

# LIBER-EBLIDA Digitisation Workshop

**Paul Ayris**

Chair LIBER Access Division, Library Services University College London,  
Gower Street, London WC1E 6BT, UK, [p.ayris@ucl.ac.uk](mailto:p.ayris@ucl.ac.uk)

## Introduction

95 participants from 23 countries attended a Workshop on the Digitisation of Library Material in Europe, from 24 to 26 October 2007 in the Royal Library, Copenhagen, Denmark.

The Workshop was jointly organised by LIBER and EBLIDA.<sup>1</sup> In a full report on the Workshop on the JISC website, the JISC highlighted that 'high on the agenda were moves towards a European Digital Library – the delivery of integrated access to the digitised collections of libraries, archives and museums across the EU.'

At a global level, this was indeed one of the main aims of the Workshop. As the Chair of the breakout groups, and the final session of the Workshop, I gave an interview for the JISC immediately after the final session where I looked at what I thought were some of the issues to be identified by the Workshop.<sup>2</sup> The purpose of this paper is to describe some of these issues in a little more detail and outline what work LIBER and EBLIDA have undertaken in the wake of the significant momentum which the Workshop has engendered for European digitisation activity.

## Top-Level Themes

A number of top-level themes emerged from the Workshop:

### Vision

One theme, which struck me, was the absence of a clear vision by European research libraries on what they wanted or expected the role of digitisation

*Fig. 1: Attenders at the Copenhagen Digitisation Workshop.*



of content to be. The Workshop reviewed digitisation activity in a number of European countries. Clearly, there is a significant amount of activity in national jurisdictions. One of the emerging findings from the Workshop is that digitisation of content has reached a level of momentum where libraries and users are expecting more and more to be delivered digitally. It is too early to say, but perhaps the Copenhagen Workshop marks a time when Europe's libraries determined that digitisation of content in libraries was a major concern for them; and that Copenhagen was where they discovered that this is a common feeling across Europe.

Perhaps the challenge from here is to try to co-ordinate all this activity across Europe by identifying a vision that all participants can share. My own attempt to outline such a vision, built on the discussions in Copenhagen, is this:

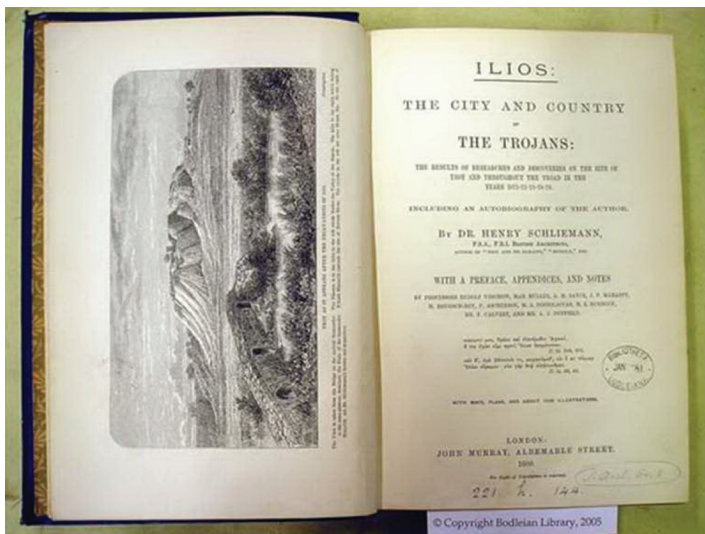
LIBER's and EBLIDA's vision should be that a user, from his/her desktop anywhere in Europe, should freely be able to access whatever content he/she needs, with no technical restrictions on bandwidth available to them, and no limitations on access to the digital materials; and that the content should be securely preserved digitally in perpetuity by trusted digital preservation services.

## Commercial Versus Institutional

Both academic librarians and commercial suppliers were present at the Copenhagen Workshop. The difference in approaches to digitisation activity between the two groups is significant.

Looking at the growth and take-up of commercial e-journals over the last ten years, it seems clear that this phenomenon has been led by the commercial world. The extensive provision of e-journal materials has revolutionised the way that researchers in Science, Technology and Medicine use libraries and literature. By and large, researchers in these subject areas are happy to stay away from physical library buildings because they can access all they need remotely from the library over the network – and provision is 24×7.

