wide Web* trouxeram naturalmente novas oportunidades para a comunicação científica e estão a catalisar alterações muito profundas na forma como os cientistas trabalham e comunicam informação entre si. Neste contexto, as fronteiras entre literatura formal e informal começam a esbater-se, resultando num *continuum* de trabalho colaborativo e interactivo [Kling e McKim 1999; Borgman 2000].

Nesta perspectiva, os progressos tecnológicos constituem-se também como força motora no cerne das transformações na comunicação científica [Ginsparg 1996; Harnard 1995; Odlyzko 1995/1996], embora estas ocorram a ritmos e de formas diferentes consoante a disciplina científica [Hurd 2000].

É neste cenário que se enquadra, a partir do início da década de 90, a criação em organismos de investigação e departamentos universitários, de repositórios na *WWW* para publicação, por iniciativa dos autores (*self-publishing*), de versões preliminares de literatura científica cinzenta (*preprints*), denominados repositórios/arquivos de *eprints*.

No entanto, se a aplicação das tecnologias da Internet traz inúmeras vantagens à comunicação científica, tendo em conta a rapidez e facilidade em difundir informação à escala global, não pode, contudo, deixar de ser feita a distinção entre o que é o debate académico e o que constitui efectivamente publicação científica. O papel e os valores associados à publicação científica e, em particular, à publicação periódica, permanecem – validação, reconhecimento, arquivo, definição de prioridade, autoria e obtenção de prestígio [Schauder 1994]. No entanto, a comunicação científica propiciada pelos repositórios de *eprints* facilita e acelera a partilha de conhecimento sem restrições e a libertação da dependência dos editores, por parte dos autores, na distribuição dos resultados dos seus trabalhos.

3. Repositórios/arquivos de *eprints*

Como se referiu anteriormente, a criação de repositórios/arquivos de versões electrónicas de documentos científicos disponíveis na *WWW* – repositórios de *eprints* – encontra-se a impulsionar grandes mutações na comunicação científica [Holtkamp e Berg 2001]. O primeiro a surgir foi o do *Los Alamos Physics Archive*, actualmente designado *arXiv.org* criado em 1991, por Ginsparg, o qual evoluiu convertendo-se num repositório de trabalhos científicos, não referendados, nos mais diversos domínios da Física, Matemática, Ciências da Computação, produzidos à escala global [Ginsparg 1996; Luce 2001; McKiernan 2000]. O sucesso deste novo meio de comunicação determinou que, ao longo da última década, o mesmo tenha sido replicado com a criação de mais de uma dúzia de *mirrors* em vários países do mundo.

O nível de utilização do *arXiv.org* e a sua importância na comunicação científica entre investigadores das áreas que cobre, podem ser apreciados pelos seguintes indicadores:

- No início de 1999, o *arXiv.org*, continha cerca de 100.000 artigos e o número destes descarregados (*downloaded*), anualmente, ascendia a cerca de 7 milhões, o que revela que, em média, cada artigo foi “utilizado” pelo menos 70 vezes (valores referentes, apenas, ao servidor em *Los Alamos*, excluindo portanto a utilização através dos *mirrors* e a que consiste apenas na visualização só para consulta em versão electrónica) [Odlyzko 2001: 14];
- O número de acessos ao *arXi.orgv* tem aumentado exponencialmente - os acessos semanais que foram cerca de 100.000 em 1994, atingiram, na primeira semana de Maio de 2000, o valor de 800.000 [Brown 2001].

Embora diversos autores tenham questionado a aceitação de repositórios de *eprints* noutras áreas científicas não tão vocacionadas para exploração dos avanços tecnológicos na comunicação de informação, como sucede nas áreas da Física [Kling e McKim 2000], o que é facto é que, entretanto, têm surgido muitos outros cobrindo diferentes domínios, salientando-se os seguintes (Quadro 1):

- *CogPrints*, na Universidade de Southampton, disponibilizando literatura na área da Psicologia, das Neurociências e da Linguística;
- *RePEc* – iniciativa com vista à disponibilização da literatura cinzenta, em formato digital, na área da Economia; presentemente faculto o acesso a mais de 140.000 *items* (*working papers*, artigos de periódicos, *software*, listas de contactos de investigadores, listas de endereço de departamentos), existindo em diversos *servers*, a nível mundial [Krichel 2000];
- *Clinmed.NetPrints* - disponibilizando o acesso a *eprints* na área da Medicina e Ciências da Saúde e desenvolvido pelo *British Medical Journal* e *HighWire Press* [Delamothe, *et al.* 1999];
- *AgEcon – Research in Agricultural and Applied Economics* – desenvolvido pela *University of Minnesota*.

