

# A EUROPEAN MODEL FOR THE DIGITAL PUBLISHING OF SCIENTIFIC INFORMATION?

*Chris Armbruster*  
Research Associate, Max Planck Digital Library,  
Max Planck Society  
Invalidenstrasse 35, D 10115 Berlin  
[www.mpdl.mpg.de](http://www.mpdl.mpg.de)

*Executive Director, Research Network 1989*  
[www.cee-socialscience.net/1989](http://www.cee-socialscience.net/1989)

## **Abstract**

The entire system of scholarly communication is in transition, with the emergence of new markets, services and players. Given what we know about this digital transition, the technical, financial and legal parameters of a future model of publishing scientific information are predictable. The contribution of the Commission of the European Communities has been to emphasise how digital access to scientific information is related to the digital preservation of the record of scientific publications and data; because if digital preservation is undertaken access cannot be guaranteed for future innovations. The proposition for the delineation of a European model is developed as follows:

1. A digital model may be developed that results in open access, while preserving and enhancing the viability of a variety of commercial publishing models.
2. The institutional players in the publishing system, namely publishers, repositories and libraries, must be ready to accept a redistribution of the key functions of registration, certification, dissemination, archiving and navigation in a manner that plays to the strength of each.
3. The condition for any successful elaboration of a digital model is that it is complementary to the technology and economics of the internet, while the litmus test is that it enhances the impact and re-use of scientific information.

Open access publishing leads to complementary relationships between publishers, libraries and repositories; unrestricted access and the widest possible dissemination; and usually facilitates the creation of value-added services as an overlay from platforms, repositories and libraries. In recent years, a number of viable full open access publishing (OAP) models have emerged: OAP where the author pays, by subscription, through a sponsoring consortium, and by way of support and sponsorship.

## **Keywords**

Digital publishing, scientific publishing, open access publishing, permission mandate, scientific information, repositories, digital libraries, peer review

## Table of contents

1. The contribution of the European Commission: an integrated perspective on access and preservation
2. Identifying the technical, legal and financial parameters for the digital publishing of scientific information
3. Redistributing registration, archiving and dissemination in Europe: permanent and open access for scientific information
4. Repository registration and dissemination: strategies in the EU and US
5. Open access publishing: communal, institutional and commercial models
  - OAP where the author pays
  - OAP by subscription
  - OAP through a sponsoring consortium
  - OAP by way of support and sponsorship
6. Non-exclusive licensing: synergies between a new kind of green open access and open access publishing?
7. How will publishers, repositories and libraries share the key functions of registration, certification, dissemination, archiving and navigation?

This proposal for a European model draws, in large parts on my earlier work, which is openly accessible via my author page: <http://ssrn.com/author=434782>. References are therefore kept to a minimum. My work is indebted to an ongoing engagement with the open access policies of the Max Planck Society, from its inception and the set-up of the Centre for Information Management (ZIM, Director: Theresa Velden), to the more recent creation of the Max Planck Digital Library (Director: Laurent Romary). In its present form, the proposal owes much to the interaction with colleagues in the process of writing a proposal for a study on open access publishing. Special thanks go to Peter Igo-Kemenes, Simon Lambert, Salvatore Mele, Neil Sandford and Ralf Schimmer.

“In order to become a truly modern and competitive economy, Member States and the EU must remove barriers to the free movement of knowledge by creating a fifth freedom. This would involve enhancing the cross-border mobility of researchers, as well as students, scientists, and university teaching staff. It would also require facilitating and promoting the optimal use of intellectual property created in public research organisations and encouraging open access to knowledge.”

Council of the European Union, 2852<sup>nd</sup> Competitiveness Council meeting, 25 February 2008 – from the press release<sup>1</sup>

## **1. The contribution of the European Commission: an integrated perspective on access and preservation**

The contribution of the Commission of the European Communities has been to emphasise how digital access to scientific information is related to the digital preservation of the record of scientific publications and data. If digital preservation is not undertaken access cannot be guaranteed for future innovations. Consequently, we have a comprehensive framework of action that incorporates:

- A. The equal importance of publishers, repositories and libraries in developing infrastructures for scientific information;
- B. The importance of access and preservation in ensuring a rapid and wide dissemination of scientific information;
- C. The internal challenge of enhancing the re-use of data and publications and attending to the issue of research evaluation;
- D. The significance of organisational, technical, financial and legal issues in the transition to digital publishing and the re-articulation of the key functions of scientific communication: registration, certification, dissemination, archiving and navigation.

The entire system of scholarly communication is in transition, with emergence of new markets, services and players. Given what we know about this digital transition, the technical, financial and legal parameters of a future model of publishing scientific information are predictable. The European Commission's communication of 2007 remained inconclusive, because of a perceived clash between players aligning in opposing camps.<sup>2</sup> Yet, peer-reviewed open access publishing has experienced several milestones and is now maturing:

- Large, commercial and full open access publishing programmes, of 100 or more journals in open access, are now run by several companies (e.g. BioMed Central, Hindawi, Bentham). This group includes the launch of a

<sup>1</sup> [http://cordis.europa.eu/search/index.cfm?fuseaction=news.document&N\\_RCN=29243](http://cordis.europa.eu/search/index.cfm?fuseaction=news.document&N_RCN=29243)

<sup>2</sup> Communication from the European Commission ‘on scientific information in the digital age: access, dissemination and preservation’, Brussels (COM (2007) 56 final)

new fleet of open access journals by Sage, the world's fifth largest journal publisher.<sup>3</sup>