Fig. 2: *Ilios*, one of the books from the Bodleian Library which has been digitised as part of the Google Library project.<sup>3</sup>



In terms of e-books, again it is (arguably) commercial companies rather than libraries which are leading the way. The Google Library project has the following objectives:

- to digitise materials from five major research libraries: Harvard, Michigan, New York Public, Oxford, Stanford;
- to create OCR'd text, with indexes for search and retrieval via Google search services and, in particular, Google Book Search;
- to provide online searching and access to hitherto inaccessible printed materials for the public worldwide.

Again, it is a commercial company which has taken the lead in this area and not libraries. Libraries are invaluable partners in this exciting adventure, as they have the content, but it is Google who initiated the work.

All this does raise an interesting question. Why are libraries not asserting their role more systematically? If it had been left to libraries, would academic researchers now have access to the wealth of e-journal literature which they currently enjoy? Is there a lesson from e-journal developments that libraries need badly to learn in tackling the digitisation agenda?

### **Collections Strategy**

The breakout groups were much concerned with the lack of an acknowledged European collections strategy for digitised collections of local content. If users were to search on the Internet, how many digitised copies of, say, Shakespeare's works would he/she find? At the time of writing, a search of Google Scholar reveals 55 versions of the text of *Macbeth*. Is this level of duplication actually required? Could not the resources, which are devoted to digitising duplicates, be better spent on digitising materials which do not exist in digital form, in the context of an over-arching European collections strategy?

Tools to help decision making do exist. In 1998, I developed a decision making matrix to help with just these questions.<sup>4</sup> This matrix is reproduced below as Figure 3:

Shaded boxes in blue ask the questions which would address the issues around the duplication of the full text of *Macbeth* on the Internet. The fact

Fig. 3: Decision-making matrix on selection of materials for digitisation.

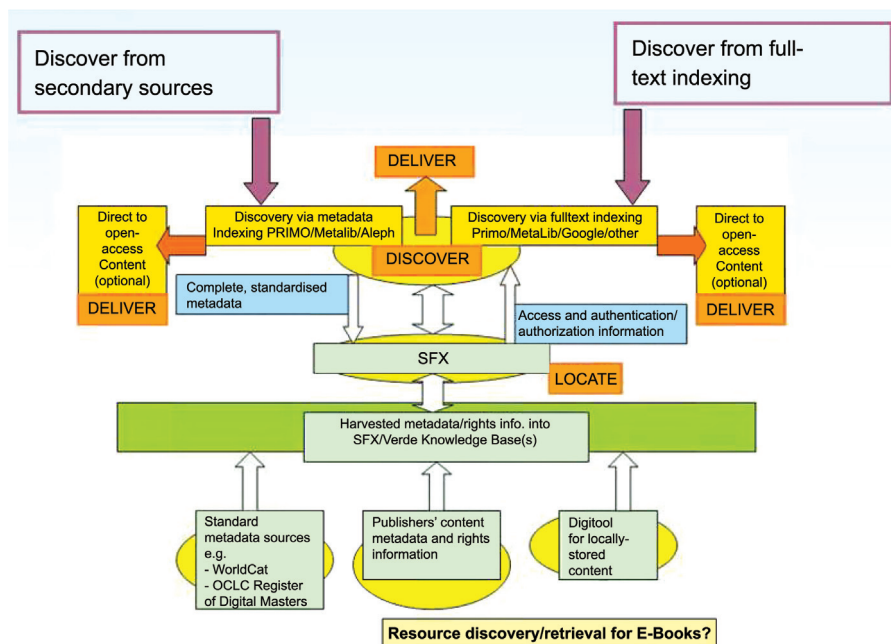
	Q1	Q2	Q3	Q4	Q5
Assessment	Is there user support?	What are local collection development policies?	Does this form a national or international contribution?	Does a similar product already exist elsewhere?	Is this conservation or preservation?
Gains	Does digitisation reduce wear on the originals or open up access?	Is the intellectual content of the work enhanced?	Is navigation easy?	Are disparate collections unified?	Is the use of damaged original material enriched?
Standards	Have suitable standards been followed?	Are the originals available from a variety of hardware platforms?	Is the software available and easy to use?	Does the metadata conform to agreed standards?	What are the archiving requirements?
Administrative issues	Do you have enough money?	Have copyright and rights issues been secured?	Does your institution have enough expertise?	Is there a partnership with a commercial provider?	Do the benefits justify the costs?

that these, or similar, questions have not been answered underlines the need, identified in the Copenhagen Workshop, for a more joined-up approach to the selection of content for digitisation across Europe.

