Establishing a UK LOCKSS Pilot Programme

This article describes the UK LOCKSS Pilot Programme, an initiative funded by the Joint Information Systems Committee (JISC) in partnership with the Consortium of Research Libraries in the British Isles (CURL). The pilot directly engages a number of selected UK HE institutions in trialling the LOCKSS technology for e-journal archiving and preservation. LOCKSS (for Lots Of Copies Keep Stuff Safe) is an initiative that has been developed by Stanford University Libraries. It involves an alliance of over 80 libraries and 60 publishers worldwide, working together to preserve persistent access to licensed e-journal content from publishers to libraries. The pilot is funded under JISC's development programme area 'digital preservation and records management'. This article describes the rationale to establish the pilot programme, the aims and objectives and, in detail, the individual components of the programme.

HELEN HOCKX-YU
Programme Manager, JISC

Introduction

Ensuring long-term preservation of, and continuing access to, scholarly and educational resources is an important strategic area for the Joint Information Systems Committee (JISC).[1] On 1 November 2005 JISC issued a call in partnership with CURL (Consortium of Research Libraries in the British Isles) to invite libraries of UK Higher Education (HE) institutions to participate in a UK LOCKSS Pilot Programme for e-journal archiving and preservation. The pilot commenced in February 2006 and will run for twenty-four months.

Lots of Copies Keep Stuff Safe (LOCKSS)[2] is a tool developed at the Stanford University Libraries to preserve continued access to web-published e-journals. The LOCKSS pilot will be funded under JISC's development programme area 'digital preservation and records management'[3] and will see the LOCKSS system deployed at twenty-five selected HE libraries. The participating libraries will be supported by a LOCKSS Technical Support Service based at the Digital Curation Centre (DCC)[4], through the provision of technical advice, software development and training. Negotiations to obtain publishers' permissions to allow LOCKSS-based archiving will be undertaken by Content Complete Ltd[5], JISC's negotiation agent for the NESLi2 national e-journals initiative.[6]

The assessment and development of services for continuing access to e-journals used by the community is a key area of work for JISC. Archiving clauses have been included in the 'Model Licence' for e-journal subscription agreements to provide the libraries with some reassurance of continued access to the content they have paid for.[7] This pilot is a step towards a practical implementation to fulfil these clauses and will help to ensure that content, once purchased, remains accessible.

LOCKSS and the problem of e-journal archiving

The move in recent years towards provision of scholarly journals in electronic form has greatly enhanced the access to, and availability of, scholarly publications. However, the arrangements for preserving long-term access to electronic journals are far from satisfactory. When subscribing to electronic journals, libraries no longer possess a local copy as they did with printed journals. They effectively lease the content of the electronic journals they subscribe to by remotely accessing it on publishers' servers over the computer network. The problem with this common practice is that access to entire back-runs of electronic journals could be lost to academic libraries when subscriptions are cancelled or when journals cease publication. The uncertainty of continuing access is a major barrier preventing libraries from moving to electronic-only subscriptions. The recent

endorsement of the statement 'Urgent Action Needed to Preserve Scholarly Electronic Journals' by organizations such as the Association of Research Libraries (ARL)[8] and the International Coalition of Library Consortia (ICOLC)[9] highlight the concern in the scholarly community over the long-term future of scholarly electronic journals. There is consensus that a solution to the e-journal archiving problem is urgently needed and that a technically and financially sustainable solution requires collaboration between libraries and publishers.

The LOCKSS initiative was designed to address the issues surrounding e-journal archiving and preservation. Over 80 libraries and 60 publishers worldwide have now joined the LOCKSS Alliance, working together to preserve persistent access to licensed content (both commercial and open access) from publishers to libraries. The open-source LOCKSS software is a peer-to-peer system that allows the libraries to collect, store, preserve, and archive authorized content locally. The local copies serve as back-ups and can be accessed when the publisher's site becomes unavailable. LOCKSS retains the libraries' traditional custodial role of scholarly information and allows libraries to 'own' the content they have paid for in much the same way as in the printed environment.

Among the approaches towards e-journal archiving, JISC has been keeping a watching brief on the LOCKSS initiative as a possible option to fulfil the archiving clauses in the model licence. The decision to establish a UK LOCKSS pilot is not a random one but the result of a series of previous activities which include a UK LOCKSS workshop in November 2003 to explore the technical development of LOCKSS and discuss its potential for preserving UK content.[10] An e-journal archiving feasibility study[11] undertaken in the same year highlighted the potential of LOCKSS and suggested that JISC carry out a detailed technical analysis of LOCKSS, which took place in 2004. The consensus was that LOCKSS is fully functional for e-journal archiving and that key advantages of LOCKSS are the little ongoing effort required to operate it, and the acceptance of many prominent academic publishers on the licensing and intellectual property front. In May 2005 the JISC Committee for Integrated Information Environment (JCIIE) gave approval to the JISC Executive to scope a programme of work that will allow the community to obtain hands-on experience with LOCKSS.

