

# **A Comparative Study of e-Journal Archiving Solutions**

---

**A JISC Funded  
Investigation**

Terry Morrow  
(Tee Em Consulting)

Neil Beagrie  
Maggie Jones  
(Charles Beagrie Ltd)

Julia Chruszcz  
(Top Class Computer  
Technologies Ltd)

Final Report - May 2008

© 2008 JISC Collections  
The authors have asserted  
their moral rights in this  
work.

## Table of Contents

1.	Foreward .....	3
2.	Executive Summary.....	4
3.	Introduction.....	6
4.	Context.....	8
5.	Approaches to e-journal preservation .....	9
5.1.	Overview .....	9
5.2.	Perpetual access or Long-term Preservation?.....	10
5.3.	Costs and Benefits .....	11
5.4.	Current Initiatives .....	13
6.	Publisher licensing and legal deposit .....	14
6.1.	Publisher Licensing .....	14
6.2.	Legal deposit .....	14
6.3.	Open Access Repositories .....	15
7.	Comparison of Six Current e-Journal Archiving Programmes .....	16
7.1.	LOCKSS .....	16
7.2.	CLOCKSS.....	17
7.3.	Portico .....	17
7.4.	e-Depot .....	18
7.5.	Electronic Collections Online (ECO).....	19
7.6.	British Library e-journal Digital Archive.....	19
8.	Practical experience of e-journal archiving solutions.....	21
8.1.	Views on continuing with LOCKSS after 2-year subsidised trial .....	21
8.2.	Views on e-journal archiving and costs .....	22
8.3.	UK LOCKSS Second Survey Results.....	24
8.4.	Manchester Workshop (Part 1) .....	25
9.	Four Scenarios .....	27
9.1.	Introduction and purpose .....	27
9.2.	Scenario 1 Library cancels subscription and requires access to past subscribed issues ..	27
9.3.	Scenario 2 E-journal or its past issues no longer available from the publisher .....	29
9.4.	Scenario 3 Publisher has ceased operation and e-publication is no longer possible .....	30
9.5.	Scenario 4 Catastrophic failure of publisher's operations/servers .....	31
9.6.	Manchester Workshop (Part 2) .....	33
10.	Criteria for judging relevance and value of new archiving initiatives .....	34
11.	Observations and Conclusions .....	36
12.	Recommendations.....	38
	Appendix 1 Table of Publisher Participation .....	39
	Appendix 2 Publishers in more than one programme .....	40

## 1. Foreward

This report is the result of a call by the JISC, issued in January 2008, for a Comparative Study of e-Journal Archiving Solutions. Recognising the emergence of other approaches to e-journal archiving during the 2 years of the UK LOCKSS pilot programme, the Invitation to Tender asked for a report that “will be published for wide use by institutions to inform policies and investment in e-journal archiving solutions.” The ITT also stated that the report should “also inform negotiations undertaken by JISC Collections and NESLi2 when seeking publishers’ compliance to deposit content with at least one e-journal archiving solution.”

The contract was awarded to Tee Em Consulting, and the work was carried out during March and April 2008. The consultants working on the report were Terry Morrow (Tee Em Consulting), Neil Beagrie and Maggie Jones (Charles Beagrie Ltd) and Julia Chruszcz (Top Class Computer Technologies Ltd).

The report has been created by reviewing existing published work in this area, by consulting with key players including existing service providers, and by holding a one day workshop at the University of Manchester on 14<sup>th</sup> April. Attendees at that workshop, who also assisted by reviewing early drafts of the report, included:

Chris Ashton	University of Sheffield Library
Joy Aspinall	Manchester Metropolitan University
Sue Bate	University of Manchester Library
Carolyn Benny	Liverpool John Moores Library
Frances Boyle	Digital Preservation Coalition
Michael Breaks	CURL/SCONUL Scholarly Communications Group
Michael Debenham	University of Keele Library
Michael Emly	University of Leeds Library
Lorraine Estelle	JISC Collections
Neil Grindley	JISC - Digital Preservation
Paul Harwood	Content Complete
Eileen Hiller	University of Huddersfield
Julie Hitchen	University of Central Lancashire Library
Emily Jesper	ALPSP
Michiel Kolman	Publishers Association
Allison Larkins	University of Huddersfield
Annette Lawrence	Lancaster University Library
Bernie Mellody	Manchester Metropolitan University
Maureen Richardson	Edge Hill University Library
Sarah Robbins	Liverpool John Moores Library
Chris Senior	University of Leeds Library
Ruth Silman	University of Manchester Library
Graham Stone	University of Bolton
Sarah Thompson	University of York

Our thanks go to John Rylands Library, University of Manchester for kindly hosting the workshop on 14<sup>th</sup> April and to all who assisted in the creation of this report and who helped to ensure that it is correct, relevant and readable.

Terry Morrow, Neil Beagrie, Maggie Jones, Julia Chruszcz  
May 2008

## 2. Executive Summary

Although there are many obvious benefits that accrue from publishing and accessing academic papers through the internet, there are costs and challenges associated with long term preservation and access which urgently need to be addressed. Finding solutions to these is the shared responsibility of all in the information chain, including authors, publishers, repository managers, librarians, subscription agents and aggregators.

**Libraries** should take the initiative in raising awareness and understanding within their institutions of the importance of perpetual access to, and archiving of, e-journals. They should work with policy makers to ensure that the recognition and support for e-journal access and archiving gets incorporated into both library and institution-wide strategies and policies for research, teaching and learning, all of which are increasingly dependent on access to this material.

**The British Library** and other legal deposit libraries are in a unique position and should continue to develop solutions that can provide a safety net, at least for all electronic journals that originate in the UK.

**Publishers**, who now deliver electronic journal content directly to readers, must acknowledge that they have a special responsibility to ensure that this material is secure for the long term future, and offer subscribing libraries clear information on their archiving and post-cancellation access policies, including which e-journal archiving initiatives they support.

**Negotiators** should use contract conditions to obtain commitments from publishers to support community recommended e-journal archiving initiatives.

**E-journal archiving solutions** must demonstrate sound financial and organisational sustainability and technical insight and expertise, in order to earn the trust of, and acceptance by, librarians and the majority of academic publishers.

This study makes the following recommendations.

1. When negotiating **NESLi2 agreements**, **JISC's negotiators** should take the initiative by specifying archiving requirements, including a short-list of approved archiving solutions.
2. To help quantify the insurance risk and the necessary appropriate investment, bodies representing **publishers** and other trade organisations should gather and share statistical information on the likelihood of the trigger events outlined in this report.
3. **Post cancellation access** conditions should be defined in the licensing agreement between libraries and publishers. Publishers should be strongly encouraged to cooperate with one or more external e-journal archiving solutions as well as provide their own post-cancellation service (at minimal cost).
4. The publisher (or subscription agent) should state their policy on **perpetual access** under the four scenarios described in section 9.
5. When **titles are sold** on to other publishers, the Transfer Code of Practice (see section 9.3.) should be followed.
6. **Archiving service providers and publishers** should work together to develop standard cross-industry definitions of trigger events and protocols on the conditions for release of archived content. Project Transfer is a potential exemplar. The ground rules

for any post-trigger event negotiation should be clear and transparent and established in advance.

7. Archive service providers must provide greater clarity on **coverage** details, including not only publishers and titles, but also the years and issues included in the archive.
8. Using the scenarios outlined in this report, libraries should carry out a **risk assessment** on the impact of loss of access to e-journals by their institution, and a cost/benefit analysis, in order to judge the value and relevance of the archiving solutions on offer.
9. Relevant UK bodies and institutions should use whatever influence they can bring to bear to ensure that archiving solutions cover publishers and titles of particular value to **UK libraries**.
10. The findings of this study should be **reviewed and updated** at regular intervals to reflect continuing developments in the field of e-journal archiving and preservation.

### 3. Introduction

**ephemeral** *adj* lasting or of use for only a short time; transitory

Electronic versions of academic journals first appeared in the mid-1990's, initially as an alternative and parallel method for the delivery of paper journals. Over the last decade, e-journals have become increasingly popular and have begun to oust paper journals as the normal mechanism for delivering papers on academic research to the community.

The "e" of e-journals of course stands for "electronic". But there is a distinct danger that the "e" could equally come to mean "ephemeral" unless we take active steps to preserve the bits and bytes that increasingly represent our collective knowledge. It is easy to think of recent history in other areas which should give us pause for thought. Obsolete hardware examples include 8 inch and 5.25 inch floppy discs, Betamax video tapes, and even reel-to-reel audio tapes. Software development is not immune to obsolescence, though the pace of change tends to be slower and more gradual than hardware. The same applies to data formats.

This experience should be telling us that we need to be looking ahead if our growing collections of e-only research publications are still going to be readable by future generations of scholars 10, 50 or 100 years from now.

Layered on top of this threat of technical obsolescence is the change of business model which results in libraries buying licences to enable their users to gain network access to a publisher's server, rather than collecting, binding and shelving paper journals. The only original copy of an issue of an e-journal now tends to be held by the publisher. But the long term preservation of that copy is arguably of greater importance to the library and research communities than the publisher. So we may need new models and sometimes organisations to ensure safe custody of these objects for future generations.

The complex issues associated with continued access to licensed e-journals has been the subject of much research and discussion in recent years. A study commissioned by JISC in 2003 investigated available options for fulfilling the clauses relating to continued post cancellation access in the NESLi2 Model Licence and found there were few practical options at that stage but promising developments in terms of nascent archiving solutions<sup>1</sup>. Three years later a CLIR survey analysed twelve archiving solutions which had met seven indicators of viability<sup>2</sup>. JISC commissioned a review of the CLIR survey intended to assess its relevance to the UK<sup>3</sup>, which in turn articulated high level principles for a national approach to e-journal archiving. This was discussed at a BL/DPC/JISC Workshop held in March 2007 and led to a JISC Briefing Paper. Related work includes a JISC funded scoping study on an e-Journals Registry.<sup>4</sup>

This study builds on previous work and reviews current thinking on the long-term preservation of formally published e-journals and e-journal articles. It also makes reference to the growing interest in using institutional repositories to store copies of some of this material. It is written primarily as a practical guide for a UK audience of practising librarians, though others may find the content useful.

Part of the context for this study is the recent JISC-supported two year trial of the UK LOCKSS Pilot Programme. 30 institutions have taken part in the trial which is being assessed in parallel with this report by Evidence Base. Some of the feedback from the assessment is included in this study. Since the UK LOCKSS Pilot was launched, other initiatives have emerged, and five of these, relevant to UK libraries, are also described and assessed to give a more complete picture.

Through the discussions that have taken place with sector professionals during the course of the study it is clear that the implications of “e” only and the licence assurances about continued access are not as straightforward as librarians had hoped. As the percentage of “e-only” journals increases librarians are looking for ways to protect their investment over time. They are looking for strategies that will provide cost effective means to ensure perpetual access. The study workshop provided some valuable insights into current and emerging priorities within institutions and these are described more fully in section 8.4.