## Discovery

How do users actually discover that materials have been digitised and are available in digital form? The Workshop considered a number of questions in their discussions. First, will discovery be by metadata or by the indexed full text of the work? The conclusion is that the question is not *either/or* but *both*. Both routes need to be available to the user. This is made apparent in the diagram below in Figure 4.

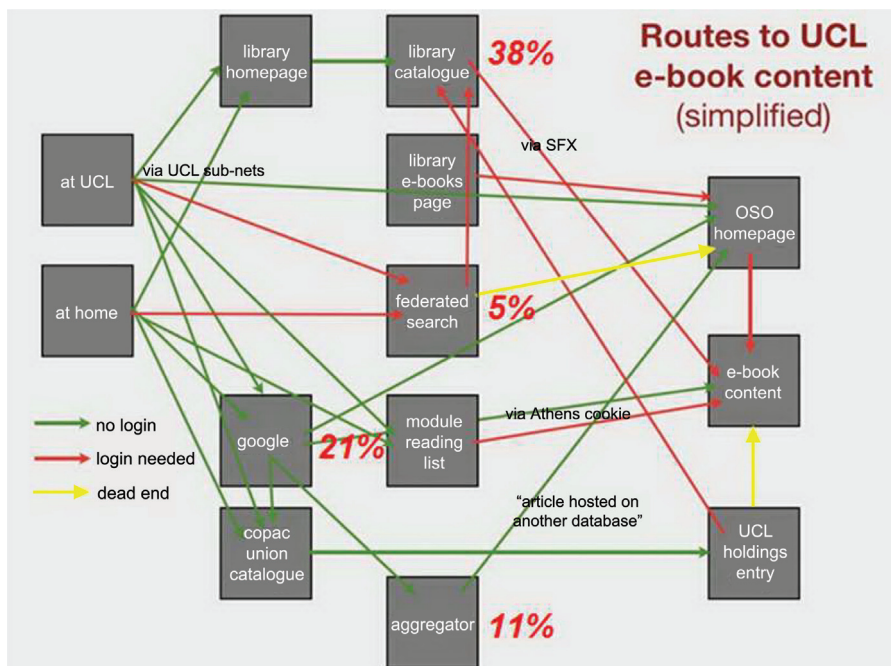
Fig. 4: Resource discovery/retrieval architecture for e-books in UCL.



The diagram is rather technical, as it illustrates from my own institution (UCL – University College London) ways in which UCL Library Services expects users to discover, locate and to receive delivery of e-books via the tools UCL has available through its suite of electronic library products. The two purple arrows at the top of the diagram illustrate that users can discover materials from secondary sources such as catalogues and listings (metadata) or from the indexed full text of material (e.g. via Google indexing).

Metadata is still a very important route for the discovery of digitised full text, a finding which has been underlined by the SuperBook project in UCL – a partnership between UCL’s School of Library, Archive and Information Studies, UCL Library Services and commercial publishers.<sup>5</sup>

Fig. 5: SuperBook's interim findings on routes to UCL e-book content.



Here in Figure 5 above, in the SuperBook project, UCL found that 38% of the users being surveyed found e-books via the library catalogue, whilst only 21% found them via Google. At the time of writing, therefore, in 2008 users seem wedded to the traditional role of the library and its catalogue as a place to discover books.

This finding makes it all the more important that European libraries act in a co-ordinated way to make metadata on digitised content available in a standard way and in a place/places where the user will be able (through one search) to find what he/she is looking for. One important finding, which the Workshop wished to emphasize, is that registries/persistent identifiers need to be established/adopted as a building block of the infrastructure of a digitised Europe.

As far as catalogues are concerned, the Workshop felt that there needs to be a high-level catalogue of the records of digitised content, *at least* at a European level. This would provide awareness of what is happening in European digitisation activity.