The UK LOCKSS Pilot Programme

The LOCKSS Pilot Programme aims first of all to raise awareness of the LOCKSS initiative. Moreover, by providing the participating libraries with the practical help to get started and to develop the skills needed to run their LOCKSS nodes beyond the pilot, it is hoped that the pilot will seed a self-sustaining base of LOCKSS users in the UK, who will collectively preserve a major proportion of the e-journals in common use in the JISC community. The pilot should also help to build experience in the community and to form a centre of expertise outside the US, which in turn will benefit the international LOCKSS community.

LOCKSS is one of the emerging developments engaged in e-journal archiving. Others include the forthcoming programme of legal deposit at the British Library, the Portico initiative currently supported by JSTOR, Ithaka[12], the Library of Congress, and the Andrew W Mellon Foundation[13] and the OCLC's Digital Archive.[14] Piloting LOCKSS should allow the community to evaluate its value and make an informed assessment of the desirability of future use of LOCKSS versus other alternatives.

The components of the LOCKSS Pilot Programme are outlined below. They were designed to focus on satisfying the conditions that must be met to preserve e-journals through the LOCKSS system:[15]

1. The publisher has granted permission
2. The publishing platform is supported by a LOCKSS *plug-in*[16]
3. There is a critical mass of libraries (at least six) that agree to collect and preserve the same title.[17]

LOCKSS Technical Support Service (LTSS)
A central UK support service is a crucial enabling factor to encourage the take-up of LOCKSS and reduce the learning curve of the participating libraries. The LTSS will build on the current infrastructure of the Digital Curation Centre and include a LOCKSS Technical Support Officer. The responsibilities of the LTSS include:

- first-line technical support
- development of publisher-specific plug-ins
- organization of training and awareness-raising events
- assistance to participating libraries in collection development
- purchase and installation of LOCKSS hardware

■ representation of UK libraries at LOCKSS Alliance.

Hardware with suitable configuration for LOCKSS nodes will be purchased collectively by the LOCKSS Technical Support Officer, who will set them up at the participating libraries' premises and liaise with local library and technical support staff to ensure that the machines work within the libraries' institutional network.

Publisher negotiation

The LOCKSS system uses a crawler to collect e-journal content from the publishers' web sites as it is published. Both written and machine-readable permissions from the publishers are required for this. Publishers are encouraged to grant libraries legal permission to cache and archive their content by means of the wording in licences or terms and conditions. In addition, a web page on the publisher's web site, called the 'publisher manifest'[18], that is recognizable by the crawler, contains a specific permission statement and permits the crawler to crawl, collect, and preserve the content. A publisher manifest is needed for each 'archival unit' (typically a volume) to be preserved through the LOCKSS system. Once permission for an archival unit has been granted by a publisher, it can be collected by all institutions with authorized access to it.

Content Complete Ltd will work together with the publishers to obtain the required permissions for the pilot programme. The negotiation will in the first instance focus on NESLi 2005 and 2006 publishers who have not yet agreed to make content available for collection and preservation by LOCKSS.[19] Negotiation with non-NESLi publishers will take place after they have been identified and agreed on by the participating libraries.

A legal appraisal of the archiving clauses in the JISC Model NESLi2 Licence will take place to ensure that these automatically and explicitly allow for future LOCKSS-based archiving. This work will be undertaken by the JISC Services Group.

LOCKSS Alliance membership

The LOCKSS Alliance[20] is a membership organization for libraries interested in using LOCKSS to build and preserve e-journal collections. The LOCKSS Alliance was established in 2004 and currently has 50 academic library members in the United States. The LOCKSS Alliance provides a co-ordinated infrastructure for software development and collection management. It is necessary for the continued existence of LOCKSS that libraries around the world support it.

A collective UK membership for the participating libraries in the LOCKSS pilot will be jointly funded by JISC and CURL to ensure UK priorities are factored into the US-led work and its future development. The participating libraries will enjoy the benefits of full members, including members-only premium content, plug-in applications for different publishing platforms and various training material and documentation. The UK's participation in the LOCKSS Alliance will not only distribute and share the efforts in working with publishers, building collections and software development, it will also allow access to the expertise of the LOCKSS team and transfer the knowledge and skills to the UK, helping to form a self-sustaining UK LOCKSS community and a centre of expertise outside the US.

Collection development

The pilot aims to recruit a critical mass of libraries to build a community in the UK that will provide the natural safe havens for jointly purchased e-journal titles and those of common interest. The above-funded components of the pilot were designed to free the libraries from all other tasks so that they can concentrate on the ones they are best placed to carry out – building and managing local e-journal library collections. This means using the web-based administration tools to configure the institutional LOCKSS node and to collect licensed and freely available content that fits the institutions' collection development policy and guidelines. LOCKSS requires very little system administration effort to run and, typically, a librarian is expected to spend at most three to five days a year to add content to the local collection and to upgrade the LOCKSS node.