This is an area which is clearly in its early stages of development, with a number of different initiatives and approaches emerging. Right now, and for the near future, it is likely to present a confusing and not wholly reassuring picture to those professionals trying to make sense of what is happening and looking for simple, clear-cut guidelines. Different communities have differing perspectives and priorities, and there is a distinction to be drawn between libraries wanting to provide perpetual access to previously licensed material (as was always the case with printed journals), and preservation for its own sake. None of the current initiatives is likely to yet fulfil all the access and archival needs of a modern library. But there are developments and emerging approaches that deserve support and there is good reason to believe that by investing in well thought through and sustainable archiving solutions, our current electronic collections, and our investment in access to them, will not be ephemeral but long lasting.

## 4. Context

Despite major concerns about e-journal archiving and the high risk that e-journals may be lost, the trend towards e-only access for scholarly e-journals is continuing. In 1996, 131 refereed e-journals were identified; by 2003, this had climbed to 14,338, representing 62% of the 23,187 peer reviewed scholarly journals. Although most are still parallel print and electronic, a British Library commissioned study by Electronic Publishing Services Ltd (EPS) concluded that half of all serial publications will become online only by 2016. A survey of ARL libraries revealed that e-journals as a percentage of total serial expenditure had risen from 5% of the total in 1995 to 42% of the total in 2004<sup>5</sup>. While concerns about long-term archiving remain an issue, libraries are increasingly cancelling their print subscriptions and moving to e-only access.

The convenience and currency of online access has made this an attractive option for many scholarly users. Libraries and publishers have responded to this demand unable to wait for assured archiving and preservation solutions. As this trend is proceeding, finding resolutions to the archiving and preservation of scholarly e-journals has assumed even greater urgency and significant efforts have been made in exploring viable options for e-journal archiving which can provide greater assurance of continued access into the future.

Although both legal deposit legislation (see section 5.2) and institutional repositories (see section 6.3) are important developments, neither of them can reasonably be expected to provide practical solutions for libraries licensing access to e-journals.

In the UK, the archiving clauses in the NESLi2 licence have provided a measure of security for libraries but in the absence of trusted repositories charged with managing e-journals, these have provided largely theoretical assurance.

A further complication is the increasingly international nature of the academic publishing business. Many publishing houses no longer have a strong association with any one country. Ideally, therefore, trusted archiving solutions need to reflect this international perspective by providing a global dimension. All current initiatives originate in particular countries, though there are signs emerging of cross-border collaboration and cooperation. For example Portico has recently reached agreement with Koninklijke Bibliotheek, the National Library of the Netherlands and operators of e-Depot, for them to host an offline copy of the Portico archive. LOCKSS of course, by virtue of its different model, is inherently more distributed.

Thus there is a pressing requirement for trusted and internationally accepted repositories focussed on archiving and preserving e-journals, which are independent of publishers, and which offer solutions which can safeguard content while sharing costs between libraries and publishers equitably. While the concerns of libraries are much the same as they were when the JISC consultancy on e-journals archiving reported in 2003, there are now a clearer set of options emerging. Over the past few years, a number of promising initiatives have been developed which provide much better prospects for continued access to licensed e-journal content and which offer cost-effective solutions for libraries and publishers. This study examines some of them in more detail.

## 5. Approaches to e-journal preservation

### 5.1. Overview

There are two broad techniques for preserving e-journals. One is to preserve the source files which constitute a journal publication. The other is to capture the presentation files (referred to as “rendition archiving”). There are advantages and disadvantages to each and it should also be noted that there can be significant variations in how each approach is implemented. The advantages of source file preservation is that it is very complete (and likely to include more content than appears in the journal); is received directly from the publisher and is frequently delivered or converted to a few normalized formats facilitating long-term preservation. The disadvantages are that it requires a large upfront investment; there is no assurance that the archive will actually be needed; and the presentation will almost certainly differ from that of the publisher.

The advantages of harvesting presentation files (rendition archiving) are that it is possible to retain the look and feel of the publication and initial costs are likely to be lower. The disadvantages of this technique are that it may be more difficult to preserve the content over time (for example, a strategy for the large scale migration of presentation files from one format to another is still untested).

A conclusion of seven Mellon-funded e-journal projects in 2002 was that there is a need to support both approaches:

“Although experience might later tell us that one approach is better suited than the other for certain kinds of material, it would not now be useful to think of them as competing approaches. We have to get used to the idea that overlapping and redundant archiving solutions under the control of different organizations with different interests and motives in collecting offer the best hope...It would be unwise at the outset to expect that only one approach would be sufficient.”<sup>6</sup>

More than four years later, it was still considered to be unwise to try to select the ‘best’ approach, as the CLIR survey of e-journal archiving notes:

“At this point, it is impossible to say which of these two approaches is the better solution to archiving. Those programs that solicit both source files and rendition copies of e-journal content (Pub Med Central, Portico, KB E-Depot, Kopal/DDB) probably are the safest archiving solution – but at a potentially greater cost.”<sup>7</sup>

Although archiving of scholarly e-journals has attracted a great deal of effort and attention over the last several years, there are a number of complicating factors which have combined to make it particularly challenging to develop effective archiving solutions.

- Changing roles and ambiguous responsibilities in a digital environment. Libraries receive online access for their subscription fee, not a physical copy as they did with print, leaving continued management of the content under the control of the publisher. This has left libraries feeling vulnerable with little or no control over the material they are purchasing and publishers needing to consider undertaking a role they did not have in the print world.
- Given the strong role of the publisher in e-journal archiving, issues around comprehensive coverage and continuity have needed to be addressed. For example, a journal title may cease or change publisher, publishers go out of business and /or merge with others. This means that, even when a publisher is part of an archiving program, it is not necessarily the case that all of their publications are covered, or

that they will all remain under the care of the archiving program if the business relationship changes.

- The link between archiving and access and how this is perceived and articulated by both libraries and publishers can cause confusion, misunderstanding, and sometimes mistrust. When publishers offer “perpetual access” does that mean only for the duration of current subscriptions or beyond? When libraries demand preservation, do they need assured preservation in perpetuity of the content they are licensing or do they only need to be assured of seamless remote online access for as long as their clients require it?
- Who should pay the costs of managing e-journals over time? Whoever pays, the fact that it needs to be factored into overall costs of e-journals was recognized by Don Waters in 2002 when he noted that:

“One aspect of the reassessment that is under way is a growing awareness that archiving has not yet been factored into the overall costs of the system, and if electronic publishing is to be taken seriously, it must be.”<sup>8</sup>

Finding an arrangement which is realistic, reasonable and equitable is by no means straightforward and much effort has been expended in recent years in developing and testing a range of options.

Whatever the differing needs, neither short to medium term ongoing access, nor long-term archiving, can be achieved without investment in, and attention to, managing the content of e-journals in such a way that the risk of loss through technical, organisational or legal problems is at least minimized if not eliminated.

In a world where it is impossible to provide cast iron guarantees, it becomes increasingly important to be able to manage risks appropriately. The risk of sudden loss of access to an e-journal is quite high, not just because of temporary technical problems or catastrophic technical failure, but because of factors which threaten continuity of care. For example, sudden changes in business relationships, such as publishers merging with others or ceasing operations entirely. This means that leaving archiving arrangements, even for the short term, entirely with the publisher, is a high risk strategy.

## **5.2. Perpetual access or Long-term Preservation?**

When discussing e-journal archiving, it is important to understand the distinction between perpetual access, and long-term preservation. These terms were defined in a JISC Briefing Paper on e-Journals Archiving and Preservation published in 2007:

### *Perpetual Access*

‘Perpetual access is most commonly associated with e-journal licence clauses designed to provide assurance of continued access to subscribed material in certain circumstances, including post-cancellation...

### *Long-term preservation*

Long-term preservation refers to the processes and procedures required to ensure content remains accessible well into the future...’<sup>9</sup>

In other words, continuing or perpetual access is an attempt to replicate the situation with paper journals where a library receives, makes available and preserves the material for ongoing reference, regardless of whether or not the subscription is continued. Long term preservation, on the other hand, can be viewed as an issue, not just for the subscribing library, but for society as a whole, ensuring that the scholarly record continues to be

accessible to future generations of scholars and mankind in general. This becomes a significant challenge as storage and delivery technologies evolve and develop.

The distinction is important as a commitment to preserve material for the long-term (as national libraries undertake) does not necessarily equate to managing widespread network access with the necessary authentication systems required to make that content available online to subscribing institutions in the event of a specific trigger. And it is unlikely that, for most library clientele in most disciplines, it would be acceptable for them to be directed to onsite use of material at the relevant national library, even if the specific title was included in the national libraries' archiving remit.

If this distinction is accepted, then inclusion of a title in a national archive should be seen as an additional bonus rather than a solution to service requirements of individual libraries who have licensed access to e-journals.

### **5.3. Costs and Benefits**

Any e-journal preservation process is going to cost money. The costs include, among other things, storage hardware systems, processing and retrieval software (all of which require regular maintenance and updating), and people to watch over and develop the systems and services. They also include promoting the service to, and negotiating with, publishers and libraries.

The costs may be covered by subscription income/membership fees from libraries or publishers, government grants, or charitable foundations or other similar organisations, or some combination. Whoever is paying the bills is going to need to weigh up the costs against the risks and potential benefits. In general, not-for-profit and publicly-funded organisations are probably less likely to have conflicts of interest in exercising stewardship of a nation's heritage than commercial organisations, though it would be sensible to encourage the latter to also play their part.

Many of us who own houses, insure the property against certain risks including, for example, total destruction by fire, flood or other disaster. Here the risk is low, but the consequences could be catastrophic. We might also take out insurance against losing our credit card. Here the risk is moderate, but the consequences are low to moderate (provided we notify the card company quickly, our potential losses are small). Finally, we might insure a pair of contact lenses, or spectacles against loss or damage. Here the risk is high but the consequence of having to buy a new pair is fairly low on the scale of everyday events.

Risk management is also a useful concept when it comes to e-journal preservation. It can be used to assess the real value of the insurance offered by an e-journal archiving solution. Figure 1 outlines four potential scenarios ranging from high risk / high impact to low risk / low impact. It suggests a few examples to illustrate the circumstances which might lead institutions to decide on whether or not to participate in a particular solution.