Designação	URL
<i>AgEcon Search: Research in Agriculture and Applied Economics</i>	http://agecon.lib.umn.edu/
<i>ArXiv.org e-Print archive</i>	http://arXiv.org
<i>Clinmed.NetPrints.org: Clinical Medicine & Health Research (British Medical Journal e HighWire Press)</i>	http://clinmed.netprints.org/home.dtl
<i>CogPrints – Cognitive Sciences Eprint Archive</i>	http://cogprints.soton.ac.uk
<i>ERA – Electronic Research Archive (The Lancet)</i>	http://www.thelancet.com/era
<i>MATH NET – Internet Information Services for Mathematicians.</i>	http://www.math-net.de/
<i>PhysNet: The Physics Departments & Documents Network</i>	http://physnet.uni-oldenburg.de/PhysNet/
<i>PrePRINT Network</i>	http://www.osti.gov/preprint/
<i>RePEc – Research Papers in Economics.</i>	http://www.repec.org/
<i>The Mathematics Preprint Server. (Elsevier)</i>	http://www.mathpreprints.com/math/Preprint/show/

QUADRO 1 – URLs de Repositórios/Arquivos de *eprints* e serviços de acesso referidos [* Conferidos em: 27 Jun. 2001].

Reflectindo a importância crescente que os repositórios de *eprints* assumem, no contexto da comunicação científica, há a mencionar, por outro lado, o desenvolvimento de novos serviços, também baseados em tecnologias da Internet e que facilitam a localização e o acesso a repositórios de *eprints* (portais/*gateways*). Salienta-se, em particular: *PrePRINT Network PhysNet; MathNet* (Quadro 1).

4. *Open Archives Initiative* – para a interoperação de repositórios/arquivos de literatura cinzenta científica

O objectivo da criação dos repositórios de *eprints* consiste em desenvolver plataformas para a efectiva comunicação de informação científica, recorrendo às potencialidades da publicação electrónica na *WWW*.

No entanto, as abordagens implementadas pelos vários repositórios são diferentes entre si, o que pode constituir um entrave à respectiva interoperabilidade e eficiência no fluxo de informação, à escala global. Assim, alguns desenvolvem-se segundo um modelo centralizado, outros promovem um modelo distribuído, a nível de instituição ou de disciplina. Uns tratam literatura científica cinzenta, exclusivamente (material que ainda não foi objecto de *peer review*), enquanto outros disponibilizam, também, metadata referente às versões já validadas (*peer-reviewed*) ou procuram estabelecer um sistema de validação à margem do sistema de *peer-review* – comentário por pares, por exemplo [Harnard 1998]. Alguns repositórios incluem apenas metadata, enquanto outros disponibilizam, para além desta, o texto integral do documento [Sompele e Lagoze 2000].

Com vista a incrementar o impacto dos arquivos de *eprints* existentes e de promover a respectiva interoperabilidade, por forma a que os mesmos se constituam como alternativas viáveis no contexto da comunicação científica à escala global, foi lançada a *Open Archives Initiative* – *OAI* - numa reunião promovida em Outubro de 1999, em Santa Fe, New Mexico, sob os auspícios do *Council on Library and Information Resources* (CLIR), o *Digital Library Federation* (DLF), o *Scholarly Publishing and Academic Resources Coalition* (SPARC), a *Association of Research Libraries* (ARL), a *Research Library*, do *Los Alamos National Laboratory* [OAI].

