e-Journal Archiving: Changing Landscape

Lars Bjørnshauge, Director
European Library Relations, SPARC Europe & Managing Director, DOAJ

Mark Jordan, Head of Library Systems,
Simon Fraser University Library & Public Knowledge Project (PKP)

Bernie Reilly, President, Center for Research Libraries

Oya Y. Rieger, Associate University Librarian
Digital Scholarship & Preservation Services
Cornell University Library

CNI Fall 2014 Membership Meeting
December 8, 2014


Life of eJournals

- Subscription cancellations and modifications
- Moving from hybrid subscription to digital only
- Changes in publishers’ business or service models

digital preservation aims to ensure the **usability**, **authenticity**, **discoverability**, and **accessibility** of content
• How do we factor in preservation status in our selection & collection building efforts?

• How much do we know about Portico, Lockss, Clockss, etc. and our roles in the success of these operation?

• What does “perpetual access” mean?

• **2CUL Study, Spring 2011**
  – Only 13%-26% of e-journal titles preserved by LOCKSS or Portico

• **Keepers Registry Study, Fall 2012**
  – Only 23-27% of e-journals with ISSNs preserved by any of 7 preservation agencies
Scale - Estimate

• 200,000 e-serial titles
  – 121,000 titles in Cornell

• 113,000 ISSNs assigned
  – 22,000 preserved
    • 13% preserved in LOCKSS or Portico

Strategies for Expanding E-Journal Preservation

2GUL Project Team

- Principle Investigators: Oya Rieger (Cornell), Bob Wolven (Columbia)
- Project Lead: Joyce McDonough, Director, Continuing and E-Resource Management, Columbia
- E-Journal Preservation Librarian: Shannon Rieger
- Project Advisory Team: Robert Wolven and Oya Rieger (the Co-Principal Investigators), Joyce McDonough (Director of Continuing and Electronic Resources Management at Columbia), Bill Kata (Director, E-Resources, Serials and Post-Cataloging Services, Cornell), Janet Gantz (Director of Preservation at Columbia), Brock Wite (Director of Library Information Technology at Columbia), Kizer Walker (Director of Collection Development, Cornell).

Project Goals

Funded by the Mellon Foundation, the project aims to identify priority content from the perspective of the research library community and make significant progress towards increasing the number and extent of e-journal archiving by major preservation programs. Ultimate goal is to develop and test methods to facilitate the continued expansion of e-journal preservation through ongoing assessment of priorities and documented practices to encourage and expand community engagement. Additional Information.

Related Papers & Presentations

- 2GUL Initiative Home Page: http://2gul.org
Part 2: Mellon-Funded Project Objectives

- Establish preservation priorities
- Contact aggregators and publishers
- Promote:
  - models for distributed action
  - model license language
  - standard means of recording access rights for preserved titles

Request to Publishers of Journals with no Preservation Arrangement

- Agree, 17, 14%
- Questions, 10, 8%
- Decline, 11, 9%
- Considering, 2, 2%
- No Response, 82, 67%

* Very Preliminary Findings *
Perspectives from Publishers & Aggregators

- What counts as preservation?
- Who will pay for preservation costs?
- I am taking care of it
- Do not have archival rights
- Worried about unauthorized access
- Others?

The Keepers Registry
supporting long-term access to journal content

Search

New!
Sign up to test the new Member Services

Current Statistics
- Journals with volumes reported as 'gestated and archived' by at least one Keeper: 25,924
- Journals with volumes reported as 'gestated and archived' by three or more: 3,541

Work on other journals (and volumes) is 'in progress'.

News and Events
20th March 2014: Data from the new keepers [View]

2nd December 2013: Library of Congress Join The Keepers Registry [View]

28th October 2013: Generating Actionable Evidence on E-journal Archiving [View]

25th October 2013: Further International Exposure for The Keepers Registry [View]

20th August 2013: The Keepers Registry in Singapore [View]
The following organisations are the Keepers of digital content, working on your behalf to ensure long-term access to the scholarly and cultural record. They provide the registry with information on their archival holdings, ordered by most recent update (date of which is shown):

- Archaeology Data Service (2 October 2014)
- British Library (2 October 2014)
- Helsinki/Trinity (1 October 2014)
- Library of Congress (1 October 2014)
- Schomburg Portal (1 October 2014)
- CLOCKSS Archive (25 September 2014)
- Global LOCKSS Network (26 September 2014)
- Portico (26 September 2014)
- National Library Board, Chinese Academy of Sciences (6 January 2014)
- e-Depot (24 April 2012)

The most up-to-date information may be available on the Keeper's website.
What have we learned so far from the Center for Research Library’s Trusted Repository Audit and Certification (TRAC)?

