

DPC Briefing Day on eJournal Preservation

RIBA, London, 30 October 2013

About the event

'Preservation, Trust and Continuing Access for e-Journals' is the latest in the Digital Preservation Coalition's (DPC) series of Technology Watch Reports. Written by Neil Beagrie, and published in association with Charles Beagrie Ltd., this report was released on 30th October at the DPC's 'e-Journals Summit' at the RIBA headquarters at 66 Portland Place, London.

The programme can be found at <http://www.dpconline.org/events/details/67-eJournals-Oct-2013?xref=73>. Below are some notes of the event intended to provide an informal briefing for members of the DPC not able to attend in person. They only represent the sections in which Angela Dappert was able to make notes. For an authoritative and comprehensive report, readers are encouraged to contact the speakers directly.

Read Neil Beagrie's report ['Preservation, Trust and Continuing Access for e-Journals'](#) here.

Presentations

Neil Beagrie, Charles Beagrie Ltd.

'Continuing' or 'perpetual access' is used to describe the rights that subscribers have. This is regardless of continuing subscription later. 'Long-term preservation' is about processes and procedure keeping content accessible.

Different types of service of Dark / Dim / Light Archives and Hosting Services. They vary in terms of current access, digital preservation, continued access, long-term access

Move to e-journal means convenience, potential space savings, improved access and user expectations that can only be met electronically

Problem: libraries often only license access rather than holding content.

Solutions: CLOCKSS, UK Legal deposit and copyright libraries, KB e-Depot, Portico

PORTICO, LOCKSS, publisher platforms

Archiving knowledge can be found in the KEEPERS registry

The key concerns are: content coverage, access to what, when and how long; library responsibilities w/r to cost, time, expertise and funding; technical approaches

Progress since 2007:

2006 ARL/CLIR study surveyed the landscape (Ann Kenney). JISC has performed archiving studies since then.

Recommendations for libraries, for publishers, and for eJournal archiving programs

Michae Seadle's analysis of archiving of eJournal listed in Directory of Open Access Journals

Trust is based on Transfer Code of Practice, Keepers Registry, model licensing clause and their enactment

Peter Burnhill, Edina, The Keepers Register

Ease and continuity of access to the scholarly and cultural record are critical aspects of eJournal preservation. Libraries subscribe to ease of access, but not necessarily to provision of perpetual access.

Peter investigated how many ISSN-known eJournals are preserved in repositories. The number has gone up since 2011 from 17% to 19 % in 2013. This seems less progress than it actually is, given that the total number of eJournal titles has gone up by 30+%. Supporting information for Peter's talk can be found in his blog 'Generating Actionable Evidence on E-journal Archiving'

<http://thekeepers.blogs.edina.ac.uk/>

As with print, a real challenge keeping tabs on the long tail of ejournals from smaller publishers / societies

Marcel Ras, KB e-Depot

STM related content with international focus; operational goes alongside DP research – feed into each other; transition to being a digital library; redesigning the technical infrastructure through redeveloping the system;

New archival infrastructure; new business model (large, small and OA publishers); trigger event and post-cancellation access need to be worked out

'Safe places network' means collaboration in coverage and research: with Portico, CLOCKSS and ??

Enhanced publications are much harder to deal with than basic publications

OA is also a challenge

New Business model in which all stakeholders take financial and other responsibilities (publisher, library and govt). But how do you deal with OA materials – who should pay?