<p><b>HIGH RISK / HIGH IMPACT</b>  <b>e.g:</b></p> <ul style="list-style-type: none"> <li>▪ Merger of two publishers – core title no longer available</li> <li>▪ Closure of small scholarly/open access publisher which publishes niche core titles</li> </ul> <p><b>Participation in one or more e-journal archiving solutions is highly recommended</b></p>	<p><b>HIGH RISK / LOW IMPACT</b>  <b>e.g:</b></p> <ul style="list-style-type: none"> <li>▪ Loss of access to non core open access title</li> <li>▪ Budget cuts mean cancellation of selected non core titles</li> <li>▪ Merger of two publishers – non-core title no longer available</li> </ul> <p><b>Participation in at least one e-journal archiving solution is desirable <i>unless</i> there is a strong commitment to preserve the scholarly record on behalf of a wider community when it becomes <b>highly recommended</b>.</b></p>
<p><b>LOW RISK / HIGH IMPACT</b>  <b>e.g:</b></p> <ul style="list-style-type: none"> <li>▪ Need to cancel core titles by major research library</li> <li>▪ Closure of major publisher</li> <li>▪ Takeover of major publisher and their key titles are no longer available.</li> </ul> <p><b>Participation in one or more e-journal archiving solution is highly recommended for large research organisations</b></p>	<p><b>LOW RISK / LOW IMPACT</b>  <b>e.g:</b></p> <ul style="list-style-type: none"> <li>▪ Closure of a major publisher to which a library subscribes very few titles</li> <li>▪ Closure of large open access publisher which includes non-core titles</li> <li>▪ Merger of two publishers – title of little research interest to the institution is no longer available</li> </ul> <p><b>Participation in at least one e-journal archiving solution is optional <i>unless</i> there is a strong commitment to preserve the scholarly record on behalf of a wider community when it becomes <b>highly recommended</b>.</b></p>

**Figure 1: Table of Risk and Impact**

Once a decision is made that signing up to or supporting an archiving solution is recommended, which solution is selected will depend on a number of factors including:

- The range of content covered (number and type of publishers, titles, year ranges)
- Costs and the basis of charging (eg by size or type of institution)
- Whether post-cancellation access is offered, and if so, how this is provided
- Whether immediate backup access is possible in the event of short term problems such as temporary failure of a publisher’s server
- Size and focus of the institution (research, teaching, mixture)
- The possible existence of teams or departments within an institution with special needs (eg a highly rated research team within an otherwise predominantly teaching university).
- The level of expertise of IT support within the library or the institution as a whole
- Usefulness of the provision of on-site access (eg at BL) when this is the only option offered.

#### 5.4. Current Initiatives

The “Metes and Bounds” report<sup>10</sup> examined a dozen different e-journal archiving initiatives which the authors were aware of when the report was written and which all met their seven indicators of viability. Of these, six have been selected for this report as being of likely relevance to UK academic libraries. The following is a brief introduction. Section 7 contains a detailed analysis and comparison of the solutions.

- **LOCKSS**  
LOCKSS (Lots Of Copies Keep Stuff Safe) enables libraries to play an active part in preservation. LOCKSS libraries maintain “LOCKSS boxes” which are used to store copies of all e-journal material that the library subscribes to (subject to publisher agreement). LOCKSS is therefore a closer analogue to paper distribution than any of the other solutions.
- **CLOCKSS**  
CLOCKSS (Controlled LOCKSS) uses LOCKSS software to create a “dark archive”. Access is enabled when the publisher authorises it, or when there has been substantial disruption to access via the publisher and the Board determines that the content can become freely available.
- **Portico**  
Portico is specifically designed as a third-party electronic archiving service. Portico’s role is as a permanent dark archive. Access is only permitted when there has been substantial disruption to access via the publisher.
- **e-Depot**  
e-Depot was established in 2003 by the government-funded Dutch national library, Koninklijke Bibliotheek (KB). Content includes Dutch university repositories, websites, and other national large-scale digitisation programmes (eg Dutch newspapers). It originally included only Dutch e-journals, but later added an international dimension including worldwide scientific publications. Access is by agreement with the publisher, but is generally restricted to on-site viewing for private research.
- **OCLC Electronic Collections Online (ECO)**  
ECO provides access to wide range of titles and publishers. ECO promises long-term access to subscribed content, subject to the library continuing to pay OCLC’s access fee.
- **British Library e-Journal Archiving Programme**  
The British Library has been building a Digital Object Management System (DOM) which can store and manage all BL’s digital content. BL began ingesting content from selected publishers into this system in 2007. Access depends on publisher requirements.

## **6. Publisher licensing and legal deposit**

### **6.1. Publisher Licensing**

Solutions are needed which allow publishers to focus on what they do best (i.e. adding value to the content they deliver), without compromising libraries' needs to be assured of continued access regardless of changes to business models and relationships. The assurance of continued access over time should not be hostage to the vagaries of commercial business decisions.

The NESLi2 Model Licence has provided a measure of reassurance for UK HE libraries for a number of years, embedding archiving within the license clauses. However, in the absence of specific preferred archiving solutions, the presence of an archiving clause alone is not sufficient. As the following quote notes, "Although some ...licenses now recognize that libraries have permanent rights to use electronic journal content, these rights remain largely theoretical. If a publisher fails to maintain its archive, goes out of business, or for other reasons, stops making available the journal on which scholarship in a particular field depends, there are no practical means in place for libraries to exercise their permanent usage rights...." <sup>11</sup>

It is therefore timely that trusted archiving solutions are beginning to emerge that could help fulfil the intent of the clauses.

The current NESLi2 licence offers three options for continued access to e-journal content following cancellation. These are:

- Continued access from the Publisher's server.
- Supplying archival copies of the licensed material to the library or to a central archiving facility operated on behalf of the UK HE/FE community or other archival facility.
- Supplying archival copies via ftp protocol.

All three options must of course rely on the cooperation of the publisher but the first one entrusts the content entirely to the publisher while the other two offer an option for the subscribing library to undertake preservation responsibility. The second option also includes the possibility of supplying an archival copy to a central or other archiving facility as well as to the individual library. Neither handing over complete responsibility to the publisher nor leaving it to individual libraries is seen as optimum and so the "central archiving facility or other archival facility", or libraries working together to build collections, offers the best assurance.

### **6.2. Legal deposit**

The Legal Deposit Libraries Act 2003 extended previous legislation to include electronic publications in the UK under secondary legislation. Many other national libraries around the world are also regarding the deposit of e-publications as a logical extension to their mandate to preserve their national heritage, though this role is not always enshrined in legislation. As legal or voluntary deposit will be expected to include some e-journals, it is tempting to believe that this will provide at least a partial solution for UK HE and FE institutions who wish to move to e-only e-journal subscriptions.

However, while legal deposit is undoubtedly a crucially important component of the digital preservation landscape, it has limited applicability to licensed e-journal content for four interrelated reasons:

- It is anticipated that, once implemented, legally deposited publications will only be accessible onsite at the appropriate national library, as is the case with deposited print materials.
- Despite several national libraries undertaking detailed preparations for legal deposit over a number of years, it is still largely untested as an effective mechanism for preserving e-publications, and it is generally accepted that it is a sensible precaution to invest in alternative solutions.
- Legal deposit legislation cannot be entirely comprehensive. It is not yet enshrined in legislation in many countries and remains subject to voluntary deposit by the publisher.
- Finally, the concept of place of publication is increasingly blurred in the digital environment, not least for e-journals, which are often multi-national in terms of management and service delivery.

### **6.3. Open Access Repositories**

The rapid development of institutional and open access repositories has been another important development which may appear to offer a preservation solution. These repositories generally include a wide range of content produced by their host institution, including, though not usually limited to, research articles and papers. However, two major factors militate against the assumption that institutional or open access repositories will provide an effective e-journal archiving solution.

Firstly, despite a powerful momentum, much peer-reviewed research literature still remains outside the realm of institutional and open access repositories. Secondly, the emphasis to date has been, unsurprisingly, on populating the repositories, rather than preserving their content, so it cannot be safely assumed that electronic research articles deposited in institutional and open access repositories are automatically preserved.

There are initiatives which are shaping the development of the archiving component of repositories in the UK. For example, JISC funded projects such as SHERPA DP<sup>12</sup> and PRESERV<sup>13</sup> are both investigating models for archiving and preserving content in distributed institutional repositories. This research will pave the way for more coherent, coordinated preservation strategies to safeguard the valuable content being held in UK institutional repositories. This is a rapidly evolving environment and justifies continued investment but should not be regarded as a substitute for other solutions designed to preserve scholarly e-journals published globally.

## 7. Comparison of Six Current e-Journal Archiving Programmes

The following six programmes have been selected as directly relevant to UK libraries. Five programmes were included in the CLIR survey as meeting criteria for trustworthy digital repositories<sup>14</sup>. The sixth, the British Library programme, was not discussed in detail in the report but was noted in the list of promising emerging programs. They are all operated by organisations with excellent credentials in this sphere. At this early stage and given the varied factors which need to be taken into account to meet the needs of individual institutions, none will completely address all requirements for UK institutions.

It may be helpful to categorise the six programmes into those which are primarily concerned with long-term preservation of the scholarly record rather than perpetual access (British Library, CLOCKSS and e-Depot), those whose primary emphasis is on immediate or perpetual access (ECO and LOCKSS) and those which combine something of both (Portico).

### 7.1. LOCKSS

#### Overview

LOCKSS (Lots of Copies Keep Stuff Safe) enables participating libraries to collect, store, preserve and provide access to their own local copies of content to which they have subscribed. The LOCKSS system was one of two very different e-journal archiving approaches (the other being what is now known as Portico) supported by the Mellon Foundation following the conclusion of seven e-journal archiving projects in 2002. The LOCKSS application is open-source software. LOCKSS enables members to harvest the web-based presentation files of the content of e-journals to which they have subscribed from participating publishers. Access is triggered whenever (and for whatever reason) the material cannot be viewed on the publisher's (or intermediary's) servers. The highly distributed nature of the approach aims to ensure that there is sufficient replication to safeguard content despite any potential disasters which might befall individual LOCKSS institutions. LOCKSS introduced the LOCKSS Alliance as a membership organization in 2005, to introduce governance to the program and address sustainability issues. In 2006, the JISC funded a two-year initiative, including a UK support post, to raise awareness of LOCKSS. A total of 24 institutions from around the UK joined the initiative, with a further 6 funding their own participation in the trial.

#### Pros

- It allows libraries to collect and exert control over the material they licence, as they have done when purchasing print journals.
- LOCKSS has increased publisher participation and the number of titles at a great rate. For example, the *Metes and Bounds* report of 2006 reported 25 publishers on board, the LOCKSS website indicates 229 as of March 2008.
- It covers a significant number of smaller and therefore probably more vulnerable publishers.
- It requires relatively modest investment in staff and equipment.
- It permits immediate access to the archive whenever there is a problem with communication with a publisher's server, even if very short term.

#### Cons

- It will need ongoing technical support.
- Although much technical support is managed remotely, staff from the subscribing institution need to monitor the local server and ensure that it is running correctly.
- Future software development will require an active LOCKSS developer community.
- It *may* be difficult to integrate with other institutional technical platforms, e.g. several UK LOCKSS pilot reports referred to difficulties of proxy integration, though only one appeared to think this would be a major barrier. [The potential to use the LOCKSS

box as an openURL source would circumvent this, and the LOCKSS Alliance is working to achieve this.]

- Not all titles of a particular publisher, nor all issues of a particular title are necessarily included in the LOCKSS collection.
- The larger STM publishers such as Elsevier have not demonstrated an enthusiasm for joining, and have expressed some concern about security of licensed content when it is distributed so widely. The resulting lack of major publisher content was a concern raised by several UK LOCKSS participants in their reports.

