Embedding Open Access into the European Landscape – the Contribution of LIBER

by PAUL AYRIS

INTRODUCTION

In 2006, my predecessor as Chair of the LIBER Access Division Raf Dekeyser published an article in the LIBER Quarterly entitled ‘The LIBER workshops on the “Open Archives Initiative” at CERN, Geneva’ (Dekeyser, 2006). In this article, Dr Dekeyser described how the LIBER Access Division grasped the challenge of Open Access in Europe and instigated a series of Workshops in CERN to act as a focus for European thought and development in this area. The purpose of this article is to continue the description and analysis of Raf’s initial paper by looking at LIBER’s contributions to the Open Access agenda and debate since 2003.

OVERVIEW OF OPEN ACCESS DEVELOPMENTS IN EUROPE

February 2007 the EU hold an important conference on the future of Open Access in Europe, hosted by the Commission in Brussels. The outcome of that meeting was interesting, in that the Commission gave encouragement to the Open Access movement whilst not mandating it de facto across Europe. It did, however, mandate the publication/deposit of research outputs into Open Access sources if funded by European research funding.

On Thursday 15 February 2007 a delegation led by Sijbolt Noorda, on behalf of the European University Association, presented the 'Guaranteed public access to publicly-funded research results' petition to Janez Potočnik, EU Commissioner for Science and Research. The petition had been signed by over 17,500 individuals and 743 institutions and called on the EC to implement the recommendations of the EU Study on the economic and technical evolution of the scientific publication markets of Europe (European Commission, 2006). The petition is still open for additional signatures and now contains over 26,000 signatories.

As Chair of the UNICA Scholarly Communications Group, the present author gave a presentation in one of the Conference Workshops. UNICA is a network of 41 universities from the capital cities of Europe, with a combined strength of over 120,000 staff and 1,500,000 students. Its role is to promote academic excellence, integration and co-operation between member universities throughout Europe. It seeks also to be a driving force in the development of the Bologna process and to facilitate the integration of universities from Central and Eastern Europe into the European Higher Education area.

In my presentation, I tried to identify what researchers want from scholarly publishing. It might be thought that access to a critical mass of scholarly research information is a given for any major university library. Sadly, that is not the case. Subscriptions can act as a barrier to use where cancellations mean a loss of access by researchers and students to key academic texts. An informal, unpublished study looked at the availability of research outputs (in commercial journals) from my own University - UCL (University College London) - to researchers and students in the University of Nottingham.
Table 1: Availability of UCL research outputs at the University of Nottingham

Here, it is clear than less than 60% of UCL’s published research outputs are available to Nottingham researchers. This is bad news - both for Nottingham and for UCL, whose research outputs are not as widely available as its researchers might expect.

Table 2 is based on a published study by Matt Cockerill on the public availability to the UK National Health Service (NHS) of its own funded research outputs. The NHS is one of the biggest employers in the world, and 90% of its funded research is available online.
Table 2: Access by the NHS to its own funded research

Only 40% of that research is available to an average NHS hospital out in the country - again not a result that the NHS might expect. In both these case studies, Open Access to the research literature would solve all these problems since everything would be available to everyone.

EUROPEAN RESEARCH PROJECTS

**DRIVER** is an international partnership which is seeking to build an interoperable network of Open Access repositories across Europe. The project is a joint collaboration between ten international partners with the intention to create a knowledge base of European research. DRIVER will put a test-bed in place across Europe to assist the development of a knowledge infrastructure for the European Research Area. The project will build upon existing institutional repositories and networks, from countries including the Netherlands, Germany, France, Belgium and the UK. The aim is for one large-scale virtual content resource to be created to access and integrate individual repositories. DRIVER will also prepare for the future expansion and upgrade of the Digital Repository infrastructure across Europe and will ensure the widest possible user involvement.

**PLANETS** is not itself an Open Access project, but it is a complementary piece of work on a European scale. PLANETS brings together European national libraries and archives, leading research institutions, and technology companies to address the challenge of preserving access to digital cultural and scientific knowledge. The PLANETS outputs will be available for use by the Open Access community. PLANETS will:

- Develop Preservation Planning services that will empower organisations to define, evaluate, and execute preservation plans. The plans will reflect the organisation’s preservation policies, as well as the content in its collections, and the way the content is used.
- Develop methodologies, tools and services for the characterisation of digital objects in order to identify the best preservation plans.
- Evaluate existing tools and services to support preservation actions and the development of innovative solutions based on the integration of existing tools and on the design and implementation of new tools where an unfulfilled requirement can be demonstrated.
• Establish a preservation test bed to provide a consistent and coherent evidence-base for the objective evaluation of different preservation protocols, tools and services and for the validation of the effectiveness of preservation plans.