- Hybrid open access publishing models (e.g. Elsevier, Wiley-Blackwell, Springer, Oxford University Press)<sup>4</sup> may be extended to provide full institutional open access. Contracts between Springer and the Max Planck Society, and with Göttingen University, imply that for authors affiliated with these institutions all publications will be on open access and, furthermore, the final published version is placed in the institution's repository for archiving and dissemination.<sup>5</sup>
- Research communities may self-organise to collect subscriptions in order to convert core journals to open access, by means of a sponsoring consortium that offers a tender. SCOAP<sup>3</sup>, led by CERN, is designed to achieve this and has already collected pledges from 12 European countries and from several important institutions in the US.<sup>6</sup>
- Those who fund research increasingly enable a direct 'author pays' model, by allowing researchers to pay for peer-reviewed publication from grants, which culminates in much of the best research becoming available in open access.
- Institutions, libraries and scholarly societies may facilitate arrangements in order to support the open access publishing of community journals, as the European Geosciences Union, Copernicus Publishers and the Max Planck Digital Library on behalf of Max Planck Society did; by taking out a subscription to EGU journals published by Copernicus which pays for the publication by all affiliated authors.<sup>7</sup>
- Some flagship open access journals have secured and maintained high impact factors (e.g. PLoS Biology, PLoS Medicine or Atmospheric Chemistry and Physics), outperforming rival journals.

Declarations, petitions, mandates and legislation in favour of open access leave no doubt as to the importance and urgency of the cause. Yet, on the whole, the logic of the situation is as much one of 'voice' as of 'exit', which means that it is not open to resolution by market competition and consumer choice factors alone. The challenge for all involved, and this includes publishers who currently base their business on exclusive copyright, is to address the organisational challenges - the most important of which is redistributing the key functions of registration, certification, dissemination, archiving and navigation among publishers, repositories and libraries.

The proposition is for the development of a European model of the digital publishing of scientific information. Such a model would enable an orderly transition to the permanent and open access of digital scientific information. The proposition is that:

---

<sup>3</sup> For example <http://www.sage-hindawi.com/>

<sup>4</sup> For example <http://www.blackwellpublishing.com/static/onlineopen.asp>

<sup>5</sup> For example [http://www.mpdl.mpg.de/news/mpdl-pressrelease-20080204\\_en.pdf](http://www.mpdl.mpg.de/news/mpdl-pressrelease-20080204_en.pdf)

<sup>6</sup> <http://www.scoap3.org/>

<sup>7</sup> See [http://www.mpdl.mpg.de/news/mpdl-pressrelease-20080125\\_en.pdf](http://www.mpdl.mpg.de/news/mpdl-pressrelease-20080125_en.pdf)

1. A digital model may be developed that will result in open access, while preserving and enhancing the viability of a variety of commercial publishing models.
2. The institutional players in the publishing system, namely publishers, repositories and libraries, must be ready to discuss the redistribution of the key functions of registration, certification, dissemination, archiving and navigation in a manner that plays to the strength of each.
3. The condition of any successful elaboration of a digital model is that it is complementary to the technology and economics of the internet, while the litmus test is that it enhances the impact and re-use of scientific information.

The line of reasoning proceeds as follows. The issues, as identified by the Commission of the European Communities in the drive to delineate technical, financial and legal parameters, are examined. Next, the redistribution of functions, as inadvertently sponsored by the Commission through its funding instruments, is investigated. In turn, moves by those funding research and by research organisations to hasten open access to peer-reviewed publications in Europe are reviewed and comparisons with the US are made. Subsequently, a variety of emerging open access publishing models is drawn out. Then, an investigation of a potential synergy between the green and gold road to open access is undertaken- if research institutions do adopt policies that spread the practice of non-exclusive licensing. Finally, a return is made to the question of how publishers, repositories and libraries may redistribute the key functions of registration, certification, dissemination, archiving and navigation.

## **2. Technical, legal and financial parameters for the digital publishing of scientific information**

The Commission of the European Communities outlined organisational, technical, financial and legal issues with regard to access and preservation of digital content and scientific information, in the communications *i2010: Digital Libraries*<sup>8</sup> (COM(2005)465) and *on scientific information in the digital age: access, dissemination and preservation* (COM(2007)56).<sup>9</sup> To date, the Commission has described many of these issues as challenges, by asking, for example, who should decide what and how much is to be preserved or what the transition costs might be in a switch to open access publishing. Yet, the Commission framework itself points towards a type of resolution, if one re-examines the issues in a logical order.

As the Commission recognises, digital technologies based on open standards and interoperability provide the opportunity to provide broad access to ever increasing volumes of scientific publications and data, while principally enabling

---

<sup>8</sup> Communication from the European Commission 'i2010 Digital Libraries', Brussels, (COM (2005) 465 final)

<sup>9</sup> op.cit.

the preservation of ever larger volumes of dynamic content. Consequently, improved and new tools may be deployed for the use and re-use of scientific information (e.g. search tools, data mining). The Commission also addresses the presently limited affordability of scientific information, because of steep price rises and rigid big deals, as well as the gap in access for less wealthy countries and institutions. This feature is not inherent in digital technology, quite to the contrary, it is due to certain organisational and contractual practices in publishing. There is concern that the future costs of digital preservation are not easy to estimate, although this is not primarily about digital storage and storage technology, which is becoming progressively cheaper, but is instead more about digital rights in copying and migrating content as well as avoiding duplication or divergence in organising preservation. Moreover, the Commission recognises that copyright law is not the issue for scientific information, but rather the contractual praxis between authors and publishers as well as publishers and libraries. Scientific data, as such, does not fall under the domain of copyright law. For the purposes of preservation, it is important that a digital copy exists, which carries no restrictions. In the absence of a contract between publishers and libraries, which provides for an unrestricted copy, the Commission envisions advocating a legal deposit law that enhances future copying and the migration of content.