## 7.2. CLOCKSS

### Overview

CLOCKSS “Controlled LOCKSS”, was launched in 2006 and is still at a relatively early stage of development. A not-for-profit collaboration between libraries and publishers, it is a dark archive based on the LOCKSS software in which a limited number of libraries take on an archival role on behalf of a broader community. CLOCKSS ingests and saves either source or presentation files, as the publisher chooses. Each library hosts two servers which create a network of dark repositories. There are 11 participating publishers and 7 libraries in the two year pilot programme, which was just concluding at the time this guide was written (April 2008). CLOCKSS intends to add a limited number of additional libraries, and more publishers in the next phase. It describes itself as complementary to LOCKSS, and a “strategic component in a multilayer, resilient, local and international preservation plan.” The 7 libraries participating in CLOCKSS are also members of the LOCKSS Alliance, 6 are U.S Based, and one is in the UK (University of Edinburgh).

### Pros

- It provides a community approach, with a small number of distributed libraries assuming responsibility for long-term archiving, working in partnership with publishers.
- Low cost. The CLOCKSS website claims that “Fees are low and for a limited period while we build the CLOCKSS Endowment which we expect to underwrite costs after five years.”
- It is likely to appeal to publishers with concerns about security of access and therefore attract a broader range of content from major publishers. While only 11 publishers were involved in the pilot, they were all major ones. The CLOCKSS website describes as a strength “...that the founding organizations, top publishers and university libraries, share a long history of survival and a deep understanding and experience with long-term sustainability.”<sup>15</sup>

### Cons

- Post-cancellation access is not supported.
- Long term costs are not clear, though there is an expressed commitment to reducing costs through the creation of an endowment.
- It is too early to be able to assess the long-term viability of the programme, including funding support.
- It is unclear what motivations (other than public good) CLOCKSS centres would have for continuing to participate in the programme long-term.

## 7.3. Portico

### Overview

Portico is the second approach to e-journal archiving supported by the Mellon Foundation. It was launched as an independent organisation in 2005, though it has been in planning and preparation since 2002 under the auspices of Ithaka and with support from JSTOR. Designed specifically as a third party service for scholarly literature published in electronic form, beginning with e-journals, it provides insurance to libraries that the e-journal content they have subscribed to will be preserved for the long-term. Portico provides access to the e-

journals they have preserved after specific 'trigger events'. Depending on what the participating publisher is able to provide, Portico preserves normalised source files and/or presentation files of e-journals and also performs rendition archiving. If no presentation files are provided, Portico creates new presentation files. Portico can become a delivery mechanism in the event of a trigger event. In addition, if a publisher has designated Portico as such, it can also serve as a potential mechanism for post cancellation access. As of 5 March 2008, 31 of the 50 publishers participating in Portico had nominated Portico as a possible mechanism to provide post cancellation access. This represents approximately 80% of the 7,596 titles committed to the archive (as of 12 May 2008). Appendix 2 illustrates which publishers participating in more than one archiving programme have nominated Portico to perform this role (as of March 2008).

#### Pros

- It removes the onus of managing the content from the library and provides assurance of long-term preservation.
- It can also provide post cancellation access providing the publisher has nominated them as a potential mechanism for this (to date more than half have).
- Publisher participation has grown at an impressive rate. The CLIR survey of 2006 indicated 13 publishers had joined Portico, which had grown to 50 as of early March 2008. Moreover some of the major STM publishers, such as Elsevier, have joined.
- Their archiving approach is very thorough, preserving source files and/or presentation files, depending on what the publisher chooses to supply.
- The recent agreement with the KB to act as a mirror site for Portico strengthens their credibility.
- It (arguably) provides an equitable business model, with publishers contributing to costs as well as libraries.

#### Cons

- Some see the dependence on publishers for revenue as a weakness.
- The fees libraries pay to Portico may be higher than other options.
- The current focus seems to be on the larger publishers
- Like CLOCKSS, it requires a cultural shift from libraries in terms of their traditional custodial role.
- Some believe that the title and publisher coverage could become rather US-centric once major international publishers have been covered.

### **7.4. e-Depot**

#### Overview

The Koninklijke Bibliotheek (KB) is the national library of the Netherlands and operates e-Depot, its archive for the Dutch national deposit collection of electronic publications and other e-content (eg Dutch newspapers). e-Depot, which is OAIS compliant, was established in 2003 and focused initially on Dutch material. Recognising the international nature of journal publishing, this has now been extended to international publications. The KB intends to conclude archiving agreements for all the journals from 20-25 of the world's largest publishers. Publishers wishing to make use of the services provided by e-Depot are required to conclude an archiving agreement with the KB and to deliver bulk content and specified metadata. The primary deposit file format is PDF. Generally, end-user access is restricted to on-site perusal for reasons of private research only and on-line access is denied.

KB, who is committed to a programme of research and development in this area, is promoting the concept of the "Safe Places Network". This recognises the power and utility of international cooperation in finding solutions to the challenges of long-term preservation. A recent example is the decision by Portico to lodge a copy of their archive with KB.

#### Pros

- e-Depot aims to cover all major STM publishers. It already has 12 major publishers on board.

- The KB has a strong reputation for leadership in digital preservation research and practice.
- e-Depot saves both rendition and source files.
- There are no costs involved for others; the service is currently underwritten by the Dutch government. KB sales literature though includes the statement *“In the coming years the KB intends to develop a sustainable business model for the e-Depot which will reflect both the public and private responsibility for our digital and cultural heritage.”*
- They have demonstrated a good understanding of technical issues.

#### Cons

- Because they are such major publishers, the trigger events which could provide open access to content archived by e-Depot are highly unlikely to occur.
- It is extremely doubtful that the KB will want to engage in managing numerous authentication systems in the case of cancellation of content, the trigger event most likely to cause disruption for most organisations licensing access.
- At the moment, assured access is only available onsite at the KB.

### **7.5. Electronic Collections Online (ECO)**

#### Overview

ECO was launched by OCLC in 1997 as a subscription service for libraries to a wide range of e-journals. It currently provides web access through OCLC’s FirstSearch service to over 5,000 titles from over 40 publishers. OCLC negotiates with publishers for perpetual access rights for subscribers to the service and for it to migrate backfiles to new formats if required.

#### Pros

- The emphasis is on continued access, which is likely to be of paramount concern to most libraries.
- It has been in operation longer than any of the others.
- It has a significant number of publishers and titles and good content coverage.

#### Cons

- Continued access is dependent on payment of an access fee. If titles are cancelled, access to past content can only be resumed if a subscription is reinstated within a five year period.
- ECO is primarily a tailored aggregator service for libraries : long term preservation is not their main mission.
- The emphasis on current access may be detrimental to long-term preservation (though some libraries find this feature makes the service easier to justify funding).

### **7.6. British Library e-journal Digital Archive**

#### Overview

Over the past few years the BL has been building a Digital Object Management System (DOM) capable of storing and managing all digital content the BL takes responsibility for, including their own digitally created content, material purchased and material acquired through voluntary and legal deposit. The DOM system, which will be OAIS<sup>16</sup> compliant, has three geographically dispersed identical nodes to provide redundancy in the event of loss of data from any one node.

The BL is currently assessing the feasibility of providing archiving and preservation solutions over and above those envisaged for e-legal deposit. The British Library began ingesting content from selected publishers during 2007, but is still finalising the exact service options that will be provided. At the time of writing, BL say that they will be testing initial voluntary deposited material in August 2008. Their intention is to launch an initial “Grey Archive solution” in the first quarter of 2009.

Access will depend on publishers' requirements. BL proposes two levels of service, Grey Archive (BL Reading Room access only), and Light Archive with content made available on the Internet worldwide following a publisher-specified trigger event.

#### Pros

- The British Library's strong reputation, experience and mandate for preservation generally;
- The archive is funded by charges to publishers and there are no other costs for libraries;
- The BL's extensive planning and preparation for digital deposit in building an infrastructure capable of dealing with both volume and complexity of material.

#### Cons

- It is too early to assess the content they will attract. As yet there is no detail of which publishers are on board. The emphasis is on recruiting publisher participation.
- The trigger events the BL use do not include post-cancellation access (apart from BL Reading Room access)<sup>17</sup>.

## **8. Practical experience of e-journal archiving solutions**

It is still premature to gauge which archiving programmes are likely to be most effective, or whether there will continue to be a need for a range of different solutions to suit a range of needs. All of the e-journal archiving programmes are relatively new. The British Library e-journal programme is still in development, CLOCKSS has only just reached the end of its 2-year pilot phase, Portico and LOCKSS were both launched in 2005, e-Depot in 2003. The oldest service is ECO, which began in 1997, but is in a different category to the others, having been established as a service to provide immediate access to titles. However a UK trial of the LOCKSS system has been underway in the UK since March 2006. A number of HEIs were funded by JISC to join LOCKSS. Pending completion of the detailed evaluation of the two-year pilot referred to in 8.3, some interim feedback is available from participants in the trial, as well as some data on the use of other archiving solutions.

### **8.1. Views on continuing with LOCKSS after 2-year subsidised trial**

Analysis of the 2007 reports from 25 LOCKSS participants revealed some interesting variations in perspectives of e-journal archiving. It is notable that there were only two institutions prepared to give an unequivocal ‘Yes’ to the question of whether they thought they would continue with LOCKSS post- the JISC trial, and four who thought they would probably not continue. The majority of participants were in various stages of indecision, though around half indicated they would like to continue, with costs and coverage being the two major factors likely to affect decision-making. Table 1 summarises responses to the question of whether or not they were likely to continue with LOCKSS.

It is worth noting that one of the purposes of the trial was to surface issues of concern to UK libraries. Since 2007, many of these concerns have been identified and have been or are being addressed, and details of the costs of the service to UK institutions after August 2008 have been circulated.