A finalidade da *Open Archives Initiative* consiste em promover mecanismos técnicos e estruturas organizativas que assegurem a interoperação dos repositórios de *eprints*, facilitando a conversão dos vários sistemas de *eprints* existentes actualmente, em elementos constituintes de um modelo inovador de comunicação científica, à escala global [Sompele and Lagoze 1999; Heery 2001].

A *Santa Fe Convention* – designação pela qual são conhecidos aqueles acordos – consiste num conjunto de acordos simples, mas que permitem um elevado grau de interoperacionalidade entre os vários arquivos de *eprints*, através do *metadata harvesting*. Estes acordos são um conjunto de princípios organizacionais e de especificações técnicas. A Convenção estabelece para os *fornecedores de dados* (*data providers*) – os arquivos individuais – mecanismos relativamente simples para tornar o metadata acessível do exterior. A existência deste metadata permite, aos *fornecedores de serviços* (*service providers*), construir funcionalidades de nível superior – os serviços intermediários (como catálogos ou portais para acesso a materiais distribuídos em diversos arquivos de *eprints*) – através da utilização do metadata disponibilizado nos arquivos desenvolvidos, de acordo com o estabelecido na referida Convenção. Esses serviços podem combinar/cruzar e processar informação proveniente dos vários arquivos e oferecer funcionalidades acrescidas para apoio à descoberta, apresentação e análise de dados originários de arquivos compatíveis.

O conceito desenvolvido no âmbito do *OAI* tem vindo, a suscitar a atenção como modelo adequado para aplicação a outros tipos de recursos digitais com interesse científico ou relevantes para o ensino superior. Para além dos *eprints* e dos textos electrónicos, incluem-se colecções de dados, materiais visuais, arquivos, dados em sistemas de informação geográfica, som, música e vídeo. A lista de repositórios desenvolvidos de acordo com as especificações do *OAI* ascende, presentemente (Junho 2001), a mais de três dezenas [OAI. *The...*].

5. Benefícios e impacto dos repositórios/arquivos de *eprints*

Embora os repositórios de *eprints* não visem substituir a publicação científica periódica por não desempenharem a globalidade dos papéis e valores que estão associados àquele, os quais foram aperfeiçoados ao longo de mais de três séculos de utilização [Schauder 1994], sem dúvida que apresentam características inovadoras, que estão a contribuir para alterar a forma como os investigadores comunicam com os seus pares e com a sociedade em geral.

Os repositórios de *eprints* propiciam a difusão rápida e automática, com base nas tecnologias da Internet, de documentos, facilitando a divulgação rápida de resultados de I&D e promovem o desenvolvimento de comunidades de pares [Boyce 2000]. Por outro lado, permitem a interligação na *WWW* com as referências

e citações incluídas no texto dos documentos, bem como a disponibilização de material multimedia, designadamente simulações, modelos a 3D, compilações extensas de dados, relevantes no âmbito da investigação científica descrita [Hitchcock *et al.* 2000].

É de esperar, no entanto, que a aceitação dos repositórios de *eprints* não ocorra, de forma uniforme, em todas as disciplinas. Diversos estudos sobre os comportamentos de procura de informação por cientistas demonstraram existirem diferenças entre as diversas áreas científicas. Em disciplinas como a Física, a Astronomia, Matemática e a Economia predomina uma cultura de difusão de *preprints* para disseminação rápida de versões preliminares de trabalhos, anteriormente à respectiva publicação. No entanto, em especialidades em que as patentes de produtos e processos constituam a norma (ex. Química e Farmacologia), a permuta de *preprints* não se encontra implantada [Hurd 2000:1282]. Nesta conformidade, é de esperar que idênticas diferenças se irão reflectir na aceitação e criação de repositórios de *eprints* [Kling e McKim 2000].

Odlyzko [2001:9] salienta ainda o efeito dos repositórios de *eprints* na alteração dos padrões da comunicação científica, decorrente da redução das barreiras de acesso à mesma. Este autor considera que é de prever o incremento da utilização de artigos disponibilizados em repositórios de *eprints*, por utilizadores de áreas científicas exteriores ao domínio científico específico a que os mesmos respeitam, sendo de admitir que a consequente redução das barreiras no acesso a este material em versão electrónica, promova a interdisciplinaridade, entre áreas que, de outra forma, não teriam meios para descobrir oportunidades para trabalho em colaboração.