Digital technology favours openness, with e-science making open access intrinsically desirable, and which enhances inclusion and the impact of scholarly communication. Moreover, on the Internet the cost ratio of the traditional publishing model is inverted: dissemination becomes cheap and peer review increasingly expensive. In turn, this favours the severance of certification from dissemination. Scientific information may be licensed non-exclusively, and subject to proper attribution, by all who require a license: publishers, repositories and libraries. Publishers will prosper, as is argued here, if they concentrate on certification and navigation services.

The technical, financial and legal parameters for the digital publishing of scientific information may be specified as follows:

- Technical parameter: Globally integrated science requires open standards, interoperability and open archiving. This does not preclude publishers from using proprietary software in the provision of services. Indeed, publishers are free to adopt any technological business model they chose - as long as scientific information (text, data) is first registered in open repositories and the published version is available in an unrestricted format for digital libraries to archive.
- Financial parameter: New and more investment in scientific publishing is needed, particularly in certification and navigation services. Several shifts make investment necessary and urgent: the worldwide expansion of research and higher education, the internet, the rise of e-science, the increasing complexity of scientific data, the global integration of scientific fields and the rise of the knowledge-based economy. All these shifts point to a growing and vibrant market for certification and navigation services. To fund these services, one important step is to reduce the

costs of online dissemination while achieving rapid and wide dissemination. All of this may be achieved simultaneously if the functions of registration, archiving and dissemination are assigned to repositories for research results (pre-prints, raw data) and libraries for publications (final published version, databases).

- Legal parameter: It is conditioned by the optimality of non-exclusive licenses to organise the certification of research results. The digital dilemma does not apply to scientific information, as authors are not seeking an income from publishing (the exception is, of course, trade-publishing of textbooks and popular science texts). Non-exclusive licensing is highly compatible with competitive markets and supports commercial publishing. Moreover, non-exclusive licensing and the deposition of certified scientific information in libraries is a model that is flexible, with regard to the business model of publishers. Publishers may opt for up-front publishing fees but are not obligated to do so. Indeed, non-exclusive licensing and deposition is highly compatible with a subscription-based model, so that the library would not be paying for access but instead would be buying the scientific information for archiving; including full rights for its dissemination, copying and its re-use.

A number of traditional publishers might object that the outlined parameters are not compatible with their current business model. Moreover, some publishers have heavily invested in digital platforms by which they seek to control the whole chain of registration, certification, dissemination, archiving and navigation. The instrument they deploy is the transfer of copyright, in conjunction with the conventional practice of accepting submissions only on the condition that they are not under consideration elsewhere. The outcome is that libraries are no longer sold a copy of the publication, and instead only purchase a license that allows their users to access the publisher's platform (or that of a host or aggregator). However, these practices are in conflict with the logic of scholarly communication and the requirements of a globally integrated science. Moreover, exclusive copyright is a model that makes the dissemination of scientific information unnecessarily expensive, not only because it inhibits competition in service provision but also because it leads to very expensive models of archiving and dissemination. Finally, exclusive copyright is in contradiction with the increasing need of scientists (in the academy and industry) to re-use, mine and otherwise exploit scientific information in an immediate and unrestricted fashion (rather than requesting and paying for permission on every occasion).

Once complementary parameters for the digital publishing of scientific information are established, then only organisational challenges remain. These are:

- A. The redistribution of functions between publishers, libraries and repositories;
- B. New contractual practices, which research institutions insist must include the right to the open dissemination of publications (and data);

- C. New ways of funding the publication of scientific information so that archiving is guaranteed cost-effectively, while wider benefits are secured by enabling innovation in certification (research excellence and evaluation) and navigation services (impact on R&D, wider public benefits).

### **3. Redistributing registration, archiving and dissemination in Europe: permanent and open access for scientific information**

On the European level, alliances have emerged which push for permanent and open access to scientific information. The Commission is funding the implementation of a digital repository infrastructure (DRIVER – Digital Repository Infrastructure Vision for European Research, FP6 and 7) that is designed to collect, aggregate and display scientific information from all member countries of the European Union. The alliance adheres, in spirit, to the *Berlin Declaration on Open Access in the Sciences and Humanities*, which means that it envisions repositories not as a dumping ground for theses and un-refereed material, but instead as a point of registration and dissemination of scientific publications and data. Characteristic of this European project is the combination of local autonomy and joint standards. A test-bed phase with 51 repositories from the Netherlands, the United Kingdom, Germany, France and Belgium was completed in late 2007. The focus is now on adding more repositories and extending services beyond search and display functions.<sup>10</sup>