Some of the building blocks are in place – EROMM (European Register of Microform and Digital Masters) and the Registry of Digital Masters (RDM) from DLF/OCLC. In May 2006 a meeting was held in Washington DC to define the minimal requirements of content for the bibliographic record, which gives the description of the digitised surrogate and the digital master. In the meantime LIBER, OCLC and EROMM have concluded an agreement on sharing information on what is being preserved in digital form anywhere. Now OCLC has opened free access to the Registry of Digital Masters. Regular exchange of records between EROMM and the RDM is currently planned.

However, Workshop attendees wondered whether this architecture was optimal. The Workshop suggested that metadata harvesting from the creating library into a central, global, database might be a better option in the long term. There is a need for comprehensive coverage for adequate resource discovery. Clearly, this is an area where more discussion at a European level needs to take place.

### **Copyright/Intellectual Property Rights**

The Workshop identified some important developments in the field of copyright and intellectual property rights.<sup>6</sup>

Despite evident resistance to re-open the question, it is clear from discrepancies in the application of the Information Society (Copyright) Directive in different Member States that the Directive is far from satisfactory. The Copenhagen Workshop recommended that the Commission looks carefully at the November 2006 report it commissioned, *The recasting of copyright and related rights for the knowledge economy*, as a prelude to revising the Directive.<sup>7</sup>

There should be a uniform copyright regime all over Europe, on a level that restores the balance of interests and builds on the Information Society Directive.



Some Member States have provided that contracts may not override the statute law of copyright. This provision is in accordance with Article 15 of the Database Directive and therefore should be part of the *acquis communautaire*. It is essential in order to preserve (in line with recitals 14 and 34 of the Information Society Directive) an encouragement for creative, educational and scientific uses which do not harm the rights-holder. The Workshop believed that this is the most important single clarification that could be brought to the copyright regime in the Community.

For older material, the Commission should encourage the adoption of a pragmatic solution by identifying a date (say a moving wall of 50 years from the date of publication) before which it is acceptable to digitise without formality, provided that any emerging legitimate rights-holder receives compensation.

### **Partnerships**

LIBER and EBLIDA attenders in Copenhagen looked forward to the positive effects of the European Digital Library (EDL) in improving the current situation concerning digitisation of content, which currently reveals no agreement, scope or real co-ordination for digitisation activity across Europe. Co-ordination is not simply a European issue. European cultural objects are held in libraries, museums galleries and archives all over the world. It is important that Europe is able to come to a view on what and where relevant digitisation activity should take place. Digitisation is not simply an add-on to what a Library currently does. Digitisation has the ability to transform the experience of the user who will use materials from Europe's cultural institutions. This in turn will help to define the meaning of European citizenship.

Responsibility for overseeing this transformation is shared between the European institutions and the Member States. LIBER and EBLIDA are happy to work with the European Commission and other partners to achieve this consensus, but feel that the scope of TEL (The European Library) needs to be expanded beyond national libraries, to embrace major academic research libraries and other important library collections. LIBER can help play an important role here in brokering partnerships.

## **Recommendations to the European Commission and Future Work for LIBER**

Following the presentations, discussions and break out groups at the Workshop, a list of 25 Recommendations has been forwarded to the European Commission. The major issues addressed in the Recommendations are described above, and the full text is given in the Appendix below. They are divided into seven categories:

1. Vision for European Digitisation Activity
2. Content
3. Discovery
4. Copyright and Intellectual Property Rights
5. Metadata
6. Business Models
7. Digital Preservation

The Copenhagen Workshop shows that the time is ripe for Europe-wide discussions on digitisation of content to take place. One of the roles which LIBER is investigating, is a partnership with the EDL whereby LIBER provides a portal for digitised content from Europe's research libraries which is made available as part of the EDL (which has a remit far wider than content from libraries). This would be a happy partnership and enable LIBER to bring its expertise and experience to bear in the important area of European digitisation activity.