• Implement an interoperability framework, within which the deliverables from each of Planet’s sub-projects can be seamlessly integrated with each other in a distributed service network.

• Undertake dissemination and take-up activities, to ensure the widest possible adoption of PLANETS deliverables among commercial tool and service providers as well as within the user community.

Clearly, Europe is alive to community-led developments which will embed new ways of thinking and new ways of doing into the European research infrastructure. Library and information provision are central to this vision and it is important to note that libraries stand at the heart of this work. It is in this context that LIBER must work to develop its own contribution to the European information landscape and to support both its member libraries and the researchers in its member institutions to take advantage of these new developments.

OAI4 AND OAI5

In his 2006 article, Dr Raf Dekeyser looked at and analysed the first three OAI Workshops in CERN, Geneva. Two further workshops have been held - OAI4 in 2005 and OAI5 in 2007. The plenary sessions from both workshops are available as videos from the CERN repository, linked from the relevant workshop programme. The aim of the workshops is to provide an environment where a mix of cutting-edge technical, publisher, library, and researcher presentations and interactions can take place, where contacts can be made and ideas exchanged and developed.

The workshops continued to show an upward trend in the number of registered delegates, with 175 in 2005 and 215 in 2007, the latter emanating from 32 countries. 217 registered attenders represents a record for OAI Workshops.

OAI5 WORKSHOP, 18-20 APRIL, 2007

The format, which has been established in earlier workshops, was continued in OAI4 and OAI5. The workshop starts with Tutorials on a number of themes. In 2007, there were seven Tutorials:

• Eprints v3 demonstration
• Advocacy and policy issues
• OAI-PMH basics
• Object models and object representation
• Access control
• Managing a repository: which legal aspects are really important?

There then follows a formal opening to the Workshop. In 2007, this was performed by Maximilian Metzger the Secretary-General of CERN, with a response by Hans Geleijnse, President of LIBER.

At both OAI4 and OAI5 the technical presentations then followed, the programme for which was coordinated on both occasions by Herbert Van De Sompel from the Los Alamos National Laboratory. The collocation of technical presentations alongside library and publisher presentations is, we believe, unique in an Open Access Conference anywhere in the world and the OAI Organising Committee is keen to maintain this tradition. At OAI5, there were technical presentations by:

• Simeon Warner – State of the OAI-PMH protocol
• Herbert Van De Sompel - OAI Object Re-Use and Exchange
• Leo Waaijers and Thomas Krichel - Author identification: national and disciplinary approaches
There are numerous examples of the need to re-use objects across repositories in scholarly communication. These include citation, preservation, virtual collections of distributed objects, and the progression of units of scholarly communication through the registration-certification-awareness-archiving chain. The last several years have brought about numerous open source repository systems and their associated communities. The Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) has been the initial catalyst for repository interoperability. However, there is now a rising interest in repositories no longer being static components in a scholarly communication system that merely archives digital objects deposited by scholars. Rather, they can become building blocks of a global scholarly communication federation in which each individual digital object can be the ore that fuels a variety of applications. Both the interest in this type of federation, and the insights gained thus far are sufficiently strong to move beyond prototypes and to support an effort to formally specify this next level of interoperability across repositories. Through the support of the Mellon Foundation, a two-year international initiative to define this interoperability fabric started in October 2006. The effort is in the context of the Open Archives Initiative, and is named Object Reuse and Exchange (OAI-ORE). OAI-ORE is intended to be a complement to OAI-PMH. OAI-ORE is co-ordinated by Carl Lagoze and Herbert Van De Sompel, and consists of international experts on Advisory, Technical and Liaison Committees. The Technical Committee held its first meeting in January 2007 and began its initial work to develop, identify, and profile extensible standards and protocols to allow repositories, agents, and services to interoperate in the context of use and re-use of compound digital objects beyond the boundaries of the holding repositories. In his presentation, Herbert Van De Sompel gave an overview of the current activities, including: defining the problem of compound documents within the web architecture, enumerating and exploring several use cases, and identifying likely adopters of OAI-ORE.

Bill Hubbard’s presentation looked at the development of a specific Open Access tool - OpenDOAR, the Directory of Open Access Repositories recording the existence and characteristics of global Open Access repositories. Using the OpenDOAR toolkit, for example, it is possible to show which repositories use which software platform, as described in Figure 1.