The Alliance for Permanent Access has emerged in response to the threat of loss of digital scientific information, if preservation is not coordinated and invested in.<sup>11</sup> Concern extends not only to early digital data at risk, but also to the present and forthcoming large volumes of research data (e.g. from the Large Hadron Collider at CERN). Moreover, large efforts are required to preserve digital publications. Noteworthy is that this alliance includes national libraries, research organisations, funding agencies and a publishers' association (STM). Some of its activities are funded by the Commission (PARSE – Permanent Access to Records of Science in Europe, FP 7). However, while it is clear that scientific data are to be preserved to enable permanent open access, subscription-based publishers that consider copyright seem keen to ensure that while publications are archived, they are not made available. On the other hand, the very same publishers' association has initiated a project proposal that would lead to the open access archiving of thousands of journal articles, from hundreds of subscription-based journals (PEER – Publishing Ecology of European Research, also to be funded in FP 7).<sup>12</sup>

There is much merit in tracing the effects of self-archiving. Significantly, major publishers seem to be ready to deposit the final manuscript in repositories on behalf of the author. Publishers worry that parallel self-archiving may undercut

<sup>10</sup> <http://www.driver-repository.eu/>

<sup>11</sup> <http://www.alliancepermanentaccess.eu/>

<sup>12</sup> [http://www.icm.edu.pl/fpspe/pres/Christine\\_Baldwin-presentation-20-11-2007.ppt](http://www.icm.edu.pl/fpspe/pres/Christine_Baldwin-presentation-20-11-2007.ppt)

their business model. Indeed, duplicate archiving on publishers' platforms and in institutional repositories generates additional costs. Moreover, the author's final manuscript is not a trustworthy and reliable source when it comes to research evaluation and overlay services (e.g. text mining). The logically complementary solution is for the final published version to be hosted by digital libraries in open archives. The NEREUS consortium of libraries, at academic institutions in Europe with leading economics departments, is achieving this goal. On the basis of unlocking publications and datasets from 500 leading economists, the consortium will build multi-lingual access and search facilities, while also pushing out the content to service providers that can create additional value (NEEO – Network of European Economists Online, FP 7 funded).<sup>13</sup>

#### **4. Open access registration and dissemination: repository deposit in the EU and the US**

The open access declarations, petitions, mandates and funding programmes signal a clear preference. It is hard to exaggerate the significance of a preference that has been aggregated to a level at which hundreds of institutional leaders, often after long internal deliberation, will support a particular policy publicly. The Berlin Declaration on Open Access in the Sciences and Humanities (launched in 2003 by the Max Planck Society)<sup>14</sup> has over 240 institutional signatories. The Petition for guaranteed public access to publicly-funded research results (launched in 2007 by DEFF, DFG, JISC, SPARC EUROPE and SURF)<sup>15</sup> has over 700 institutional signatories.

The European Research Council is the new flagship that funds research. Controlled and staffed by the community, it disburses large and prestigious principal investigator grants to senior and junior scholars alike. The ERC is in a position to set a standard for emulation. It is noteworthy that the Open Access mandate, passed on 17 December 2007 applies not just to publications but to data as well, making it an open access mandate for scientific information. For scientific data, the ERC identifies repositories as a primary registry and for publications at least as a co-registry. Given an embargo of six months at the most, the ERC clearly also envisions a dissemination function for repositories.<sup>16</sup>

---

<sup>13</sup> [http://www.nereus4economics.info/about\\_neeo.html](http://www.nereus4economics.info/about_neeo.html)

<sup>14</sup> <http://oa.mpg.de/>

<sup>15</sup> <http://www.ec-petition.eu/>

<sup>16</sup> [http://erc.europa.eu/pdf/ScC\\_Guidelines\\_Open\\_Access\\_revised\\_Dec07\\_FINAL.pdf](http://erc.europa.eu/pdf/ScC_Guidelines_Open_Access_revised_Dec07_FINAL.pdf)

### Box 1 - The European Research Council on Open Access

1. The ERC requires that all peer-reviewed publications from ERC-funded research projects be deposited on publication into an appropriate research repository where available, such as PubMed Central, ArXiv or an institutional repository, and subsequently made Open Access within 6 months of publication.
2. The ERC considers essential that primary data - which in the life sciences for example could comprise data such as nucleotide/protein sequences, macromolecular atomic coordinates and anonymized epidemiological data - are deposited to the relevant databases as soon as possible, preferably immediately after publication and in any case not later than 6 months after the date of publication.

Source: ERC web site

The Council of the European University Association, on 25 January 2008, adopted recommendations by an EUA Working Group on Open Access, which called on universities to build institutional repositories, to mandate open access and to manage copyright contracts more pro-actively so that the institution has deposit, usage and dissemination rights. National Rectors' Conferences in Europe are encouraged to work with those funding research on open access mandates. These should follow the best practice model of the ERC and others who offer funding, like RCUK and NIH. While the EUA recommendation encompasses only publications and not data, the EUA itself has committed to work towards a 'self-archiving mandate for all research results' for all EU research funding, extending the ERC mandate.<sup>17</sup>

The National Institute of Health asked the United States Congress to pass a provision that would require all investigators funded by the NIH to deposit an electronic copy of their final peer-reviewed manuscript in the National Library of Medicine's Pubmed Central, which would be made publicly available no later than 12 months after publication. The provision was signed into law on 27 December 2007. The NIH also envisions the repository as an agent of registration as well as dissemination, albeit limited to peer-reviewed publications with an embargo of up to 12 months. The NIH has designated a single central repository.<sup>18</sup>