One immediate outcome from the Copenhagen Workshop is likely to be a series of follow-up workshops to build on the foundations which LIBER has laid in the first Workshop. At its Annual Conference in Turkey in 2008, on 1–5 July, LIBER will devote the Pre-Conference session to the theme of European digitisation. LIBER is also minded to have at least a further workshop in 2008, devoted to digitisation, in order to develop some of the themes which were aired in Copenhagen. It is an interesting time and LIBER is willing, and determined, to play its role in constructing a digital Europe for all European citizens.

## **Appendix**

Text of the Recommendations from the LIBER/EBLIDA Digitisation Workshop submitted to the European Commission.



LIGUE DES BIBLIOTHÈQUES EUROPÉENNES DE RECHERCHE  
ASSOCIATION OF EUROPEAN RESEARCH LIBRARIES

### **Recommendations from the LIBER/EBLIDA Digitisation Workshop, October 2007**

On 24–26 October 2007, 95 participants from 23 countries attended a Digitisation Workshop in the Royal Library in Copenhagen, organised by LIBER and EBLIDA. In the course of the three days, the attenders reached a number of conclusions and, following further work, LIBER and EBLIDA wish to make the following Recommendations to the European Commission.

LIBER and EBLIDA are convinced that concerted action on digitisation in Europe is required to support European researchers, students and citizens to gain access to Europe's scholarly and cultural heritage.

LIBER and EBLIDA are happy to work with the European Commission, the emerging European Digital Library (EDL) and other bodies to help realize this challenging vision.

### **Recommendations**

#### **1. Vision for European Digitisation Activity**

1. LIBER and EBLIDA look forward to the positive effects of the European Digital Library (EDL) in improving the current situation concerning digitisation of content, which currently reveals no agree-

ment, scope or real co-ordination for digitisation activity across Europe.

2. Co-ordination is not simply a European issue. European cultural objects are held in libraries, museums, galleries and archives all over the world. It is important that Europe is able to come to a view on what and where relevant digitisation activity should take place.
3. Digitisation is not simply an add-on to what a Library currently does. Digitisation has the ability to transform the experience of the user who will use materials from Europe's cultural institutions. This in turn will help to define the meaning of European citizenship.
4. Responsibility for overseeing this transformation is shared between the European institutions and the Member States
5. LIBER and EBLIDA are happy to work with the European Commission and other partners to achieve this consensus.

## **2. Content**

6. There is a need for formal selection criteria for the identification of material to be digitised, as this would aid co-ordination across European countries.
7. There is a corresponding need for a European Collections Strategy. LIBER and EBLIDA would be happy to work with the Commission and the EDL to take this work forward.
8. The construction of a critical mass of digitised content needs to be achieved sooner rather than later and a road map should be constructed identifying how this will be accomplished.
9. It is important that the wishes of the user are part of this decision-making process and steps need to be taken to identify what the European citizen actually wants in terms of digitised content.

## **3. Discovery**

10. Funding is needed to allow for more conversion of non-automated metadata for library/archive materials. Catalogues which remain in analogue form effectively obstruct access to the material they represent. Libraries need to assess their priorities in this area, seeking partnership funding where appropriate to help them.

11. There should be a 'European Resource Space', a portal to all European resources, which unites the offerings of EDL, TEL, EROMM, RDM and other European databases.
12. The scope of TEL needs to be expanded beyond national libraries, to embrace major academic research libraries and other important library collections. LIBER can help play an important role here in brokering partnerships.
13. More encouragement is needed to bring the work of archives, museums, and libraries closer together.

#### 4. Copyright and Intellectual Property Rights

14. Despite evident resistance to re-open the question, it is clear from discrepancies in the application of the Information Society (Copyright) Directive in different Member States that the Directive is far from satisfactory. We recommend that the Commission looks carefully at the November 2006 report it commissioned, *The recasting of copyright and related rights for the knowledge economy*, as a prelude to revising the Directive.
15. There should be a uniform copyright regime all over Europe, on a level that restores the balance of interests and builds on the Information Society Directive.
16. Some Member States have provided that contracts may not override the statute law of copyright. This provision is in accordance with Article 15 of the Database Directive and therefore should be part of the *acquis communautaire*. It is essential in order to preserve (in line with recitals 14 and 34 of the Information Society Directive) an encouragement for creative, educational and scientific uses which do not harm the rights-holder. We believe that this is the most important single clarification that could be brought to the copyright regime in the Community.
17. For older material, the Commission should encourage the adoption of a pragmatic solution by identifying a date (say a moving wall of 50 years from the date of publication) before which it is acceptable to digitise without formality, provided that any emerging legitimate rights-holder receives compensation.