Response	Number	Rationale + Selection of relevant quotes
YES	2	<ul style="list-style-type: none"> <li>▪ “Yes, we would anticipate continuation of LOCKSS usage. A fee-based model would be acceptable, provided it was within our means and of good value.”</li> <li>▪ “Yes, we plan to continue &amp; would do so if fee is payable.”</li> </ul>
PROBABLY	9	<ul style="list-style-type: none"> <li>▪ “[LOCKSS] is the only model for preserving e-journals which allows control by individual libraries of their collections but it suffers from a lack of critical mass...”</li> <li>▪ Concerns about lack of key content, despite generally favourable response. [mentioned by 4 of the 9]</li> <li>▪ Concerns about fee [raised in 7 of the 9 responses], e.g. “We would very much hope that we continue with LOCKSS...However it would very much depend on the level of fees required.”</li> </ul>
POSSIBLY	10	<ul style="list-style-type: none"> <li>▪ Under review</li> <li>▪ Need more details on costs</li> <li>▪ Need more details on service support</li> <li>▪ Increasing competition, e.g. “Initiatives like Portico and CLOCKSS have grown rapidly and seem to have the commercial publishers pretty much sewn up. The Open Access titles are probably at a higher risk of disappearing completely, but we are not convinced that we should be paying to preserve them (would we pay for subscriptions to them?)...”</li> <li>▪ Unclear of business case, e.g. “I would need a much clearer understanding of LOCKSS and how it works (as mentioned in 6) in order to convince budget holders that we should pay to continue with LOCKSS.”</li> <li>▪ Concerns about disk space</li> <li>▪ Lack of key content [referred to by 5 of the 10, usually in conjunction with other factors]</li> </ul>
UNLIKELY	4	<ul style="list-style-type: none"> <li>▪ Fees, e.g. “Probably not, especially if fee-based”</li> <li>▪ Other options, combined with fee, e.g. “If the cost were anything other than minimal, we would prefer to spend money on a service such as Portico.”</li> <li>▪ Lack of content, combined with fees, e.g. “...it is unlikely that we would be willing to pay a fee unless the coverage of the system can include the more important journals. ... We would be looking to the national libraries (e.g. British Library) to take on e-journal preservation as part of their remit.”</li> <li>▪ Difficulties with setting up proxy server, combined with fees, e.g. “Given that end user access in our institution may be problematic without a proxy server, it is uncertain that the distributed model of preservation will work for us in the long term. A fee-based model ...would be a further barrier...”.</li> </ul>

**Table 1:** Summary of Responses to Q 10 – “The JISC supported LOCKSS Pilot Programme concludes at the end of February 2008. Do you anticipate a continuation of LOCKSS usage? To sustain the LOCKSS support service beyond this date, a fee-based model is required. How would this affect your decision in continuing to use LOCKSS?”

## 8.2. Views on e-journal archiving and costs

Nine of the 25 respondents said they were participating in another e-journal archiving programme and three indicated they were maintaining a watching brief. Portico was cited more than any other alternatives (five already participating and one indicated probable subscription), one was in CLOCKSS, and one each mentioned JSTOR, an E-Print Repository, and mounting a publisher archive locally. None mentioned the other four

archiving programmes under review, except obliquely, through comments relating to waiting for national libraries.

It would have been useful to explore these responses further. For example, it is unclear whether the 13 who answered 'No' were unaware of other options or uninterested in them. It would also have been interesting to explore further the rationale behind those who were participating in another solution and what they hoped to gain from it.

The responses from the 25 reports suggests that UK academic libraries are generally not yet wholly persuaded they need the safety net provided by e-journal archiving solutions [or at any rate are not confident they can build a convincing case for funding] and are waiting to see what develops. There appear to be very different views in terms of willingness to pay for archiving solutions but most expressed at least some concern, and for some it was an absolute barrier.

One library reported that they were preparing a digital preservation strategy for all their digital assets, including licensed e-journals, and that they would be evaluating LOCKSS against other solutions, such as Portico. Several others indicated a similarly cautious approach.

The reluctance to pay may well stem from the lack of serious consequences experienced at this early stage of e-journals, in that there have probably been few disruptions experienced because of lack of a trusted archiving system. Libraries will need to be convinced that the threat is sufficiently real over the medium to long term to justify an annual expenditure – even a relatively modest one, from stringent budgets.

It may appear to be valid risk management to defer any expenditure relating to insuring continued access to e-journals unless and until such time as there is an obvious negative impact from not doing so. There may also be a certain amount of willingness to wait and see what happens, especially in relation to what role national libraries might play, as evidenced in the comment on the British Library in Table 1. There is also an understandable reluctance to commit to one archiving solution until it is clearer which one is likely to best meet the needs of an individual institution.

A counter argument to this approach is that:

- Expenditure and dependence on e-journals is increasing at a rapid rate.
- A relatively small percentage of the overall budget could act as insurance against unforeseen circumstances and undue dependence on publishers to fulfil a role they have not traditionally taken.
- It should be possible to divert anticipated savings from binding, storing and managing print journals.

In the short term, this last benefit may be less evident as libraries will need to manage an interim period where they will see little advantage in terms of space savings. One respondent in the LOCKSS trial expressed some frustration with the lack of progress in this area, "In due course, as with many other academic libraries, we will need to deselect considerable quantities of print-based stock to free up shelf and study space. Originally, our vision was that LOCKSS would provide a suitable means of archiving based on permanent e-content for titles where publishers were unable to guarantee such access if subscriptions ceased. That momentum still seems to be lacking (the number of core titles is far too few) and we could only support LOCKSS if this particular benefit is to be realised reasonably soon".

To a large extent, the success of any e-journal archiving solution will depend on the willingness of libraries to factor the costs into their overall budgets, much as they did for storage of print journals and to view e-journal archiving as a legitimate and necessary investment. Whether or not national libraries accept archiving e-journals as part of their remit (as some, including the BL are), there is still likely to be a need for perpetual access, in addition to long-term preservation.

### **8.3. UK LOCKSS Second Survey Results**

At the same time as this study was being prepared, the 2-year UK LOCKSS Pilot Programme was being evaluated. Three general questions about preservation services were added to the questionnaire sent to all participants in the programme, and initial results from the survey have been shared with the authors of this report.

Taking each of the questions in turn:

**Have you investigated or signed up to any other e-journal preservation services e.g. Portico? Please provide details.**

Of the 22 responses, four stated they had signed up to Portico. Several are keeping a watching brief on Portico, though one expressed concern at Portico's costs. Two responses mentioned CLOCKSS (one of them was the University of Edinburgh which is hosting a CLOCKSS server). One response read *"Not yet, but we will be looking at, and evaluating, Portico over the coming months. However, I see this as a complementary service to LOCKSS, not as a replacement for it. Portico has more content and more of the bigger publishers, but it cannot give quick access to content in the way that LOCKSS can."*

Another made the following reference to BL: *"We have not signed up to other services but as a legal deposit library we participated in the pilot programme led by the British Library to test e-journal deposit in anticipation of the extension of legal deposit to this format and are currently participating in discussions with the BL and other legal deposit libraries concerning the scope of the voluntary scheme for e-journal deposit via the BL."*

The next question was:

**Which of the e-journal preservation services you use has given you the most confidence in their long term sustainability?**

There was no clear cut answer to this question. Only 3 mentioned LOCKSS without qualification, and four suggested Portico. Most answers were more nuanced, eg *"LOCKSS is rooted in the library and open-source communities. CLOCKSS and Portico have both demonstrated that they can make preserved content available and this may help to establish them."* Another said *"It is early to say – the buy-in from major publishers means that services like CLOCKSS and Portico have an advantage, but they do not offer Libraries assurance of access where there are temporary system failures or subscription problems. At this stage Libraries may choose to opt for multiple approaches, depending on costs and local resource issues."*

LOCKSS and Portico both had their supporters. *"Given the participants involved and consequent funding we would say Portico."* Another said: *"As indicated above, all solutions are in their infancy, and it is not possible to have confidence in the long term in any of them. LOCKSS is potentially sustainable given the investment already made in it by the community, and if strategic leadership is maintained."*

But the clearest indication of the current state of uncertainty came from the respondent who said: *"There is no e-journal preservation service where we are yet confident of long term sustainability."*

The third question read:

**Approximately how many of your e-journal titles are not, as far as you are aware, currently covered by any e-journal preservation services? What proportion is this of your total e-journal collection?**

Many respondents, the majority, didn't seem to have the exact figures to hand ("*I haven't been able to get a figure for this*"; "*Not sure – this has not, as yet, been investigated*"). The general view was that they think that only a relatively small fraction of their current subscriptions are covered by current preservation solutions. Figures of around 90 - 95% were suggested for titles not covered.

One respondent said: "*We have realised in the course of our internal review that we don't know this accurately, so it is another piece of work we need to do. We have approximately 20,000 e-journals including all the bundled deals, and approximately half of these come from the major NESLi deals where the publishers are at least involved in preservation initiatives.*"

Another respondent said:

*"We have around 60 titles being archived in our LOCKSS box. Our link resolver knowledge base has 35,000 journal titles. Around 17,000 of these are from the three biggest full-text aggregator databases to which we subscribe, and over 5000 are from directories of open access journals. Nonetheless, the proportion of journals covered by LOCKSS is currently tiny. At least another 2,500 subscribed titles will be covered by CLOCKSS (the ScienceDirect and Blackwell packages to which we subscribe). However, CLOCKSS coverage is of minimal interest to us as we are mainly concerned about post-cancellation access not long-term preservation. We cannot comment on the coverage from any other e-journal preservation service such as Portico, as we do not know the level of overlap with our journal holdings."*

#### **8.4. Manchester Workshop (Part 1)**

In the first part of the workshop held in Manchester on the 14<sup>th</sup> April, the six selected solutions were reviewed. Three separate discussion groups examined and commented on the approaches and the main points were then shared in a report-back session. The following is a summary based on notes taken during the event.

##### **General Issues**

During the discussions on the six preservation solutions, some general issues emerged. These included:

- Although there is a trend towards "e-only" options, restraining factors include some academics distrust of e-publications and resistance to change.
- Drivers for a move to e-only include access, space, user expectations, and 24-7 availability.
- NESLi2 deals had helped to build confidence in e-only.
- Libraries will need to lobby within their institutions to gain support for funding a contribution to archiving programmes.
- It would be helpful if a cost-benefit analysis could be carried out comparing the savings from hardcopy cancellations (space, binding etc) with e-only licensing plus archiving costs.
- Which solution, if any, is chosen will depend on several factors including critical mass of content, numbers of publishers, numbers of titles, issue coverage, and costs.
- It should be acknowledged that publishers themselves may be able and willing to provide ongoing post-cancellation access, though this may be a chargeable service.

## **LOCKSS**

This is seen as closest to the paper subscription model, and direct library involvement and instant availability in the event of temporary problems with accessing the publisher's server are viewed as positives. It also tends to be used by smaller publishers whose journals are arguably more in need of archiving solutions than the big, wealthy publishers who can look after their own, and costs are relatively low. On the other hand some publishers are wary of joining LOCKSS because of concerns about the security implications of multiple copies of their material being distributed to many LOCKSS boxes (the feature that is the major selling point of LOCKSS). The limited (though growing) number of publishers, and title and year coverage were currently seen as negatives. Some felt that LOCKSS needed a clearer business model and a formal service level agreement. Finally, there was some concern about who would be responsible for migrating content to new formats as and when these emerge, making earlier formats obsolete [Response from LOCKSS – LOCKSS will take responsibility].

## **CLOCKSS**

This is perceived to be a "preservation for the community" solution, rather than preservation for an individual institution. Some concern was expressed about the funding model (a review of CLOCKSS was scheduled for about the same time as this report was being prepared). Its effectiveness compared with Portico was questioned by some, but it was acknowledged that it was more attractive to some of the major publishers than LOCKSS. The fact that it is a "dark archive" means that it can only ever be part of the solution for universities, who also want, for example, perpetual access to cancelled titles.