---

<sup>17</sup> <http://www.eua.be/index.php?id=396>

<sup>18</sup> <http://publicaccess.nih.gov/>

## Box 2 - The National Institute of Health on Public Access

1. The NIH Public Access Policy applies to all peer-reviewed articles that arise, in whole or in part, from direct costs funded by NIH, or from NIH staff, that are accepted for publication on or after April 7, 2008.
2. Institutions and investigators are responsible for ensuring that any publishing or copyright agreements concerning submitted articles fully comply with this Policy.
3. PubMed Central (PMC) is the NIH digital archive of full-text, peer-reviewed journal articles. Its content is publicly accessible and integrated with other databases.
4. The final, peer-reviewed manuscript includes all graphics and supplemental materials that are associated with the article.
5. Beginning May 27, 2008, anyone submitting an application, proposal or progress report to the NIH must include the PMC or NIH Manuscript Submission reference number when citing applicable articles that arise from their NIH funded research.

Source: NIH web site

While Europe has invested much in a string of networked institutional repositories, the US is home to many successful disciplinary and centralised repositories. The ERC mandate refers to PubMed Central and ArXiv, but more repositories exist in computer science (CiteSeer), economics (RePEc) and the social sciences more generally (SSRN, launching a humanities division).<sup>19</sup> Institutional repository managers might like to note the additional awareness (e.g. editorial services, subject specific distribution lists, new papers, top lists) and evaluation services (e.g. usage and citation metrics) that centralised repositories offer. Moreover, overlay services may be built on these repositories to offer services to readers (e.g. [gopubmed.com](http://gopubmed.com), based on social networking and folksonomy). Institutional repositories might achieve similar utility, but this requires social organization beyond federation and interoperability, namely the agreement of high-quality standards.

If dissemination is a function of repositories, will archiving also become one? Insofar as the repositories function as (or within) a digital library, this might imply a mission for archiving. As the ERC has mandated acting as a data deposit, repositories must clarify if they are to become data libraries too, guaranteeing long-term access. For publications, it might depend on whether the repository holds the final published version or just the author's final manuscript. While the latter is relevant to scholarly communication, it is unlikely to be considered as part of the authoritative scientific record. Publishers' may hope that their platforms have a revenue generating function for the term of the copyright, but this would also oblige them to invest in permanent access. Principally, such a solution has become possible in the US through a publishers and libraries consortium called CLOCKSS (Controlled LOCKSS™, Lots of

<sup>19</sup> <http://www.pubmedcentral.nih.gov/>; <http://arxiv.org/>; <http://citeseer.ist.psu.edu/>; <http://repec.org/>; <http://ssrn.com/>

Copies Keeps Stuff Safe).<sup>20</sup> CLOCKSS is a not-for-profit partnership that deploys open-source software to build a comprehensive digital archive of final published versions, but it only releases these to the public if the publisher chooses to do so (i.e. trigger events have been defined as the publisher ceases operation; or the publisher ceases to publish a title; or the publisher no longer offers back issues; or there is a catastrophic and sustained failure of a publisher's platform – always decided jointly).

Strategy in the US is not as comprehensive as in the EU. CLOCKSS gives publishers long-term control over archiving, while the NIH mandate requires public access within 12 months. However, the NIH version (author's manuscript) is not recognised as an authoritative scientific record. HighWire Press (by Stanford University Libraries, hosting more than 70 of the 200 most highly cited journals)<sup>21</sup> introduces yet another model for the natural sciences with a mix of free journals and free back issues (corresponding to an 18 month embargo). Although innovation has occurred in the US (besides the repositories mentioned, JSTOR and ARTSTOR are also significant central solutions),<sup>22</sup> the present situation is characterised by a lack of coordination and a multitude of solutions that make it difficult for anyone outside the cartel of participating institutions to launch new services. For the EU, this situation allows for the opportunity to increase its competitive advantage by coordinating and implementing a distribution of functions that enables more innovation to happen faster.

## 5. Open access publishing: commercial, institutional and communal models

The aim of the open access movement is well expressed by the text adopted in the Budapest Initiative<sup>23</sup>, namely “... *free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself.*” The following is what is understood by the expression ‘full open access’. The aim is to apply this principle, notably, to the results of scientific research described in peer-reviewed journal articles.

---

<sup>20</sup> <http://www.clockss.org/>

<sup>21</sup> <http://highwire.stanford.edu/>

<sup>22</sup> <http://www.jstor.org/>; <http://www.artstor.org/>

<sup>23</sup> <http://www.soros.org/openaccess/read.shtml>

To be sure, even publishers who frequently sign declarations against the idea of open access often do publish in open access. The most widespread options are:

- Open access for public institutions in low-income countries (HINARI - a selected range of journals on health; AGORA – on agriculture; OARE - on ecology);<sup>24</sup>
- The so-called hybrid open access publishing models whereby the portion of articles for which a publication charge has been paid is made available (e.g. large commercial STM publishers); and
- Embargoed open access publishing whereby content is released after a certain number of months (e.g. HighWire Press).

All of these initiatives are important steps in the transition to full open access publishing.

Over the past years a number of viable full open access publishing models have emerged. These may be seen as open access publishing (OAP) as author pays, through subscription, through a sponsoring consortium, and by way of support and sponsorship.

#### OAP where the author pays

'Author pays' implies a financial transaction for every article that has been accepted for publication after peer review. Typically, publishers set a fee, which is waived in case of demonstrable hardship. 'Author pays' works for a wide variety of publishers, large and small; be they commercial or society publishers.