New ingest workflows in 2015; agreements with 80 publishers in 2018; prioritising Dutch Big Deal publishers;

Audit and certification process planned for 2015

Mirror for other partners in case of emergency

Delft University Library are helping to define what to keep: the problem are OA and small scale publishers

Who should benefit: researchers, libraries, govt,

Kate Wittenberg, PORTICO's perspective

Needed: Transparency, reliability with trigger responses, willingness to be audited and certified; willingness to collaborate to leverage each other's work; research into future preservation needs;

PORTICO: 220 publishers, 17k journals; eBooks: 54 publishers; D-Collections; participating Libraries: 851; triggered titles: 10

Partnership with BL; pilot in 2012 waiting for legal deposit to be passed in April 2013; roll out preservation copies to access in reading rooms

Preservation at scale workshop at iPres drew many participants; a lot of organisations had to ramp up to work on scale;

Portico and LOCKSS will lead a discussion at the annual CrossRef members' meeting to involve publishers, especially small publishers who do not participate in preservation services (13 Nov 13 in Boston)

New forms of scholarly content

Exponential growth of data in need of preservation (that is not a supplement) but data and executable files that are integral to the publication

New organisations, services, technologies are popping up and are not well understood but need to be taken seriously. They are not competition, but need collaboration. (WHAT ARE THEY?)

Adam Rusbridge, UK LOCKSS Alliance

Each institution spends about £2,000 per month on a LOCKSS box. Stanford founded; different levels: libraries run infrastructure; EDINA supports, Stanford does content processing and development

Libraries have cost and continuing access concerns as main worry

Model license for moving to e-only and discard traditional, and for support for local archiving;

Local collections are built. They have different local context; post-cancellation for current collection material post 1990 is wanted; for preservation it is more for older materials; for short-term access recent content; etc.

Technical principles: Libraries have to reduce impact of failing service providers

Having custody of content under local jurisdiction is important

LOCKSS: providing access statistics; Shibboleth; Publishers get first refusal of access supply; in some cases new material is served from the publisher and older material from LOCKSS

In the longer term it is an infrastructure platform; other deployment types, management, technical models (WHAT ARE THEY ????)

Make it easier to illustrate/visualise information and how to use the tools; when do you lose access; what do you have access to.

Suite of tools that contribute to the use cases: Transfer code of practice; KB+ to show changes to subscription; JUSP, KEEPERS Register, COUNTER statistics

3 case studies can be found on the LOCKSS blogs in 2013 which highlight use cases

Randy Kiefer, CLOCKSS

Use LOCKSS architecture

Has been there for 3 years; core values still hold

Community governed 12 libraries and 12 publishers; equal vote for gathering the right content; global approach to decentralised preservation all over the world – distributed and duplicated;

Commitment to open access

7-8 nodes minimum; they now have 12, going on 13

704 participating libraries, 170 publishers; 12 nodes

A studyboard for the long tail (OA and society publishers) has been established

Challenge: the one-time cost is the same as for large publishers (PROVIDE STANDARDS FOR SMALL PUBLISHERS ???)

Luckily the (surviving) browsers are backwards compatible allowing access to old material; writing plugins is very expensive for smaller publishers

Ability to charge publishers is based on their revenues; what small publishers pay does not cover CLOCKSS' cost; large publishers and libraries are subsidising them: How do we deal with the long tail is the key question

What is an OA trigger if publishers disappear but the platform remains

Stanford and Edina have taken on hosting roles

Define who is not an archive (commercial hosting platforms)

Preservation in a cooperative way

No single organisation can preserve it all

Discussion

Do you catalogue / model at article or title (stream) level at ISSN to see whether you have all the volumes and issues? => both is needed for different access models

Peter Burnhill: There is a working group for small publishers investigating the issue

Neil Grindley: What do we know about these small publishers; how many are classified as scholarly? What groupings can be made? How do they produce these publications? How do we get this intelligence? => Ask libraries for their subscription lists; outreach to a lot of education publishers; use publisher multiplier organisations who bring a lot of publishers with them;

Collection development policy has changed – it has enlarged and is shaped differently;

Journals and eBooks are the same with respect to technology

Symbiosis with OA: if they go away they are still OA

Karen Wittenberg: PORTICO intern will go through CrossRef of unpreserved journals to analyse the long-tail publishers; Mellon is funding is as well looking into the 85% unpreserved titles

Brian Hole, Ubiquity Press: You should only be a member of organisations if you follow a minimal best practice

Adam Rusbridge: use criteria as to how journals should be submitted. At point of deposit they should specify what agency they belong to. OJS problem is its customisability- have preservation compliant vanilla interface. They have OAIPMH interface. But permissions are the stumbling block

Paul Conway: lack off incentive structures is well documented. What are offers that cannot be refused? Batches of quality; Big deal publishers had incentives that threatened them for non-compliance.