CNI Project Briefing
December 8, 2014

Bernard F. Reilly
CRL

Three Things Are Important:
Three Things Are Important:

Optics

Three Things Are Important:

The Process
Three Things Are Important:

Realism

Next Steps:
Next Steps:

More transparency
Next Steps:

Analysis
Next Steps:

Engagement

CRL

2CUL Study

– Cornell

• 18% of our 8,279 OA e-journal titles in the catalog have identifiable preservation strategies

• 2011 study by Michael Seadle
  – 8% LOCKSS & CLOCKSS
  – 5% Portico
DOAJ Efforts in Increasing Awareness of OA Publishers on Preservation Matters

(Remote) Presentation for the CNI 2014 meeting, December 2014
Lars Bjørnshauge, DOAJ
www.doaj.org

Lars Bjørnshauge
Brief Background

• Launched in May 2003 at Lund University, Sweden with 300 journals
• Since January 2014 operated by IS4OA
  —non-for-profit company registered in the UK
www.is4oa.org

Founded by
Caroline Sutton,
Alma Swan &
Lars Bjørnshauge

The challenges

• Bringing DOAJ up to date and respond to demands and expectations
• Help OA-publishers to enter the mainstream of scholarly publishing
• Involve the community in the operations of the DOAJ
• Develop sustainable funding
Tighter criteria

• We have worked with OASPA, COPE & WAME and agreed on principles of transparency and best practice in scholarly publishing

• The new DOAJ criteria are very much inspired by the above principles.

The new criteria!

• Much more information is required regarding:
  – The editorial process
  – The peer-review process
  – Rights (reader rights, reuse rights, remixing rights etc.)
  – Persistent identifiers
  – Discoverability
  – Archiving
DOAJ –
A diverse collection

- Rather few OA-publishers with 50 or more journals
- Thousands of publishers publishing 1-2 journals
- Several not for profit aggregaters in place:
  - PKP (Public Knowledge Project – OJS)
  - SciELO (Latin America)
  - Redalyc (Latin America)
  - INASP JOL (Journals on Line)
- But still thousands of “homeless” journals

The long tail
Preservation

Does your organisation or your journal(s) have an arrangement for LPTA or partake in any LPTA program? Would you be interested in DOAJ providing/facilitating a fee-based LPTA service?

<table>
<thead>
<tr>
<th>Answer Options</th>
<th>Response Percent</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes (Add the name in box below)</td>
<td>13,8%</td>
<td>127</td>
</tr>
<tr>
<td>No - I'm not interested</td>
<td>41,3%</td>
<td>381</td>
</tr>
<tr>
<td>I am interested. Tell me more!</td>
<td>49,3%</td>
<td>455</td>
</tr>
<tr>
<td>LPTA scheme name:</td>
<td></td>
<td>97</td>
</tr>
</tbody>
</table>

The challenge related to preservation

• Many, many journals
  – lack the financial & technical resources to go beyond just publishing the content.
  – haven´t addressed the preservation issue yet, but would like to do so, provided smart and cheap solutions are available.

• **Preservation** of OA-journals (as well as provision of persistent identifiers and standardized copyright information) is a significant problem.
So – what are we doing?

• The new criteria and archiving/preservation:
• Journals now have to provide information about:
  – Whether the journal has an archiving arrangement
  – If, yes: with which organisation &
  – specify the URL providing this information

  – Answering No, does not disqualify the journal from being listed

Archiving/Preservation

• From the new application form - http://doaj.org/application/new

  23) What digital archiving policy does the journal use? *
  [□] No policy in place
  [□] LOCKSS
  [□] CLOCKSS
  [□] Portico
  [□] PMC/Europe PMC/PMC Canada
  [□] A national library
  [□] Other

  Select all that apply. Institutional archives and publishers' own online archives are not valid

  24) Enter the URL where this information can be found *
  [ ]

  This field is optional if you have only selected “No policy in place” above

  25) Does the journal allow anyone to crawl the full-text of the journal? *
  [□] Yes
  [□] No
Archiving arrangement – a qualifier for the DOAJ SEAL

A note about the qualifiers for the DOAJ Seal

DOAJ promotes best practice in Open Access publishing. To that end, we are introducing the DOAJ Seal for Open Access journals.

The qualifiers for the Seal will highlight features related to the openness, indexability and discoverability of the journal and, as such, have nothing to do with the scholarly quality of the papers published in the journal.

To qualify for the Seal the journal must:

1. have an archival arrangement in place with an external party. **(scroll to digital archiving policy question)** If ‘No policy in place’ is selected, the journal will not qualify for the Seal.
A lot is going on!

- CLOCKSS - +/- 1000 DOAJ listed journals are in the CLOCKSS programme
- PKP:
  - Private LOCKSS Network
  - Agreement with CrossRef will cater for archiving
- Still a lot of small OA-publishers who need help.
- We help the publishers to address the archiving issue – that is an important step in the right direction.

Summing up!