The 'authors pays' model may support publishers and journals that are fully open access, as well as publication in journals that offer open access by choice. This model is much supported by those who offer funding who permit their grantees to pay publication charges from grant budgets, e.g. the Wellcome Trust, Research Councils UK, the European Research Council and so on.<sup>25</sup>

For example, the number of commercial publishers offering open access journals has increased rapidly, with BioMed Central offering more than 180 journals, Hindawi Publishing Corporation now publishing over 100 journals, and Bentham Science Publishers committed to launching more than 200 titles. At the same time Hindawi and Sage (the world's fifth-largest journal publisher with more than 480 titles) have agreed to launch another suite of open access journals. This indicates that commercial publishers regard open access publishing to be sustainable, at least for the STM sector. Significantly, Hindawi and Sage will publish under a Creative Commons attribution license, whereby authors retain copyright. All journal content from the series will be archived in the not-for-profit electronic archive Portico<sup>26</sup> for permanent open access.

<sup>24</sup> <http://www.who.int/hinari/en/>; <http://www.aginternetwork.org/en/>; <http://www.oaresciences.org/en/>

<sup>25</sup> For example, the Wellcome Trust "will provide grantholders with additional funding, through their institutions, to cover open access charges, where appropriate, in order to meet the Trust's requirements." <http://www.wellcome.ac.uk/About-us/Policy/Policy-and-position-statements/WTD002766.htm>

<sup>26</sup> <http://www.portico.org>

### OAP by subscription

'Subscription' implies that an institution pays a contractually agreed sum to secure open access. This model may be operated in a number of ways. An institution (e.g. university, academy, principally also a national research council) may contract with a publisher to convert all publications of affiliated (or sponsored) authors to open access. A variation with a publisher or journal that offers only hybrid OA is that the institution maintains a subscription to the 'closed' content while at the same time buying open access for affiliated authors. In these scenarios the author no longer pays and there is only one annual transaction (not one for each publication).

Yet another variant is open access publishers who run subscription programmes that involve waiving or reducing publication charges. Significantly, the subscription model of paying for open access provides income security for the publisher as an annual sum is agreed upon. It seems that the 'subscription' model of open access publishing might be used by consortia of research libraries or research councils to convert publications to open access on a large scale.

For example, research institutions sign a hybrid subscription and open access deal. The Max Planck Society and Springer have agreed a deal, by which all of SpringerLink's content is licensed, while simultaneously all publications from MPS researchers are published in open access. The significance of this agreement is that the subscription fee is used to convert all publications to open access. Prior to this agreement, hybrid publication models were based on individual choice and an additional fee for open access. By signing a corporate deal, the publisher continues to receive income, but the research organisation transitions its certified research output to open access.

### OAP through a sponsoring consortium

A sponsoring consortium is based on collecting funds from relevant sources to finance peer reviewed publishing of the output of a specific community or research field. In this model, firstly, a mechanism must be found to identify who is to contribute and how much they will donate. Secondly, the sustainability of the scheme must be guaranteed by long-term pledges. The idea is to redirect libraries subscription monies from paying for reading, to paying for publishing (and make reading a free exercise). Libraries or those offering funding to libraries contribute.

The significance of this model is that sponsors cooperate to raise funds and might finance journals through a call for tender for the peer-review service. This would favour competition and introduce the 'best value-for-money' concept in scholarly communication.

For example, the Sponsoring Consortium for Open Access Publishing in Particle Physics (SCOAP<sup>3</sup>, funding agencies, research organisations and libraries) has been formed to convert publishing in high-energy physics to open access. Whereas publishers have derived income from journal subscription, they will henceforth have to bid in response to a tender from SCOAP<sup>3</sup>. The consortium is raising €10m per year, country-by-country, according to the distribution of articles over the past years. The target of €10m is based on a count of about 5000 HEP publications in a typical year and an estimate that a certification service by publishers would cost up to €2k per article. This community-based model is set on converting five core HEP journals and sections of two broadband journals, which together publish almost all of the relevant articles. This community-controlled approach asks publishers to perform a service: certification for the purpose of measuring quality and productivity of individuals, groups and institutions. The budget allocated for this service is not simply arrived at by counting present subscription costs, but by calculating the true costs of certification. In this, the model assumes that initial registration occurs in a repository (e.g. ArXiv) and libraries handle archiving. By using a tender, SCOAP<sup>3</sup> effectively controls the cost of publishing while enhancing competition.

#### OAP by way of support and sponsorship

There are instances in which support and sponsorship is essential to open access publishing. Generally, it may be that a sponsor (or consortium of sponsors) finds value in supporting open access publishing. More specifically, in the social sciences and humanities a large number of journals exist that are sponsored by institutions (e.g. university departments, research centres) and supported by volunteer effort. It has also been the case that foundations have sponsored the founding of flagship open access journals (e.g. Public Library of Science). A variant of this approach has been the release of publications to open access after a specified embargo period. Particularly society publishers thus forego revenue over the longer term, in the interest of sponsoring open access.

For example, The Open Access Law program is a community-based initiative to convert law journals, run in conjunction with the Science Commons.<sup>27</sup> The model is based on a set of principles as expressed in open access declarations, which have been adopted by more than 35 journals. The journal makes available to the author the final, published version with the expectation that it will be posted in an open access repository. What the journals ask for in return is a limited-term exclusive licence for commercial publication and subsequent attribution (as the first publication). Law journals are frequently sponsored and run by law departments and some prominent journals from Berkeley, Duke, Harvard, Indiana and Michigan have joined. The spirit of performing a certification service, while not seeking to control archiving, is much the same as

---

<sup>27</sup> <http://sciencecommons.org/projects/publishing/oalaw>

in SCOAP<sup>3</sup>. As law journal publishing is distributed and often controlled locally, the route to open access is via institutional sponsorship.