O impacto dos repositórios de *eprints*, na comunicação científica, pode ser apreciado segundo duas vertentes:

i) Citação de trabalhos publicados em repositórios de *eprints* em periódicos científicos

Brown [2001] constatou o incremento da importância de *eprints* na comunicação científica em Física e Astronomia, através de um estudo recentemente realizado, sobre a citação de trabalhos existentes em repositórios de *eprints*, por artigos publicados em publicações periódicas científicas. Os dados recolhidos evidenciam que tem aumentado o número de citações a *eprints*, num número considerável dos mais prestigiados periódicos nos domínios da Física e da Astronomia [Brown 2001: 196]. No mesmo trabalho, também foram pesquisadas as atitudes dos editores científicos (de cerca de cinquenta títulos seleccionados entre os mais prestigiados nos domínios científicos em apreço designadamente, *American Institute of Physics – AIP*, *American Physical Society*; *Institute of Physics - IoP*, *American Astronomical Society*) com respeito à inclusão, em artigos aceites para publicação, de citações a *eprints*. Aquela autora conclui que, embora as políticas variem ainda desde uma atitude de resignação (“porque não?”) até à de preocupação pelo facto dos *eprints* não terem sido objecto de *peer review* para validação da qualidade, o que é facto é que o número de títulos que aceita citações a *eprints* tem aumentado significativamente, nos últimos anos.

ii) Contributo dos repositórios de *eprints* para o aumento da visibilidade e impacto dos mesmos

Lawrence [2001], do *NEC Research Institute*, em estudo realizado no âmbito das Ciências da Computação, demonstrou que os artigos que se encontram acessíveis gratuitamente na *WWW* são citados com muito maior frequência que os restantes, pelo que a inclusão de documentos em repositórios de *eprints* apresenta também vantagens em termos do aumento de visibilidade e impacto dos mesmos, na base global de conhecimento (nº de citações).

6. Conclusão

Como o afirma Boyce [2000], os repositórios de *eprints* “vieram para ficar”. A sua função não será a de substituir a publicação periódica, mas a de a complementar, tirando partido das tecnologias da Internet e *WWW*, para promover a rápida e eficiente difusão de resultados de investigação e criar *fora* de debate alargado, entre comunidades de cientistas dispersos à escala global, ao mesmo tempo que facilitam a disponibilização de resultados de I&D, que apenas possam ser distribuídos em formato electrónico.

As possibilidades trazidas pela *Open Archive Initiative*, com vista a promover a interoperacionalidade de repositórios/arquivos de *eprints* desenvolvidos de acordo com este enquadramento tecnológico, proporcionarão a implementação de uma base global do conhecimento científico disponível,

gratuitamente, na WWW, facultando a interdisciplinaridade e concedendo visibilidade a trabalhos científicos, à escala global.

Finalmente, existem vários indícios que apontam no sentido dos editores passarem a aceitar, crescentemente, esta nova forma de comunicação entre cientistas, por permitirem a submissão de trabalhos que previamente tenham sido disponibilizados em repositórios de *eprints*, como também por alterarem as regras de transferência de direitos de autor autorizando a manutenção nesses repositórios de versões preliminares (*eprints*) de trabalhos que já se encontram publicados em publicações periódicas. Ilustrativo desta tendência são algumas contribuições no debate promovido pela *Nature* sobre *Future e-access to primary literature*, em curso desde Abril de 2001 [Butler 2001], bem como o aparecimento de novos repositórios de *eprints* implementados por editoras comerciais – *ERA - Electronic Research Archive*, da *The Lancet* e o *The Mathematics Preprint Server*, da Elsevier.

Do exposto decorre, ainda, a relevância da iniciativa LIZA, lançada pela Biblioteca Nacional e que visa a criação de um repositório de literatura científica cinzenta, em formato digital, produzida pela comunidade científica nacional, interoperável com iniciativas similares a nível internacional. Este constituirá uma forma de preservar esta parte do património científico nacional e de lhe conceder a adequada visibilidade, ao mesmo tempo que facilitará a aquisição de competências, a nível nacional, na gestão de recursos de informação científica disponível na WWW [Correia e Borbinha 2001; Correia e Neto 2001].