## **5. Standards and Policies: Metadata**

18. Registries/persistent identifiers need to be established/adopted as a building block of the infrastructure of a digitised Europe.
19. Guidelines are required to support selection and use of the full spectrum of metadata; minimum standards need to be defined; although it should be noted that retrieval of content is both by assigned metadata and also by the indexing of the full-text
20. Libraries should support the development of designated EC Centres of Competencies.
21. There needs to be a high-level catalogue of the records of digitised content. This would provide awareness of what is happening in European digitisation activity. Some of the building blocks are in place and WorldCat is aiming to provide comprehensive coverage in this area. Libraries need to be encouraged to create local records of digitised materials, for harvesting by services such as WorldCat.

## **6. Standards & Policies: Business Models**

22. The European Commission does not currently fund the creation of digitised content, but it should re-consider this position. The Commission recommends the formation of Public/Private partnerships but Public/Private partnerships require funding in order for public institutions to make an offering.
23. Private funding of digitisation activity may restrict access by the user, which is not in the interests of the European citizen. LIBER and EBLIDA do not see this as the preferred route for funding the digitisation of content.

## **7. Digital Preservation**

24. National activities across Europe – e-Depot (Netherlands), Nestor (Germany), ABES (France) and EU projects such as PLANETS should form the basis of a more co-ordinated approach to digital curation activities across Europe. LIBER would be happy to contribute to a dialogue in Europe on how this join-up can be achieved.
25. What is the future model for European digital preservation services? There needs to be a road map to help identify the way forward. Trust is a key issue in identifying the journey.

*The Organisations making this Submission*

LIBER is the principal association of the major research libraries of Europe. See <http://www.libereurope.eu>; EBLIDA is the European Bureau of Library, Information and Documentation Associations. See <http://www.eblida.org/>.

## **Websites Referred To In The Text**

EBLIDA, European Bureau of Library, Information and Documentation Associations, <http://www.eblida.org>

EROMM, European Register of Microform and Digital Masters, <http://www.eromm.org/>

EU Information Society (Copyright) Directive, [http://en.wikipedia.org/wiki/EU\\_Copyright\\_Directive](http://en.wikipedia.org/wiki/EU_Copyright_Directive).

JISC, Joint Information Systems Committee, <http://www.jisc.ac.uk>

LIBER Annual Conference 2008, [http://www.ku.edu.tr/ku/index.php?option=com\\_content&task=view&id=2011&Itemid=2917&lang=en](http://www.ku.edu.tr/ku/index.php?option=com_content&task=view&id=2011&Itemid=2917&lang=en)

RDM, Registry of Digital Masters from DLF/OCLC, <http://purl.oclc.org/DLF/collections/reg/OCLCservice>.

TEL, The European Library, <http://www.theeuropeanlibrary.org/portal/index.html>.

Workshop report on the JISC website, <http://www.jisc.ac.uk/news/stories/2007/10/digi.aspx>.

## **Notes**

---

<sup>1</sup> The powerpoints from the Workshop are mounted on the LIBER website at <http://www.libereurope.eu/node/264>

<sup>2</sup> The podcast is available at <http://digitisation.jiscinvolve.org/>

<sup>3</sup> See [http://digital.casalini.it/retreat/retreat\\_2006.html](http://digital.casalini.it/retreat/retreat_2006.html)

<sup>4</sup> See <http://eprints.ucl.ac.uk/archive/00000492/>

<sup>5</sup> See <http://www.ucl.ac.uk/slais/research/ciber/superbook/>

<sup>6</sup> I am very grateful to Toby Bainton, Secretary of SCOUNL in the UK, for providing the following text on copyright and intellectual property rights.

<sup>7</sup> See [http://ec.europa.eu/internal\\_market/copyright/docs/studies/etd2005imd195recast\\_report\\_2006.pdf](http://ec.europa.eu/internal_market/copyright/docs/studies/etd2005imd195recast_report_2006.pdf)