In summary, not only has the authoritative list of peer-reviewed open access journals of the Directory of Open Access Journals surpassed 3000 items,<sup>28</sup> but also some of Europe's largest publishers have adopted open access as a viable business model. The sketch of OAP models clarifies that commercial publishing may continue to thrive, even as open repositories and open archives continue to spread. Indeed, we expect commercial publishing to grow and expand in response, not only to further growth of research and higher education, but also because of the need for enhanced navigation and certification services.

## **6. Nonexclusive licensing: Synergies between a new kind of green open access and open access publishing?**

The so-called green road to open access is also known as self-archiving movement. Indeed, the dissemination and awareness services of some repositories are very effective and efficient, so much so, that hundreds of scholars volunteer their time to carry out these tasks (e.g. SSRN, RePEc, ArXiv). Open access to the final published version is currently only provided by open access publishers. Open access to the final published version enables the seamless, coordinated and cost-effective archiving of all scientific information for combination, search, mining and re-use (which, for example, is the stated aim of the Alliance for Permanent Access).

'Green' OA is problematic in that it introduces a new, 'artificial' version into the publishing system (often called Version 2), the author's peer-reviewed final manuscript, which is then archived alongside the working paper (Version 1) and the published paper (Version 3). Green OA archives an 'inexact' copy. While this may enhance access for researchers on a day-to-day basis, there are numerous problems with Green OA, besides the extra cost incurred. The most important issues are, that these not-quite-published manuscripts are rejected for research assessment (in the last UK RAE) and that their use for all kinds of value-added services like metric research evaluation, text mining, navigation services and science education is limited. However, the Faculty of Arts and Sciences of Harvard University<sup>29</sup> has passed a mandate that could create a synergy between Green OA and open access publishing. The key is that Harvard has established the principle of non-exclusive licensing. The same type of policy is being considered at the University of California.<sup>30</sup>

---

<sup>28</sup> <http://www.doaj.org/>

<sup>29</sup> [http://www.fas.harvard.edu/~secfas/February\\_2008\\_Agenda.pdf](http://www.fas.harvard.edu/~secfas/February_2008_Agenda.pdf)

<sup>30</sup> <http://osc.universityofcalifornia.edu/openaccesspolicy/>

### **Box 3 - The Harvard Faculty of Arts and Sciences policy of nonexclusive licensing**

Each Faculty member grants to the President and Fellows of Harvard College permission to make available his or her scholarly articles and to exercise the copyright in those articles. In legal terms, the permission granted by each Faculty member is a nonexclusive, irrevocable, paid-up, worldwide license to exercise any and all rights under copyright relating to each of his or her scholarly articles, in any medium, and to authorize others to do the same, provided that the articles are not sold for a profit. The policy will apply to all scholarly articles written while the person is a member of the Faculty except for any articles completed before the adoption of this policy and any articles for which the Faculty member entered into an incompatible licensing or assignment agreement before the adoption of this policy. The Dean or the Dean's designate will waive application of the policy for a particular article upon written request by a Faculty member explaining the need.

To assist the University in distributing the articles, each Faculty member will provide an electronic copy of the final version of the article at no charge to the appropriate representative of the Provost's Office in an appropriate format (such as PDF) specified by the Provost's Office. The Provost's Office may make the article available to the public in an open-access repository.

The Office of the Dean will be responsible for interpreting this policy, resolving disputes concerning its interpretation and application, and recommending changes to the Faculty from time to time.

Resolution adopted unanimously by the Harvard Faculty of Arts and Sciences, February 12, 2008

Source: Harvard web site

The significance of this policy is that the author retains the copyright and grants Harvard a non-exclusive license. In consequence, and on principle, publishers can no longer ask for the full transfer of copyright and they cannot embargo the release of the Harvard open access version. As Harvard encourages authors to retain copyright, publishers are asked to shift their business model to a 'license to publish'. But, that license is non-exclusive at least in the sense that a near-identical version will be archived and disseminated by Harvard in open access.

While it must be acknowledged that Harvard will also post the problematic 'Version 2', it does so with a plan for the future. An institutional policy of non-exclusive licensing regulates the market for scientific publishing. It does so in a way that provides incentives for publishers to re-consider their role. Harvard's new copyright regime leaves publishers no grounds on which to prohibit the resulting open access, now or in the future. If research institutions opt to take on and to underwrite the functions of registration, archiving and dissemination, then publishers have much to gain by concentrating on certification and navigation services. In short, Harvard-type open access policies instigate a transition to

open access publishing. It remains to be seen whether this will be translated into a competitive advantage in the US or adopted in the EU.

## **7. How will publishers, repositories and libraries share the key functions of registration, certification, dissemination, archiving and navigation?**

It is proposed that repositories focus on the registration of knowledge claims and early dissemination of research results, in services that are free to readers and authors: publishers undertake certification of research results and their publication, followed by navigation services targeted at specific audiences; and libraries-cum-repositories archive the certified and published version of scientific information (publications and data), guaranteeing long-term preservation and open access.