![Usage of Open Access Repository Software - Europe](http://liber.library.uu.nl/)

**Figure 1:** Use of repository platforms, as mapped by OpenDOAR

The future importance of OpenDOAR from the SHERPA team probably lies in the ability of the Directory to be interrogated directly through machine-to-machine interaction. OpenDOAR complements the SHERPA Romeo database, which acts as an authoritative source for discovering publisher copyright policies.
Day 2 at OA5 saw the library and publisher presentations and a very active poster session with 31 registered posters. The formal presentations were from:

- David Prosser - Repositories, policy and politics
- Norbert Lossau - DRIVER: Building a Sustainable Infrastructure of European Scientific Repositories
- John Robertson - The repository ecology: an approach to understanding repository and service interactions
- Jens Vigen - On the golden road: Open Access publishing in particle physics
- MacKenzie Smith - Open Access Forever -- Or Five Years, Whichever Comes First: Progress on Preserving the Digital Scholarly Record
- Alma Swan - Business models for digital repositories
- Andrew Wray - Publishers and repositories

David Prosser gave an overview of political developments in the Open Access arena, underlining the important role that SPARC-Europe has to play in keeping track of, and contributing to, this agenda. MacKenzie Smith gave a very approachable introduction to the concept of digital curation and looked at offerings in this area across the world. 31 posters were registered for presentation at the Workshop - again a record for the event - and these can be seen in the online poster gallery.

Jens Vigen outlined the work at CERN in developing new Open Access business models. The particle physics community has over the last 15 years achieved so-called full green Open Access through the wide dissemination of preprints via arXiv, a central subject repository managed by Cornell University. However, green Open Access does not alleviate the financial difficulties of libraries, who are still expected to offer access to versions of records of the peer-reviewed literature. For this reason the particle physics community is now addressing the issue of gold Open Access by converting a set of the existing core journals to Open Access. A working party is now working to bring together funding agencies, laboratories and libraries into a single consortium, called SCOAP3 - Sponsoring Consortium for Open Access Publishing in Particle Physics (SCOAP3, 2007). This consortium will engage with publishers towards building a sustainable model for Open Access publishing. In this model, subscription fees from multiple institutions will be replaced by contracts with publishers of Open Access journals where the SCOAP3 consortium is a single financial partner.

The end of day 2 saw 6 Breakout Groups meet to discuss topics of mutual interest to the 200+ attenders. These were:

- This business of repositories: the case, the cost, the challenges
- Content aggregation and network architectures
- Repositories in an institutional context
- Scientific quality assurance I: Research assessment and new metrics
- Scientific quality assurance II: Peer review – technical and organizational issues
- Institutional repositories - Policy and risk management

Reports from these Breakout Groups have been loaded into the CERN repository, linked from the Program.

The final day saw a series of presentations from end-users. These were:

- Doctoral e-Theses; experiences in harvesting on a national and European level - Gerard Van Westrienen
- Science commons – John Wilbanks
Gerard van Westrienen showed that in many European countries, doctoral e-theses are an integrated part of institutional repositories, which have been set up in recent years. In some countries they have been harvested on a national level, such as the Netherlands with the Promise of Science portal. In October 2006, the SURF Foundation (Netherlands), JISC (UK) and DIVA (funded through BIBSAM in Sweden) started a common project to harvest repositories with e-theses on an international scale and to set up a freely-accessible European portal, to test in practice the interoperability of such repositories. The project, DART-Europe, which aims at creating a value-added service for doctoral e-Theses, will finish in June 2007. The presentation underlined lessons learned from the work, and the first results of the demonstrator, the DART-Europe E-thesis Portal (DEEP), in five countries: Denmark, Germany, the Netherlands, Sweden and the UK. The speaker also described developments regarding the European co-ordination of doctoral e-theses through the GUIDE Working Group, funded by JISC and SURF.

Researchers in cosmology and astrophysics depend on the arXiv repository for the registration and dissemination of their work, as well as for current awareness. Nonetheless, they continue to submit papers to journals for review. Could rapid quality certification be overlaid directly onto the arXiv repository? This presentation introduced the RIOJA (Repository Interface to Overlaid Journal Archives) project, on which a group of cosmology researchers from the UK is working with UCL Library Services and Cornell University. The project is creating a tool to support the overlay of journals onto repositories, and will demonstrate a cosmology journal overlaid on top of arXiv. RIOJA will also work with the cosmology community to explore the social and economic aspects of journal overlay in this discipline: what other value, besides the quality stamp, does journal publication typically add? What are the costs of the ideal overlay journal for this community, and how could those costs be recovered? Would researchers really be willing to submit work to a new journal overlaid on the arXiv repository?