The priority of the Pilot is to ensure all NESLi2 content is available via LOCKSS. However, the programme will also seek to include non-NESLi content important to the UK HE research community. Participating libraries will take part in survey(s) which will be carried out to identify and agree on titles/publishers of common interest, particularly those of interest to the UK that are at risk. Negotiations with the publishers of selected content will be carried out by Content Complete Ltd. Plug-ins needed for these publishers will subsequently be developed by the LTSS.

Since the objective of the pilot is to seed a viable UK LOCKSS community for the future, it is important that the libraries are committed to collection development and maintenance of their LOCKSS nodes beyond the pilot. This is one of the important criteria that will be used to evaluate libraries' applications. The participating libraries are required to develop the necessary skills locally so that they can continue to run their LOCKSS node once the pilot had ended.

Conclusion

The LOCKSS Pilot Programme provides a unique opportunity for libraries to take part in a community programme which will ensure that the journals they have subscribed to in digital form can continue to be accessed. By undertaking publisher negotiation and providing comprehensive support covering the technical, training and advocacy aspects, the pilot will ensure that the participating libraries have a much less steep learning curve and are fully supported in acquiring valuable skills in e-journal collection management.

It is hoped that the benefits of LOCKSS will become evident to the participating libraries and they will continue to maintain their LOCKSS nodes as a self-sustaining community using the skills gained through the pilot.

While keeping a watching brief on the other emerging approaches, piloting LOCKSS will allow JISC and the community to make an informed assessment of the best approaches towards e-journal archiving and preservation.

Notes and references

1. Joint Information Systems Committee: http://www.jisc.ac.uk

2. LOCKSS Programme: http://lockss.stanford.edu/

3. Digital Preservation and Records Management Programme at JISC: http://www.jisc.ac.uk/index.cfm?name=programme_preservation

4. The Digital Curation Centre is a UK-based centre funded by JISC and the UK e-Science Programme to support expertise and practice in data curation and digital preservation across communities of practice: http://www.dcc.ac.uk/

5. Content Complete Ltd: http://www.contentcomplete.com/

6. NESLi2: http://www.nesli2.ac.uk/

7. Archiving options suggested by the Model Licence are to make content available on servers of the Publisher or a third party, or for an archival copy to be delivered to the Licensee or to a central archiving facility operated on behalf of the UK HE community. Clause 2.2.2 of the Model Licence reads: 'After termination of this Licence, the Publisher will provide the Licensee and its Authorised and Walk-in Users with access to the full text of the Licensed Material which was published and paid for within the Subscription Period, either by continuing online access to the same material on the Publisher's server or a third party server or by supplying an archival copy in an electronic medium mutually agreed between the parties which will be delivered to the Licensee or to a central archiving facility operated on behalf of the UK HE community without charge. For the avoidance of doubt continuing archival access (in whatever form) and use shall be without charge and is subject to the terms and conditions of this Licence.' Other archiving-related clauses are 5.4.1, which refers to the need for an archive for licensed material and 5.4.2, which reiterates the need for continued access following termination, as specified in 2.2.2. See: http://www.nesli2.ac.uk/NESLi2_Licence.doc for full text of the Model Licence.

8. http://www.arl.org/arl/pr/presvejrnloct05.html

9. http://www.library.yale.edu/consortia/ICOLCPresRelease11-16-05.htm

10. UK LOCKSS Workshop report: http://www.jisc.ac.uk/uploaded_documents/LOCKSSWorkshopReport.doc

11. Archiving E-journals Consultancy – Final Report: http://www.jisc.ac.uk/uploaded_documents/ejournalsfinal.pdf

12. Ithaka: http://www.ithaka.org/

13. http://www.portico.org/

14. http://www.oclc.org/digitalarchive/default.htm

15. http://lockss.stanford.edu/librarians/building.htm

16. The plug-ins are software modules that use journal-specific information to make more efficient the searches for new content and for damage to preserved content. Each online publishing platform will require a separate plug-in module.

17. Tentative data suggests that each title should have at least six redundant copies on the LOCKSS system.

18. A *publisher manifest* example:
    http://www.lockss.org/pubmanifest.shtml

19. A list of participating publishers in NESLi2 is available at:
    http://www.nesli2.ac.uk/offers.htm.
    Some NESLi publishers are already participating in the LOCKSS initiative as a result of negotiations that took place in the US and elsewhere.

20. The LOCKSS Alliance:
    http: /lockss.stanford.edu/alliance/alliance.htm

Article © Helen Hockx-Yu

■ **Helen Hockx-Yu**
Programme Manager
JISC Office
Kings College London
Strand Bridge House (3rd Floor)
138–142 Strand
London WC2R 1HH, UK
Tel: +44 (0)20 7848 1803
Mobile: 07813 024633
E-mail: h.hockx-yu@jisc.ac.uk

To view the original copy of this article, published in *Serials*, the journal of the UKSG, click here:

http://serials.uksg.org/openurl.asp?genre=article&issn=0953-0460&volume=19&issue=1&spage=47

For a link to the table of contents for the issue of *Serials* in which this article first appeared, click here:

http://serials.uksg.org/openurl.asp?genre=issue&issn=0953-0460&volume=19&issue=1