One big incentive is library requirements for publishers (must have DOI, etc.) Publishers are effectively encouraged by library customer requests.

Liam Earney: The long tail comes up in all discussions: usage statistics, DP, etc.; How easy is it to have libraries make it a priority? They have competing priorities. There needs to be a return to stewardship (See Peter's comment)

Liz Mallett: OU is not a physical library- don't necessarily think about themselves as a holding library, (Peter Burnhill: but as an access library which can be the appropriate role for a subset of libraries and has been in the pre-digital world)

Michael Day, BL: E-Legal deposit constrains the BLs role; they can collect but can make it available in a limited way. Access only in the reading rooms.

Brian Hole: Ubiquity is a small publisher, spun out of UCL in 2012; also service provider to universities; power university presses; provide production support;

Adam Rusbridge: new features to be offered: holdings comparison feature: upload list of titles and see which have been preserved: used to make decisions on collecting; tools to be released this year; often libraries don't report back to help improve the tools;

Keepers registry has been funded in project mode; will change to have a service funded model; R&D and service development need to be dealt with separately; This is a UK decision; the rest of the

world needs to develop international objectives and plans. Using IFLA 2014 as a venue and target time can be a goal.

Peter Burnhill: 2 new Edina features are technically possible, but need to have determined how one best benefits from them: any person can upload ISSN – KR normalises it to title level; How much info do you give back? It may be too much. How do you economically go through title by title to work out the archival status in order to make collection decisions?

3rd party (union catalogue) can select serial status to find out the archival status.

The KEEPERS are very diverse. The extent of how well they deliver varies. They only have to say that they have the intent to preserve, not how well they will do it. Keepers Registry does not want to audit and certify, but keepers should show certificates if they have them.

Tim Devenport: persistent identifiers. Objects of desire defined by DOIs. Emphasis of containers of volumes and issues is not as important. Now much is published before it is paginated. Assign DOI before the linking to ISSNs and containers happens.

EDItEUR developed the initial metadata structure depending on volumes and issues. They realised that they should have spoken about releases instead in order to have a more flexible organisational structure.

CLOCKSS have archival unit delimiters e.g. a year under an aggregate title as minimum structure. If publishers make modifications in the next year they need to pay duplicate fee. But it cannot be floating data without any context.

Peter Burnhill: Holdings statement: Peter argues that a year is an attribute. ISSN goes to things that issues content in parts – this could be a repository. ISSN is great for leveraging/finding out what you don't have. By concentrating on article and dismissing the hierarchy (year, title etc.), it is difficult to know what is missing. Aggregates make the information layers more tractable. It is an ergonomic feature. URNs, etc. can also be used

Tim Devenport: UKSG and others run training courses. There are also NISO webinars offered.

Ann McDonald: There is not enough technical training offered in UK post-graduate programs

Paul Conway: in the US there are 30 iSchool, not all have depth; 5 full-blown post-graduate programs. In 2 years you can have more technical breadth than in UK 1 year programs. Institute for Museums and Libraries IML funding push helped; same applies to research data management; Provided practical internships; technical infrastructures.

Paul Conway: message to academic libraries in the DPC Techwatch Report underplays the importance of collaborative work – don't do it alone – assign some staff to this activity -

About this document

Version 1	Written on day	31/10/2013	AD
Version 2	Distributed	31/10/2013	To DPC members

Document Distribution Note
Release to Members: Immediate
Release to Public: 31/10/2013