## **Portico**

This service is perceived as expensive by some, though the transparency of the pricing was appreciated. [Note: UK pricing of LOCKSS and CLOCKSS services are in the process of being agreed at the time of publication of this report.] One contributor made the point that the costs are not that high when compared with their journal binding budget. Their thorough approach to long-term preservation was a positive, as was the growing list of publishers and titles, though there seem to be few small publishers. There was some concern about a possible US bias in the longer term, but the recent deal with KB was encouraging. A question was raised about authenticated access from the UK.

## **e-Depot**

The fact that e-Depot is currently funded by the Dutch government was seen as a positive, as was the fact that all the major STM publishers were participating. Their pragmatic approach and apparent good understanding of the technical issues were appreciated. But there was a question about lack of participation by smaller publishers, and concerns about restrictions on access. [In response, KB say they are starting a project – the Directory of Open Access Journals – DOAJ – which includes 3000 small publishers. Open access journals and publishers, such as Biomed, are of course freely accessible.]

## **OCLC ECO**

This was seen as a different sort of service to the others, and not primarily a preservation solution. Nevertheless it had good content coverage and perpetual access to cancelled titles (on payment of a subscription) was a useful function.

## **BL e-Journal Digital Archive**

At the time of the workshop, little was known about the latest developments at BL (see elsewhere in this report for an update on BL's plans). So much of the discussions focused on the lack of clarity about business models, access options, content coverage etc. It was assumed that their primary focus might be on UK material, which raised the issue about their effectiveness as a backup for international publications. On the other hand they were perceived to be a trusted organisation who should be able to attract a wide range of publishers.

## 9. Four Scenarios

### 9.1. Introduction and purpose

There are a variety of reasons why an electronic journal might become unavailable to the members of a subscribing institution. These may be predictable (library budget cuts mean a title is cancelled), or unplanned (a title is withdrawn by a publisher). They may be short term (temporary failure of the network or hardware) or much longer term, or even permanent (catastrophic loss of systems and hardware supporting a remote service as a result of fire, flood, explosion or terrorist action). Some are more likely than others, and the consequences will vary from minor inconvenience to severe disruption, depending on a range of factors.

In order to illustrate the potential problems and the ability of the different solutions to cope with them, the following four trigger events have been identified:

- Cancellation of an e-journal title by a library.
- E-Journal is no longer available from a publisher (title discontinued or sold to another publisher).
- Publisher has ceased operation and access to their e-journal servers is no longer possible.
- Catastrophic failure of publisher's operations/servers.

Each of the trigger events is compared across six archiving programmes in **Figure 2**, and is described and discussed in greater detail below.

### 9.2. Scenario 1

#### **Library cancels subscription and requires access to past subscribed issues**

This is the only one of the four scenarios where the event is triggered by the library. Before taking this step, the library will need to have developed its own risk assessment and a contingency plan for perpetual access to past subscriptions. This will include negotiating, where possible, with publishers and agents for post-cancellation access, agreeing any additional fees and the mechanism for access (eg via publisher's servers, or copies for loading on local systems). If copies are provided for local hosting, the nature of those copies needs to be agreed. For example it is unlikely that most libraries would want the basic source files (text, illustrations, tables etc) that were used to create an issue of a journal. They are more likely to want the presentation files.

Of course the reasons for cancellation, combined with the needs of the primary clientele, may influence what (if any) alternative arrangements are necessary. For example:

<b>The title is being cancelled because it is no longer considered core to the primary clientele.</b>	
<i>Requirement</i>	<i>Action</i>
Continued access to back issues is still needed for researchers who will need timely and convenient access to data contained in them for the foreseeable future.	Prior investment by the library in an archiving solution which provides post-cancellation access to the title. Both risk and impact are high
Demand for past issues is likely to be low but for a key part of the library's clientele, they will be essential to their research and they will need timely access.	Investment in a trusted archiving solution is highly advisable. Risk is low but impact would be high
Demand for past titles is likely to be fairly low and intermittent and/or demand can probably be satisfactorily met from other sources, e.g. interlibrary loan or onsite access at the BL.	Investment in a trusted archiving solution is optional, especially if the title is readily available elsewhere and speed of access is not a major issue. Both risk and impact are fairly low
<b>The title is being cancelled because of severe budget pressure.</b>	
<i>Requirement</i>	<i>Action</i>
The Library needs assurance that their previous expenditure has not been wasted.	Investment in a suitable archiving solution would enable continued access to past issues (if the publisher does not support access)

As the following table shows, only some of the solutions reviewed are set up to provide access where the trigger event is cancellation of a subscription.

<b>Scenario 1: Library cancels subscription and needs access to past issues to which they subscribed</b>		
<b>Solution</b>	<b>Triggered?</b>	<b>Comment</b>
LOCKSS	Yes Immediate	LOCKSS assures members of access to participating publishers' content provided they have purchased or licensed that content.
CLOCKSS	No	CLOCKSS only makes titles openly accessible following a defined trigger event. Subscription cancellation is not such a trigger.
PORTICO	Yes Delayed	Providing the publisher has nominated Portico as a provider of post-cancellation access.
e-Depot	No	Except for onsite access.
OCLC ECO	Yes Immediate	ECO provides continued access on payment of an access fee.
BL E-J archive	No	This is not one of the trigger events included.

### **Support for Scenario 1**

### 9.3. Scenario 2

#### E-journal or its past issues no longer available from the publisher

This is a highly likely scenario as publishers merge or change their business models, or as larger publishers review and adjust their portfolio of titles. Journal titles are also sometimes traded between publishers, which may mean that access to past issues is no longer supported by the previous owner.

The UKSG TRANSFER initiative<sup>18</sup> has produced a Code of Practice aimed at easing the problems created when journal titles move between publishers. Of relevance are the following paragraphs contained in the latest version of the draft (Version 1.0 – April 2008):

It is very common for journal content to be included in one or more archiving services. The Receiving Publisher will not remove content that was previously deposited in an archive, or archives, even if the Receiving Publisher will not be continuing to deposit content in the archive, or archives. The Receiving Publisher is encouraged to continue the existing archiving arrangements for a journal after the Effective Transfer Date.

The Receiving Publisher will honour any perpetual access rights to previously published content which have been granted by the Transferring Publisher with the authority of the journal owner.

The real-life scenario of the publisher Sage no longer offering its publication *Graft* has provided an opportunity to demonstrate the success of two archiving solutions, Portico and CLOCKSS. In this case each is able to continue to offer access to the issues they hold, either as open access (CLOCKSS) or else as a service to members (Portico). While it cannot be guaranteed that the archive will include all back-issues of the title (as with *Graft*), participation in an archiving solution which covers at least some issues will significantly reduce the risk of disruptions to continuity of service.

Thus, summarising the risks and impacts of this particular scenario:

<u>Risk of event happening:</u>	High
<u>Impact on subscribers</u>	
Title is core to primary clientele	High Impact
Retrospective access is needed	High Impact
Requirement is mainly for most recent issues	Low Impact

Scenario 2: E-Journal or its past issues are no longer available from the publisher		
Solution	Triggered?	Comment
LOCKSS	Yes Immediate	LOCKSS assures members of access to participating publishers' content provided they have purchased or licensed that content.
CLOCKSS	Yes Delayed	The title would be made openly accessible to all.
PORTICO	Yes Delayed	The title would be opened up to all active participants, regardless of whether they previously subscribed to the content.
e-Depot	Yes Delayed	All archived titles will always be available onsite at the KB. This trigger would result in the title being made openly accessible, subject to publisher agreement.
OCLC ECO	Yes Immediate	ECO's Business Model is to continue to provide access to journal titles on payment of an access fee.
BL E-J archive	Yes Delayed	All archived titles will always be available onsite at the BL. Subject to publisher agreement, this trigger would result in the title being made openly accessible.

#### Support for Scenario 2

#### 9.4. Scenario 3

##### **Publisher has ceased operation and e-publication is no longer possible**

In this scenario, the publisher is no longer in business and therefore unable to support a service providing access to their collection of previously published journal issues. Suitable strategies for coping with this event depend on risk management techniques. The probability of this trigger event for large publishers is arguably low. However, the impact for research institutions may be unacceptably high.

The risk and impact relationship is more complex than for some of the other scenarios, as illustrated in **Figure 1**, which includes this risk in all four risk/impact quadrants.

<b>Scenario 3: Publisher has ceased operation and e-publication is no longer possible</b>		
<b>Solution</b>	<b>Triggered?</b>	<b>Comment</b>
LOCKSS	Yes Immediate	LOCKSS assures members of access to participating publishers' content provided they have purchased or licensed that content.
CLOCKSS	Yes Delayed	The title would be made openly accessible to all.
PORTICO	Yes Delayed	The title would be opened up to all active participants, regardless of whether they previously subscribed to the content.
e-Depot	Yes Delayed	All archived titles will always be available onsite at the KB. This trigger would result in the title being made openly accessible, if agreed with publisher.
OCLC ECO	Yes Immediate	ECO's Business Model is to continue to provide access to journal titles on payment of an access fee.
BL E-J archive	Yes Delayed	All archived titles will always be available onsite at the BL. Subject to publisher agreement, this trigger would result in the title being made openly accessible.

#### **Support for Scenario 3**

## 9.5. Scenario 4

### Catastrophic failure of publisher's operations/servers

In this scenario, access to a publisher's e-journals suffers a major disruption. This is a somewhat unlikely, but not entirely implausible event. Examples can be thought of that might include fire, flood, explosion, lightning strike or terrorist action, which result in the total destruction of a publisher's servers and, perhaps, damaged or destroyed backup materials. It should be added that major publishers are likely to take steps to protect themselves against such events by running mirror sites at various locations around the world, but smaller publishers might not have the resources to do this. Off-site or secure fire-safe storage of backup media is also common practice, but perhaps not universal.

Rebuilding a service following such an event might take many months. An agreement between a publisher and an archiving solution that this is an acceptable trigger to open up access to the archive would enable subscribers to continue to access back copies. For example Portico specifies the following condition for opening access: *"Licensor has stopped publishing or providing access to the Publication for a period longer than ninety (90) days due to technical difficulties or any business interruption, bankruptcy, insolvency, receivership or business failure."* Note, however, that if it is assumed that the publisher will be able to resume operations at a future date, such a service would only open up access to the material on the preservation service's servers until such time as the publisher was able to resume their own service. The normal legal protections against downloading bulk data for access on other servers would still have to apply.

This scenario depends on risk management techniques.

#### Risk of event happening

Major Publishers	Low
Medium publishers	Low
Small Publishers	Medium

#### Impact on subscribers

Core Title	High Impact
Current access required	High Impact
Non core title	Low Impact

Scenario 4: Catastrophic failure of publisher's operations/servers		
Solution	Triggered?	Comment
LOCKSS	Yes Immediate	LOCKSS assures members of access to participating publishers' content provided they have purchased or licensed that content.
CLOCKSS	Yes Delayed	Content moved to a hosting platform and made freely available
PORTICO	Yes Delayed	Portico subscribers get free access to content as long as publisher is unable to provide a service
e-Depot	Possibly Delayed	All archived titles will always be available onsite at the KB. Network access depends on agreements with publishers.
OCLC ECO	Yes Immediate	ECO provides continued access on payment of an access fee.
BL E-J archive	Possibly Delayed	All archived titles will always be available onsite at the BL. Network access depends on agreements with publishers.