Referências Bibliográficas

- ARL. *Principles for Emerging Systems of Scholarly Publishing*, 2000.
URL: <http://www.arl.org/scomm/tempe.html> [Conferido em: 27 Jun. 2001].
- BACHRACH, S. *et al.* Intellectual property: who should own scientific papers, *Science*, 281 (5382), 1459 – 60.
URL: <http://www.sciencemag.org/cgi/content/full/281/5382/1459> [Conferido em: 27 Jun. 2001].
- BORGMAN, C. Digital Libraries and the continuum of scholarly communication. *Journal of Documentation*, 56, 4 (2000), pp. 412 – 430.
- BOYCE, P. For better or for worse. Preprints are here to stay. *College and Research Libraries News*, 61, 5 (2000), pp. 404 – 407(414).
URL: <http://www.ala.org/acrl/scholcomm.html> [Conferido em: 15 Jun 2001].
- BROWN, C. The E-volution of Preprints in the Scholarly Communication of Physicists and astronomers. *Journal of American Society for Information Science and Technology*, 52, 3 (2001), pp. 187 – 200.
- BUTLER, D. Future e-access to the primary literature. *Nature web debates*, 2001.
URL: <http://www.nature.com/nature/debates/e-access> [Conferido em: 27 Jun. 2001].
- CORREIA, A. M. e BORBINHA, J. Deposit of Scientific and Technical Gray Literature in Portugal: a Case Study. In ISAIAS, P. (ed.). *New Developments in Digital Libraries, Proceedings of the 1st International Workshop on New Developments in Digital Libraries (NDDL-2001)*, in conjunction with *International Conference on Enterprise Information Systems. (ICEIS)*, Setúbal, Portugal, July 2001, pp. 8 – 19.
URL: <http://www.isegi.unl.pt/docentes/acorreia/preprint/TextoSetubalVF.pdf> [Conferido em: 27 Jun. 2001].
- CORREIA, A M. e NETO, M. The role of Eprint archives in the access to an dissemination of scientific gray literature: LIZA – A case study by National Library of Portugal. Paper submitted to the Workshop *Electronic Media in Mathematics*, Coimbra, Set. 2001.
- CRONIN, B. e OVERFELT, K. E-Journals and Tenure. *Journal of the American Society for Information Science*, 46, (9) 1995, pp. 700- 703.
- DAY, C. The economics of electronic publishing: some preliminary thoughts.
URL: <http://arl.cni.org/symp3/day.html> [Conferido em: 27 Jun. 2001].
- DELAMOTHE, T *et al.*. Netprints: the next phase in the evolution of biomedical publishing. *British Medical Journal*, 319 (1999), pp. 1515 – 6.
URL: <http://www.bmj.com/cgi/content/full/319/7224/1515> [Conferido em: 27 Jun. 2001].