CONCLUSIONS FROM THE WORKSHOP?

As chair of the OAI5 Organising Committee, it was my privilege to give the round-up presentation. I highlighted a number of conclusions and identified some gaps in the work going on in the global Open Access arena. I offered a new definition, if not a new term, for Scholarly Communication - an inclusive
definition which opens up the importance of Open Access in the scholarly world. I suggested that the definition could be phrased as follows:

- The authoring, publishing, dissemination, and reading of information produced for teaching, learning or research in whatever format;
- with the tools, measures and systems needed to provide access to and store these materials in perpetuity

The additions to the traditional definition of Scholarly Communication are in italic and, I suggested, cover a lot of the work which was being discussed in OA15.

I then looked at a gap survey recently commissioned by the Research Information Network (RIN) in the UK. The survey identifies gaps in our knowledge base, judged by the existing published literature. For Open Access, the Report concluded inter alia that the deposit of material in Open Access sources seems to result in more, and earlier, citations of that work. It concluded that longitudinal studies were needed to take this work further (Electronic Publishing Services, 2006).

I then looked at the need for advocacy, to get more content into repositories. This is a theme which had been identified by several speakers throughout the workshop. The need was underlined by my analysis of the content of five specimen UK repositories, the results of which are given in Table 3. Bar possibly one repository, these figures are hardly impressive. The Joint Information Systems Committee (JISC) in the UK has invested over £17 million in a public repository programme. Do these figures represent a good Return on Investment? The answer, to date, must surely be no.

<table>
<thead>
<tr>
<th>Repository</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repository 1</td>
<td>821 records</td>
</tr>
<tr>
<td>Repository 2</td>
<td>2521 records</td>
</tr>
<tr>
<td>Repository 3</td>
<td>512 records</td>
</tr>
<tr>
<td>Repository 4</td>
<td>93 records</td>
</tr>
<tr>
<td>Repository 5</td>
<td>69 records</td>
</tr>
</tbody>
</table>

Table 3: Counts of the contents of 5 specimen SHERPA repositories

I then picked up the theme of digital preservation and looked at the work of the LIFE project - a LIBER collaboration between its Access and Preservation Divisions, which is funded by the JISC in the UK. LIFE has developed a lifecycle costing model for digital curation and a generic preservation model to cost digital preservation. LIFE defines preservations thus:

\[
\text{Preservation} = \text{Technology watch} + \text{Preservation frequency} \times \text{Overall preservation action}
\]

Figure 3: The LIFE preservation model

Looking at the actions which make up those activities, it is possible to attribute costs to each of them. The actions, which the LIFE project defined, are:
Table 4: Actions which populate the LIFE model

Having defined each of these actions, it is possible to cost them and to arrive at a cost for the long-term preservation of digital objects. Table 5 gives a breakdown of such costs for the long-term archiving of a website:

Table 5: Exemplar costs for the long-term curation of a website

Such costings are a first attempt to arrive at actual costs for digital curation activity. In one sense they are almost certainly wrong. On the other hand, they do seem to indicate that the underlying formulae and methodological approach are sound. LIFE has received a further round of funding from the JISC and will firm up the economic modelling with the Swedish Business School, looked at more exemplar case studies and provide more examples of costs. The SHERPA-LEAP network of Open Access repositories in London will form one of the case studies.

OAI SOCIAL ACTIVITIES

To balance the formal sessions of the Workshop, there are lighter social moments. The second SPARC Europe award for outstanding achievement in Scholarly Communication was awarded by the Chair of the SPARC Europe Board, Bas Savenije, to the SHERPA team in the University of Nottingham during the meeting. Furthermore the Director General of CERN, Dr Robert Aymar, was formally presented with sponsorship funding for CERN’s SCOAP3 development project by the University of Patras in Greece to inaugurate fundraising in this area.

http://liber.library.uu.nl/ Volume 17 Issue 2 2007
One of the hallmarks of the OAI Workshops is the famous (or perhaps infamous?) ‘drinks sharing’ evening. Every participant in the OAI Workshops is invited to bring a bottle of wine or beer or lager from their own country to share in a drinks party. In the course of one evening, it is possible to travel across the whole of Europe and beyond, tasting (probably drinking) alcoholic beverages from many, many nations. Of course, this is all undertaken to build greater understanding and to exchange news of developments, right until 04.00 in the morning when the drinks have run out.