### Support for Scenario 4

**Figure 2: Comparative Analysis of E-Journal Archiving Programs – Trigger Events**

Trigger Event	Access Arrangements					
	LOCKSS	CLOCKSS	PORTICO	e-Depot	OCLC ECO	BL E-jnl archive
1. Library cancels subscription and needs access to past issues to which they subscribed	<b>Yes</b>	<b>No</b> This is not one of the trigger events included	<b>Yes</b> Providing the publisher has nominated Portico as a provider of post-cancellation access.	<b>No</b> Except for onsite access.	<b>Yes</b> ECO provides continued access on payment of an access fee.	<b>No</b> This is not one of the trigger events included.
2. E-Journal or its past issues are no longer available from the publisher	<b>Yes</b>	<b>Yes</b> The title would be made openly accessible to all.	<b>Yes</b> The title would be opened up to all active participants, regardless of whether they previously subscribed to the content	<b>Yes</b> At least onsite access. Open access following trigger if agreed with publisher.	<b>Yes</b> ECO's Business Model is to continue to provide access to journal titles on payment of an access fee.	<b>Yes</b> At least onsite access. Open access following trigger if agreed with publisher.
3. Publisher has ceased operation and e-publication is no longer possible.	<b>Yes</b>	<b>Yes</b> Ditto above	<b>Yes</b> Ditto above	<b>Yes</b> Ditto above	<b>Yes</b> Ditto above	<b>Yes</b> Ditto above
4. Catastrophic failure of publisher's operations/ servers	<b>Yes</b>	<b>Yes</b> As long as publisher is unable to provide a service.	<b>Yes</b> As long as publisher unable to provide a service.	<b>Possibly</b> Depends on agreements with publisher.	<b>Yes</b> Ditto above	<b>Possibly</b> Depends on agreements with publisher.

## **9.6. Manchester Workshop (Part 2)**

In the second half of the workshop, the four scenarios described above were examined in turn. A general issue that was raised concerned access mechanisms and link resolver services needing to be able to recognise new routes to preserved titles when a trigger event occurs.

### **Scenario 1**

#### **Library cancels subscription and requires access to past subscribed issues**

The point was made that the first option would normally be to negotiate post-cancellation access with the publisher (which may entail access fees). Ideally this should be embedded in the original contract. Cancellation might mean that the title is considered of low value in which case post cancellation access would not be important. Titles from smaller publishers which are typically not covered by big deals, such as NESLi2, were seen to be most at risk due to lack of resources and expertise, and where preservation solutions could be most valuable.

### **Scenario 2**

#### **E-journal or its past issues no longer available from the publisher**

The Transfer project report helps to clarify responsibilities when a title moves between publishers. Packaging of bundles of titles was highlighted as a problem. Sometimes libraries are faced with a choice where none of the bundles includes all the titles they want (eg they may want the new titles included in a new bundle, but not lose access to previously subscribed titles). Can an archive help in this situation? A concern was raised that technology developments, such as, for example, PDF or other file formats becoming obsolete, might result in some publishers choosing not to migrate some low use titles or year-ranges.

### **Scenario 3**

#### **Publisher has ceased operation and e-publication is no longer possible**

Small publishers were recognised as being at greatest risk. There is a possible important role for the British Library here in preserving access to material from this high risk group.

### **Scenario 4**

#### **Catastrophic failure of publisher's operations/servers**

Some doubted that this is a very likely event. It was pointed out that a major failure of aggregator services could also have an impact. Concern was expressed at the ability of some backup services to cope with very high demand if a popular service failed. Big publishers are more likely to have effective strategies to cope with major failures (eg by distributed multiple servers) than small ones.

## 10. Criteria for judging relevance and value of new archiving initiatives

At one extreme, there might be a small institution with limited resources which subscribes to a limited number of journal subscriptions, most, if not all, of which are either also held in print and/or held by a major archiving institution, such as the British Library, which is reasonably close by. In these circumstances, the institution may well decide not to subscribe to any archiving solution but to take the risk that, should they lose access because of a specific trigger event, their clients will be able to gain access to the material by either visiting the British Library or else through inter library loan.

At the other extreme, a large, well-funded institution has a world class research program which relies heavily on timely access to a number of e-journals, many of which are not available in print form. They are also concerned that the scholarly record is preserved long-term. In this case, it may be considered expedient to invest in more than one solution – either because this is the only way that complete coverage of the titles of interest can be achieved, or else because it is considered wise to invest in judicious overlap at such an early stage of e-journal archiving.

In between these two extremes there are of course numerous permutations, and it is unlikely that any scenario will be as clear-cut as these two. For the most part, it is likely that the designated community of any institution will be best served by having some form of guaranteed continued access to journal content to which they have subscribed.

Factors which might impact on decision making for archiving solutions could include:

- Does the service include titles which are core to my institution's primary clientele/designated community?  
*As indicated in Appendix 1 and 2, there is relatively small overlap between the archiving solutions to date. It will also be important that significant date ranges are covered, as well as titles*
- Does the overall bundle of titles in the archive meet my needs?
- Is my Library planning [or has already commenced] a move to e-only access?
- Does my Library intend to deselect print journals as electronic backfiles become available?  
*The space savings will constitute a cost saving but will need to be offset against the need to be assured of continued access. As one respondent to the JISC LOCKSS pilot noted "Part of the reason we are purchasing backfiles like this is to enable us to discard print journal volumes, as we are chronically short of space. However, the future stability of publishers is not guaranteed, and we might find that the publisher who sold us a backfile package has gone out of business and that the content we purchased in perpetuity is no longer available to us. Furthermore, we would have discarded any print volumes we had, leaving us with absolutely nothing."*
- Do I need assurance of continued *online* access?  
*Or will it be sufficient to provide offline access, via another archiving institution which permits onsite access?*
- Do I need assurance of continued access for an indefinite period/longer than 10 years?

*Or is the usage pattern of the designated community primarily interested in only the most recent literature?*

- Can I afford the service and does it represent value for money?  
*Cost will inevitably be a factor in decision-making. However, it also needs to be balanced against overall investment in e-journals and also the prospects of cost savings in terms of storing, binding, and managing large print collections.*

If the answer is yes to most of the above, then it will certainly be worthwhile investing in a trusted archiving solution, such as LOCKSS, Portico, and/or CLOCKSS.

Assuming the title is covered in more than one archive, what factors might affect which solution to go for?

- Do I want control over the titles that I have subscribed to in a similar way to print ? [e.g. by retaining copies of the journals my library subscribes to, such as LOCKSS]
- Do I need seamless access following a trigger event?  
*Or would it be acceptable to wait for a period of weeks or possibly months?*
- Do I want to delegate responsibility for archiving and perpetual access to a trusted third party? [e.g. Portico]
- Am I more concerned about the vulnerability of titles from smaller publishing houses rather than the major players?

If there is a combination of the content being of key importance and concern about the early stage of archiving solutions, it could be advisable to subscribe to more than one, assuming the content of primary interest is offered.

## 11. Observations and Conclusions

Although there are many obvious benefits that accrue from publishing and accessing academic papers through the internet, there are costs and challenges which need to be addressed. These are the shared responsibility of all in the information chain, including authors, publishers, repository managers, librarians, subscription agents and aggregators. It is also arguable that society as a whole has an interest, and therefore governments should have a role to play.

High amongst the list of challenges is ensuring that this material can continue to be read and appreciated for the indefinite future. Traditionally this has been dealt with by librarians and archivists who received, bound and stored paper journals. But electronic journals present new concerns and new responsibilities. A particular challenge is that, while libraries and librarians might be assumed to be the natural custodians of these materials, and to have the greatest interest in their long term preservation, the owners of the files are, in the majority of cases, the publishers. Subscriptions usually buy internet access and not, normally, the receipt of anything tangible (the LOCKSS model being an exception).

Furthermore, preservation inevitably incurs costs. These costs should, arguably, be shared by all who benefit from this new mode of delivery, including both consumers and suppliers.

This study has examined six different approaches to archiving and preservation. All deserve support and encouragement, but none currently offers the typical academic library a complete solution to their archival needs. Nor do any of them currently cover the greater proportion of the journal titles being published today.

Three of the six (LOCKSS, Portico, ECO) require the library to pay a subscription (the recent LOCKSS UK trial was subsidised by the JISC). Portico and CLOCKSS also receive income from publishers. e-Depot and the British Library currently get their income from government grants.

Taking each of the key players in turn, this study makes the following observations:

### Libraries

- Libraries should take the initiative in raising awareness and understanding within their institutions of the importance of perpetual access to, and preservation of, e-journals.
- Libraries should work with policy makers to ensure that perpetual access and preservation gets incorporated into both library and institutional strategies and policies for research, teaching and learning.
- Using the scenarios outlined in this report, libraries should carry out a risk assessment on the impact of loss of access to e-journals by their institution, and a cost/benefit analysis, in order to judge the value and relevance of the archiving solutions on offer.
- While preservation costs must be factored into library budgeting, it is acknowledged that libraries need to make value judgements on how much investment in archiving and preservation solutions is appropriate for their circumstances.
- The more that preservation costs can be shared throughout the community, the lower the charges for any individual library.
- The British Library is in a unique position in the UK. It can, and should, provide a safety net, at least for all electronic journals that originate in the UK, if not beyond. It is encouraging that it is collaborating with others through the Planets and other digital preservation R&D projects. International cooperation is another important dimension to e-journal archiving. However BL needs to be clearer on what it intends to offer end users in the academic community in the UK and elsewhere. For now, libraries and

publishers would be advised to continue to support one or more of the other solutions which appear better defined and more mature.

### **Publishers**

- In accepting their role in managing the delivery of electronic journal content directly to the reader, publishers must also acknowledge that they have a special responsibility to ensure that this material is secure for the long term future.
- They must offer subscribing libraries clear information on their archiving and post-cancellation access policies. Archiving and perpetual access must become essential parts of the subscription package offered to customers. In particular they must be able to tell libraries how they would handle each of the scenarios described in section 9.
- Although large publishers may be able to create and manage their own archiving processes and strategies, smaller organisations are likely to need to rely on one or more archiving solutions, such as those detailed in this report. Regardless, all publishers should sign up to one or more approved e-journal archiving initiatives.

### **Negotiators**

- By specifying conditions in their contracts, negotiators, such as Content Complete, have the potential to have a significant influence on arrangements for post-cancellation access and the rate of take-up of archiving solutions.
- When a community wide deal is agreed, an essential part of the package must be a clear commitment by the publisher to support one of more approved archiving solutions, whether or not they have their own preservation processes. Again, their position on each of the scenarios in section 9 must be clear.