- GARVEY, W. D. e GRIFFITH, B. C. Communication and information processing within scientific disciplines: Empirical findings for psychology. *Information Storage and Retrieval*, 8, (1972), pp. 123 – 126. citado em HURD, J. Models of Scientific Communications Systems. In CRAWFORD, S. et al. *From Print to Electronic: The transformation of Scientific Communication*. Medford, NJ: ASIS, 1996, pp. 9 – 33.
- GINSPARG, P. *Winners and losers in the global research village*. Invited Contribution, UNESCO, Paris, 19-23 Feb. 1996.
URL: <http://xxx.lanl.gov/blurb/pg96unesco.html> [Conferido em: 27 Jun. 2001].
- HARNARD, S. Electronic Scholarly Publishing: Quo Vadis? *Serials Review*, 21, 1 (1995), pp. 78 – 80.
URL: <http://www.cogsci.soton.ac.uk/~harnad/Papers/Harnad/harnad95.quo.vadis.html> [Conferido em: 27 Jun. 2001].
- HARNARD, S. Learned Inquiry and the Net: the Role of Peer Review, Peer Commentary and Copyright. *Learned Publishing*, 11, 4 (1998), pp. 183- 192.
URL: <http://www.cogsci.soton.ac.uk/~harnad/Papers/Harnad/harnad98.toronto.learnedpub.html> [Conferido em: 27 Jun. 2001].
- HARNARD, S. The invisible hand of peer review. *Exploit Iterative*, 5, (2000).
URL: <http://www.exploit-lib.org/issue5/peer-review/> [Conferido em: 27 Jun. 2001].
- HEERY, R. OAI Open Meeting. *Cultivate*, 4, (2001).
URL: <http://www.cultivate-int.org/issue4/oai/> [Conferido em: 27 Jun. 2001].
- HITCHCOCK, S. et al. Developing services for open eprint archives: globalisation, integration and the impact of links. *Proceedings of the 5th ACM Conference on Digital Libraries*, San Antonio, Texas, 2000.
URL: <http://www.cogsci.soton.ac.uk/~harnad/Papers/Harnad/harnad00.acm.htm> [Conferido em: 27 Jun. 2001].
- HOLTKAMP, I. e BERG, D. A. *The impact of Paul Ginsparg's ePrint arXiv (formerly known as xxx.lanl.gov) at Los Alamos National Laboratory on Scholarly Communication and Publishing: a selected bibliography*, 2001.
URL: <http://lib-www.lanl.gov/libinfo/preprintsbib.htm> [Conferido em: 27 Jun. 2001].
- HURD, J. Models of Scientific Communications Systems. In CRAWFORD, S. et al. *From Print to Electronic: The Transformation of Scientific Communication*. Medford, N. J. Information Today, 1996, pp. 9-33.
- HURD, J. The Transformation of Scientific Communication: A Model for 2020. *Journal of the American Society for Information Science*, 51, 14 (2000), pp. 1279 – 1283.
- IST - Information Society Technology Programme (Multimedia Content and Tools).
URL: <http://www.cordis.lu/ist/ka3/home.html> [Conferido em: 27 Jun. 2001].
- JISC. *e-Lib - Electronic Libraries Programme*.
URL: <http://www.ukoln.ac.uk/services/elib/> [Conferido em: 27 Jun. 2001].
- KING, D. W. e TENOPIR, C. Economic Cost Models of Scientific Scholarly Journals. In *Economics, Real Costs and Benefits of Electronic Publishing in Science – A Technical Study*. Oxford: Proceedings of ICSU Press Workshop, 1998.
URL: <http://www.bodley.ox.ac.uk/icsu/kingppr.htm> [Conferido em: 27 Jun. 2001].
- KLING, R. e McKIM, G. Not just a matter of Time: Field Differences in the Shaping of Electronic Media in supporting Scientific Communication. *Journal of the American Society for Information Science*, 2000.
URL: <http://xxx.lanl.gov/ftp/cs/papers/9909/9909008.pdf> [Conferido em: 27 Jun. 2001].
- KLING, R. e McKIM, G. Scholarly Communication and the Continuum of Electronic Publishing. *Journal of American Society of Information Science*, 50, 10, (1999), pp. 890 – 906.