Further arguments in support of this redistribution of functions may be summarised as follows:

- **Registration:** The idea of a journal, as a registrar of knowledge claims, has always been flawed if the journal rejects submissions, for then authors' knowledge claims initially would exist only in a 'limbo' of rejection and resubmission until acceptance. The preferable digital solution is public registration in a repository, as any deposit would be immediate and independent of any intermediary. Incidentally, a direct and open deposit also better protects the author's moral rights – including claims to priority and against plagiarisers, as well as protection against the taking of ideas from 'not yet published' texts, which is particularly important for younger authors.
- **Archiving:** A journal has only ever been an archive for the first published version. In many fields the subsequent modification and re-issue of an article has not been uncommon, especially of those articles that are considered to be groundbreaking. Archiving is ultimately the function of libraries. If publishers assume the function of Internet archiving, they have to worry about retro-digitisation of print material and the maintenance of archives, as well as facing potentially large future financial risks due to the obsolescence of current technology. The age-old tradition of the deposit library may be revitalised by designating digital deposit libraries.
- **Dissemination:** In the print world, the printed journal was the premier vehicle for dissemination. Industry experts judge first copy costs, on average, to be \$3,500-4,000. OA author-side charges, where raised by the publisher, tend to range between \$1,500-3,000 (i.e. the range from BMC to PLoS and Springer); this would imply that if the first copy is produced in-house in the traditional way, costs remain substantial. An informal enquiry with some of the major repositories yielded an average 'first copy cost' below \$5 for pre-prints and post-prints posted by the author and pushed out through editorial services. Repositories certainly do not, for this sum, provide as extensive editorial services as publishers, but if, upon enquiry, arXiv, RePEc and SSRN all say that their equivalent 'first copy costs' would be below \$5, this indicates, even if allowing for a large margin of error, that publishers should consider shedding dissemination.

- Certification: The internal logic of science requires not only the certification of the validity and reliability of knowledge claims by peer review, but also, over time, the ranking of knowledge claims by value. Reputation, status and prestige are peer produced in an iterative process that includes not just publication in prestigious journals, but also keynote speeches, research grants, scholarly awards and, ultimately, the Festschrift, the lifetime achievement award and the Nobel Prize (or its substitutes). Particularly in the social and cultural sciences, as well as the humanities, it has been a long-standing practice to continuously reprint books and articles that are highly valued by successive generations of scholars. The certification of knowledge claims is and will remain labour intensive and expensive.
- Navigation: The knowledge-based society thrives on the (commercial) re-use of research, creating a case for treating research results as information in the R&D pipeline as well as making commercial re-use of research articles and data desirable. For that to happen, research results must be made available on the basis of a non-exclusive license, so that a multitude of service providers may target wider and new audiences, particularly in the knowledge-intensive industries and services.

Key to this redistribution of functions is that publishers shift their mission from content holding to service provision, based on non-exclusive licensing to ensure market competitiveness. Service provision by publishers may well be commercial and profit-oriented. Indeed, the digital model here proposed is highly compatible with the funding of publishers by capital markets. Repositories and libraries take over the functions of registration and archiving, respectively, and in doing so enable open access dissemination of the pre-print (working paper) as well as of the final, published version.

This European model distributes functions among publishers, libraries and repositories in such a fashion that permanent and open access ensure. Moreover, the redistribution of functions proposed is congruent with the norms of science, is cost-effective, it simplifies version control (preprint and published version only), it maximises access, inclusion and impact in scholarly communication and it opens up the commercial re-use of scientific information, including services aimed at research evaluation.

Addendum: Published articles and working papers upon which I draw are the following:

Moving Out of Oldenbourg's Long Shadow: what is the future for society publishing?  
Learned Publishing 20, 4 (2007) pp.259-266

Cyberscience and the Knowledge-based Economy, Open Access and Trade Publishing:  
from contradiction to compatibility with nonexclusive copyright licensing.

Winner Yale Law Information Society Project writing competition 'Access to Knowledge'.

International Journal of Communications Law and Policy, Issue 12, 2008  
(republished in: Policy Futures in Education, Vol. 6, No. 4, 2008)

Research Universities: autonomy and self-reliance after the Entrepreneurial University.  
Policy Futures in Education, Vol. 6, No. 4, 2008

On Cost-Sharing, Tuition Fees and Income Contingent Loans for Universal Higher  
Education: a new contract between university, student and state?

Policy Futures in Education, Vol. 6, No. 4, 2008

The Rise of the Post-Doc as Principal Investigator? How PhDs may advance their career  
and knowledge claims in the New Europe of Knowledge.

Policy Futures in Education, Vol. 6, No. 4, 2008

Open Access in Natural and Social Science: the correspondence of moves to enhance  
access, inclusion and impact in scholarly communication,

Policy Futures in Education, Vol. 6, No. 4, 2008

All research articles are available in open access through the Social Science Research  
Network author page: <http://ssrn.com/author=434782>

The following working papers are also available:

Armbruster, Chris, "Five Reasons to Promote Open Access and Five Roads to  
Accomplish it in Social and Cultural Science" (November 12, 2005).

Available at SSRN: <http://ssrn.com/abstract=846824>

Translated and to be published in Italian soon.

Armbruster, Chris, "Access, Usage and Citation Metrics: What Function for Digital  
Libraries and Repositories in Research Evaluation?" (January 29, 2008).

Available at SSRN: <http://ssrn.com/abstract=1088453>