### **Archiving Solutions**

- For an archiving solution to survive in the long term, it needs to be sustainable, which implies a sound and transparent financial model, visibility to the widest possible community (which implies self-promotion), and buy-in from as many publishers as possible (ideally representing the great majority of the titles a typical library subscribes to).
- Archiving solutions need to earn the trust of publishers. The source files and rendition files of a publisher's journals represent their core business assets. They will be understandably wary of releasing these to third parties without clear and verifiable guarantees and commitments that they will be safeguarded from unauthorised disclosure. Clear and unambiguous terms and conditions need to be agreed for triggering the opening up of access.
- A successful archiving solution also needs technical insight, planning, and expertise, and effective communication with publishers in order to deal with issues such as appropriate data structures and formats for storage and retrieval.
- Any archiving service providers also need to recognise that publishers are likely to want to add value to their e-publications by introducing features that cannot be replicated in the paper version, such as forward citations, multimedia clips and so on. Retaining these features in any archiving service will be challenging.
- Archiving solutions need to be clear about exactly what they are able and willing to provide in the way of (potentially high volume) end-user access services following a trigger event, and whether this is open access to all, access only to subscribers to the archiving service, or access only to previous subscribers to the title. If authentication is required, how this is administered must be specified.

## 12. Recommendations

From the study and its observations and conclusions, we can draw the following major recommendations.

1. When negotiating **NESLi2 agreements**, **JISC's negotiators** should take the initiative by specifying archiving requirements, including a short-list of approved archiving solutions.
2. To help quantify the insurance risk and the necessary appropriate investment, bodies representing **publishers** and other trade organisations should gather and share statistical information on the likelihood of the trigger events outlined in this report.
3. **Post cancellation access** conditions should be defined in the licensing agreement between libraries and publishers. Publishers should be strongly encouraged to cooperate with one or more external e-journal archiving solutions as well as provide their own post-cancellation service (at minimal cost).
4. The publisher (or subscription agent) should state their policy on **perpetual access** under the four scenarios described in section 9.
5. When **titles are sold** on to other publishers, the Transfer Code of Practice (see section 9.3.) should be followed.
6. **Archiving service providers and publishers** should work together to develop standard cross-industry definitions of trigger events and protocols on the conditions for release of preserved content. Project Transfer is a potential exemplar. The ground rules for any post-trigger event negotiation should be clear and transparent and established in advance.
7. Archiving solutions must provide greater clarity on **coverage** details, including not only publishers and titles, but also the years and issues included in the archive.
8. Using the scenarios outlined in this report, libraries should carry out a **risk assessment** on the impact of loss of access to e-journals by their institution, and a cost/benefit analysis, in order to judge the value and relevance of the archiving services on offer.
9. Relevant UK bodies and institutions should use whatever influence they can bring to bear to ensure that archiving solutions cover publishers and titles of particular value to **UK libraries**.
10. The findings of this study should be **reviewed and updated** at regular intervals to reflect continuing developments in the field of e-journal archiving and preservation.

## Appendix 1

TABLE OF PUBLISHER PARTICIPATION as of 5/03/08

LOCKSS	CLOCKSS	PORTICO	E-DEPOT	ECO
229	11	50	12	39 [Note: webpage indicates "over 70" but the list of publishers totals 39 as of 12/03/08]

*Note:*

Publisher participation in the BL's E-Journal Digital Archive is not currently known.

Also note that these figures are from the websites of each e-journal archiving solution provider and are not compiled on an identical basis e.g. the publisher BioOne can be considered as one or 109 separate publishers depending on the approach adopted.

Only **three** publishers, albeit fairly major ones (OUP, Sage, Taylor & Francis) are in all five archiving programmes. A total of **thirty** publishers are participating in more than one archiving programme, as indicated in **Appendix 2** (March 2008). A total of **282** discrete publishers are represented by one or more archiving programmes.

However, these figures provide only part of the picture and can be misleading when viewed in isolation. Those archiving programmes with more publishers do not necessarily hold more content, though the latter is not easily determined. As the authors of the CLIR survey discovered, establishing actual content coverage is very complicated. The following factors were cited by them:

- Presence of a publisher does not necessarily indicate all titles published by them are included in the programme.
- Even when the titles are listed, date spans are not necessarily included [of the five programmes under discussion, only ECO provided this information as of 1 July 2006]
- The pace of consolidation within scholarly publishing houses in particular, makes accurate and up-to-date information on titles coverage problematic.

Because of these and other factors, the authors of *Metes and Bounds* concluded that 'Thus, the publisher listings presented here should be viewed as no more than a fuzzy snapshot of circumstances on July 1, 2006.' Despite this, their tables reveal an interesting variation between publishers and titles archived. For example, their report indicated that, of the five archiving programmes under discussion, ECO had both the largest number of publishers and titles at that time (c. 5,500 titles from c. 40 publishers). Portico, on the other hand, though only listing around half the number of publishers at that time as LOCKSS (13 compared to 25), had more than twice as many titles included (c. 3,500 titles compared to c. 1,500 titles).<sup>19</sup>

Portico have recently introduced a new Holdings Comparison service which enables libraries to check whether their titles are included in Portico's archive, at no cost to participating or non participating libraries. By providing Portico with a list of ISSN's associated with holdings, Portico will then check whether they hold the title. Although the standard report does not include date ranges, these can be added on request.

## Appendix 2

### Publishers in more than one programme (as of March 2008)

Publisher	LOCKSS	CLOCKSS	PORTICO	E-DEPOT	ECO
American Anthropological Association	√		√*		
<b>American Chemical Society</b>		√	√		
American Meteorological Society			√		√
<b>Annual Reviews</b>	√		√		
Berkeley Electronic Press	√		√*		
BioMed Central	√			√	
Brill				√	√
Brookings Institution Press	√				√
<b>Cambridge University Press</b>			√		√
Duke University Press	√		√*		
<b>Elsevier</b>		√	√*	√	
Hindawi Publishing Corporation	√		√*		
<b>Institute of Physics</b>	√	√	√		
IOS Press				√	√
Johns Hopkins University Press	√		√		
Mary Ann Liebert			√*		√
MIT Press	√		√*		√
<b>Nature</b>		√	√*		√
Ohio State Uni Press	√				√
<b>Oxford University Press</b>	√	√	√*	√	√
Palgrave Macmillan			√*		√
<b>Project Muse</b>	√				√
Royal Society of Medicine Press	√		√*		
<b>SAGE Publications</b>	√	√	√*	√	√
Seismological Society of America	√		√		
<b>Springer</b>		√	√	√	√
<b>Taylor &amp; Francis</b>	√	√	√*	√	√
University of California Press			√*		√
University of Chicago Press	√		√*		
<b>Wiley-Blackwell</b>	√	√	√*	√	√

### KEY

**NESLI2 Publishers** [Note: There are seven other 2007-2009 NESLI2 publishers who are not participating in more than one archiving programme. Of these, American Institute of Physics is a member of Portico, BMJ Publishing Group and New England Journal of Medicine are in LOCKSS, and British Psychological Society + Royal Society of Chemistry are in ECO, leaving only two – AAAS Science Online and Informa Healthcare who were not participating in an archiving programme as of March 08.]

\* Publishers who have nominated Portico as one possible access mechanisms for post-cancellation access [Note: A total of 31 out of 50 Portico publishers had made this commitment as of 5/3/08]

## References

---

- <sup>1</sup> Archiving E-Journals Consultancy. October 2003.  
[http://www.jisc.ac.uk/uploaded\\_documents/ejournalsfinal.pdf](http://www.jisc.ac.uk/uploaded_documents/ejournalsfinal.pdf)
- <sup>2</sup> Kenney, A et al [2006]. E-Journal Archiving Metes and Bounds: A Survey of the Landscape.  
<http://www.clir.org/pubs/reports/pub138/contents.html>
- <sup>3</sup> Jones, M [2007]. Review and Analysis of the CLIR report E-Journal Archiving Metes and Bounds: A Survey of the Landscape  
[http://www.jisc.ac.uk/whatwedo/programmes/programme\\_preservation/ejournalarchiving.aspx](http://www.jisc.ac.uk/whatwedo/programmes/programme_preservation/ejournalarchiving.aspx)
- <sup>4</sup> Muir, A [2008]. E-Journal Registry Scoping Study.  
[http://www.jisc.ac.uk/whatwedo/programmes/programme\\_preservation/2008ejournalregstudy.aspx](http://www.jisc.ac.uk/whatwedo/programmes/programme_preservation/2008ejournalregstudy.aspx)
- <sup>5</sup> Kenney, Anne R; Entlich, Richard; Hirtle, Peter B; McGovern, Nancy Y. and Buckley, Ellie L (2006). *E-Journal Archiving Metes and Bounds: A Survey of the Landscape*. p.5  
<http://www.clir.org/PUBS/abstract/pub138abst.html>
- <sup>6</sup> Waters, Donald (2002). Good Archives Make Good Scholars: Reflections on Recent Steps toward the Archiving of Digital Information. In *The State of Digital Preservation: an International Perspective*.  
<http://www.clir.org/pubs/reports/pub107/waters.html>
- <sup>7</sup> Kenney, Anne R et al(2006). Op.Cit.p.37
- <sup>8</sup> Waters (2002). Op Cit.
- <sup>9</sup> Jones M (2007). *JISC Briefing Paper: e-journals Archiving and Preservation*.  
[http://www.jisc.ac.uk/publications/publications/pub\\_ejournalspreservationbp.aspx](http://www.jisc.ac.uk/publications/publications/pub_ejournalspreservationbp.aspx)
- <sup>10</sup> Kenney, Anne R et al(2006). Op.Cit.p.37
- <sup>11</sup> Waters, Donald (2005). 'Urgent Action Needed to Preserve Scholarly Electronic Journals'  
<http://www.diglib.org/pubs/waters051015.htm>
- <sup>12</sup> JISC SHERPA DP Project  
[http://www.jisc.ac.uk/whatwedo/programmes/programme\\_preservation/programme\\_404/project\\_sherpa2.aspx](http://www.jisc.ac.uk/whatwedo/programmes/programme_preservation/programme_404/project_sherpa2.aspx)
- <sup>13</sup> PRESERV 2 JISC Project  
[http://www.jisc.ac.uk/whatwedo/programmes/programme\\_preservation/programme\\_404/project\\_sherpa2.aspx](http://www.jisc.ac.uk/whatwedo/programmes/programme_preservation/programme_404/project_sherpa2.aspx)
- <sup>14</sup> Kenney, Anne et al (2006). Op Cit.
- <sup>15</sup> CLOCKSS Website. <http://www.clockss.org/clockss/Home>
- <sup>16</sup> Open Archival Information System; see <http://www.ukoln.ac.uk/repositories/digirep/index/OAIS>
- <sup>17</sup> Sourced from The British Library e-Journal Digital Archive on 2/04/08  
<http://www.bl.uk/reshelp/atyourdesk/docsupply/publisher/blejournal/index.html>
- <sup>18</sup> The UKSG Transfer Project: <http://www.uksg.org/transfer>
- <sup>19</sup> Kenney, A et al (2006). Op Cit. p.30-31