- KRICHEL, T. Working towards an Open Library for Economics: The RePEc project. In PEAK. *The Economics and Usage of Digital Library Collections*, Ann Arbor, Mi, 2000.
URL: <http://openlib.org/home/krichel/myers.html> [Conferido em: 27 Jun. 2001].
- LANCASTER, F. W. Towards paperless information systems. London: Academic Press, 1978 citado em HURD, J. The Transformation of Scientific Communication: A Model for 2020. *Journal of the American Society for Information Science*, 51, 14 (2000), pp. 1279 – 1283.
- LAWRENCE, S. Free online availability substantially increases a paper's impact. In *Nature web debates*, 2001.
URL: <http://www.nature.com/nature/debates/e-access/Articles/lawrence.html> [Conferido em: 27 Jun. 2001].
- LUCE, R. E-prints intersect the Digital Library: inside the Los Alamos arXiv. *Issues in Science and Technology Librarianship*, Winter 2001.
URL: <http://www.library.ucsb.edu/istl/01-winter/article3.html> [Conferido em: 27 Jun. 2001].
- McKIERNAN, G. arXiv.org: the Los Alamos National Laboratory e-print server. *International Journal of Grey Literature*, 1, 3 (2000), pp. 127 – 138.
- MEADOWS, A. J. *Communicating Research*. San Diego, Ca: Academic Press, 1998.
- MORTON, B. Is the Journal as We Know It an Article of Faith? An Open Letter to the Faculty. *The Public – Access Computer Systems Review*, 8, 2 (1997), pp. 1-8.
URL: <http://info.lib.uh.edu/pr/v8/n2/mort8n2.html> [Conferido em: 27 Jun. 2001].
- Newsletter on Serials Pricing Issues*.
URL: <http://www.lib.unc.edu/prices/> [Conferido em: 27 Jun. 2001].
- OAI – *The Open Archive initiative: List of OAI - Conformant repositories*.
URL: <http://oaisrv.nsd.cornell.edu/Register/BrowseSites.pl> [Conferido em: 27 Jun. 2001].
- OAI. *The Open Archives Initiative*.
URL: <http://www.openarchives.org> [Conferido em: 27 Jun. 2001].
- ODLYZKO, A M. *The Rapid Evolution of Scholarly Communication*. Aceite para publicação em *Learned Publishing*, e em LOGEE, W. e Mac-KIE MASON, J (eds.). *Bits and Bucks: Economics and Usage of Digital Collections*, MIT Press, 2001.
URL: <http://www.research.att.com/%7Eamo/doc/rapid.evolution.pdf> [Conferido em: 27 Jun. 2001].
- ODLYZKO, A. M. The Economics of Electronic Journals. *First Monday*, 8, 2 (1997) e In EKMAN, R. e QUANDT, R. (eds.). *Technology and Scholarly Communication*. Univ. Calif Press, 1998.
URL: <http://www.research.att.com/~amo/doc/complete.html> [Conferido em 27 Jun. 2001].
- ODLYZKO, A. M. Tragic loss or good riddance? The impending demise of traditional scholarly journals. *J. of Human-Computer Studies*, 42 (1995) e em PEEK, R. P. e NEWBY, G. B. (eds.). *Scholarly publishing: the electronic frontier*. Cambridge, MA: MIT Press, 1996, pp. 91-102 (versão condensada).
URL: <http://www.research.att.com/~amo/doc/complete.html> [Conferido em: 27 Jun. 2001]
- OKERSON, A. Rights, Copyrights and Licenses: the basics and what they mean. In *Informação: o desafio do futuro*. Actas do 7º Congresso Nacional BAD, 2001.
- OPPENHEIM, C *et al.* The future of scholarly journal publishing. *Journal of Documentation*, 56, 4 (2000), pp. 361-398.
- ROWLAND, F. The need for management of electronic journals. In PEEK, R. P. e NEWBY, G. B. (eds.). *Scholarly publishing: the electronic frontier*. Cambridge, MA: MIT Press, 1996, pp. 243 –250.
- SCHAUDER, D. Electronic Publishing of Professional Articles: Attitudes of Academics and Implications for the Scholarly Communication Industry. *Journal of the American Society for Information Science*, 45, 2 (1994). pp. 73-100.

- SEVERINS, T. *et al.* *PhysDoc* – A distributed network of Physics Institutions Documents. *D-Lib Magazine*, 6, 12, (2000).
URL: <http://www.dlib.org/dlib/december00/severiens/12severiens.html> [Conferido em: 27 Jun. 2001].
- SOMPEL, H. e LAGOZE, C. The Santa Fe Convention of Open Archives Initiative. *D-Lib Magazine*, 6, 2 (2000).
URL: <http://www.dlib.org/dlib/february00/vandesompel-oai/02vandesompel-oai.html> [Conferido em: 27 Jun. 2001].
- TENOPIR, C. e KING, D. W. *Towards Electronic Journals: Realities for Scientists and Publishers*. Washington: Special Libraries Association, 2000.
- WALKER, T. Free Internet Access to Traditional Journals, *American Scientist*, 86, 5 (1998).
URL: <http://www.sigmaxi.org/amsci/articles/98articles/walker.html> [Conferido em 27 Jun. 2001].