In addition to these events, which take place during the Workshop itself, OA15 had a choice of 8 additional social activities - ranging from a tour of the CERN Laboratory to a coach tour of Chamonix-Mont-Blanc in France.

OAI ORGANISATION

The organisation of such Workshops is an enormous piece of work, and it could not take place without lots of enthusiastic help from the Organizing Committee.

![Figure 4: The OA15 Organising Committee in CERN, Geneva](image)

Very special thanks must be reserved for the Local Organising Committee - Joanne Yeomans, Anne Gentil-Beccot and Susanne Schaefer, for whom the organising of OAI Workshops represents a tremendous contribution. They are situated on the front row in the picture in Figure 4 along with Jens Vigen, Scientific Information Officer at CERN, situated second left on the back row.

The Workshops would not be possible without generous contributions from our sponsors, who for OA15 were:
Ex Libris is a leading worldwide developer and provider of high-performance applications for libraries, information centres, and researchers, with installations around the globe.

JISC (Joint Information Systems Committee) supports UK Further and Higher education and research by providing leadership in the use of Information and Communications Technology in support of learning, teaching, research and administration.

LIBER (Ligue des Bibliothèques Européennes de Recherche) is a non-governmental association of the research libraries of Europe, formed under the auspices of the Council of Europe and having the status of a consultative body with that Council. LIBER represents and promotes the interests of the research libraries of Europe.

Open Repository is a service from BioMed Central that builds, hosts and maintains repositories on behalf of institutions and organizations.

Open Society Institute's Information Program has two aims. First, to enhance the ability to access, exchange, and produce knowledge and information for key open society constituencies and disadvantaged groups, especially in the poorer parts of the world. Second, to use new tools and techniques to empower civil society as a force for open society in national, regional and global arenas.

SHERPA-LEAP (the London E-prints Access Project, affiliated to SHERPA) was founded in 2004 to develop Open Access e-print repositories for member institutions of the federal University of London.

SPARC and SPARC Europe are international alliances of research institutions, libraries and organizations that encourage competition in the scholarly communications market.

We thank them for their generosity and for their willingness to work with the community in this exciting way.

THE FUTURE FOR OAI WORKSHOPS

The future for OAI Workshops seems secure, and planning is already underway for OAI6 in 2009. LIBER is proud to have initiated and to support this series of workshops as we believe that it provides a laboratory in which the library, technical, publishing and research communities can come together to explore new developments and to exchange ideas in a secure and supportive environment.

In partnership with the University of Glasgow, the LIBER Access Division has also launched a companion workshop series called Open Scholarship. The first meeting was held in Glasgow in 2006, with over 200 delegates attending. A second workshop, integrated as a strand into the international Open Repositories Conference in Southampton in 2008, is being planned.
As Europe grows ever larger, with new countries seeking membership of the Community, the development of a sustainable information infrastructure is essential to deliver the benefits of membership to all European citizens. LIBER is proud to lead in this important area on behalf of European research libraries.

REFERENCES


WEBSITES REFERRED TO IN THE TEXT

DART-Europe. http://elib-a.ucl.ac.uk/About/
DiVA, the Academic Archive Online. http://www.diva-portal.org/

JISC - Joint Information Systems Committee. http://www.jisc.ac.uk/
LIBER - Ligue des Bibliothèques Européennes de Recherche. http://www2.kb.dk/liber/


Petition for guaranteed public access to publicly-funded research results. http://www.ec-petition.eu/
RIN - Research Information Network. http://www.rin.ac.uk/
SHERPA - Securing a Hybrid Environment for Research Preservation and Access http://www.sherpa.ac.uk/
SHERPA-LEAP - the London E-prints Access Project, affiliated to SHERPA. http://www.sherpa-leap.ac.uk/
SHERPA Romeo database. http://www.sherpa.ac.uk/romeo.php
SURF Foundation. http://www.surffoundation.nl/
UNICA - Network of Universities from the Capitals of Europe. http://www.ulb.ac.be/unica/

NOTES


2 The full social Programme is listed at http://oai5.web.CERN.ch/oai5/Social_events.html

3 See http://www.lib.gla.ac.uk/openscholarship/

4 See http://openrepositories.org/2008/