- Journals are now required to provide information about archival arrangements
- This information will be publicly available
- This is a considerable motivator for journals to move forward on the archiving issue
- Archiving is not a requirement of being listed – for the moment!
- But it will probably be a requirement for the next revision of the criteria – due 2016.
Thank you for your attention!
Lars Bjørnshauge
lars@doaj.org

The Public Knowledge Project
Private LOCKSS Network

Mark Jordan, Head of Library Systems,
Simon Fraser University Library & Public Knowledge Project (PKP)
Global Open Journal Systems usage

OJS journals with 10 or more articles
Background to the network

• Demand for/interest in preservation exists among journal managers
• Some OJS journals exist in the Global LOCKSS Network and in other PLNs
  – IBICT
  – COPPUL
  – ADPNet
• Most OJS journals will never be preserved in the Global LOCKSS Network (GLN)

Why not the GLN?

• Currently, the GLN preserves content from 111 OJS titles (812 volumes and 28 OJS instances)
• Many “publishers” to negotiate with
• Quality, consistency, and legitimacy sometimes unclear
• Not top priority for LOCKSS Alliance libraries
Rationale for PKP PLN

- Gap in preservation services for OJS journals
- PKP is starting to develop coordinated services for the OJS community, e.g. CrossRef, Altmetrics
- Several PKP Development partners are members of the LOCKSS Alliance
- A PLN for OJS journals is consistent with LOCKSS’ “community” focus

Service model

- Any OJS journal that meets minimum requirements can participate at no cost
- Primary journal contact agrees to sign off on several requirements
- Content is included in PLN automatically
- After “trigger event,” content will be made available under the CC-BY license
Staging server
Provides list of issues to be made accessible.

1. PKP PLN staff, with the assistance of automated tools, decides to trigger preserved content.

2. A list of the journal issues to make accessible is generated from the staging server.

3. The list is used by the import tool to extract the content from the PLN and import it into one or more OJS access instances.

4. The import tool can generate KBART title lists for use in link resolver knowledge bases.

5. Users access preserved content directly at OJS instances or indirectly via link resolvers.

Journal Manager agrees...

- I have the legal and contractual authority to include this journal’s content in a secure preservation network and, if and when necessary, to make the content available in the PKP PLN.
- I agree to allow the PKP-PLN to make post-trigger event content available under the CC-BY (or current equivalent) license.
- I agree to allow the PKP-PLN to include this journal’s title and ISSN, and the email address of the Primary Contact, with the preserved journal content.
- I confirm that licensing information pertaining to articles in this journal is accurate at the time of publication.
- I acknowledge these terms may be revised from time to time and I will be required to review and agree to them each time this occurs.
- I agree not to violate any laws and regulations that may be applicable to this network and the content.
- I agree that the PKP-PLN reserves the right, for whatever reason, not to preserve or make content available.
Trigger events

- **Journal-initiated**
  - Event: Primary Journal Contact indicates that the journal content should be triggered for access
  - Response: PLN staff will confirm with the Primary Journal Contact that a trigger event is justified

- **Automated**
  - Event: Journal has gone dark (no response) for two months, and we receive a negative response from email verification to Primary Journal Contact
  - Response: PLN staff will investigate whether the journal is accessible and assess whether or not to trigger the content for access

OJS plugin

- Distinct from existing OJS LOCKSS plugin
- Works specifically with the PKP PLN
- Provides a user interface to journal manager for sign off and for viewing preservation status
- On issue publication, packages content and informs the staging server via SWORD deposit
- Shipped with OJS 2.4.5 in September
Organizational model

- Acceptance by participating PLN nodes
- PLN Advisory Panel - represent the broad range of potential stakeholders and participants in this service
- Light administrative and organizational structure
- Technical and administrative coordination on behalf of PKP by SFU and UBC Libraries
  - PKP PLN coordinator (Bronwen Sprout)
  - PKP PLN primary technical consultant (Mark Jordan)
  - PKP PLN overall lead (Brian Owen)

Charter nodes in the PLN

- Standard Private LOCKSS Network
- Nodes at:
  - University of Alberta
  - University of British Columbia
  - Indiana University
  - University of Pittsburgh
  - University of Victoria, and
  - Simon Fraser University
- Pilot will start with 2 TB storage
- Configuration managed by LOCKSS-O-Matic
Pilot phase

- **Stage 1**: (Sept 2014 – April 2015)
  - release of OJS version 2.4.5
  - test the new features of OJS 2.4.5 relevant to the PLN
  - test the “staging server”
  - finalize organizational components of the PLN
  - participating institutions must identify two or three OJS journals they currently host that will be good candidates to join the PLN
- **Stage 2**: (April – July 2015)
  - extend testing to LOCKSS-O-Matic
  - participating institutions must procure and set up a LOCKSS server

Advisory panel

- Lars Bjornshauge, Directory of Open Access Journals (DOAJ)
- Sioux Cumming, International Network for the Availability of Scientific Publications (INASP)
- Sarah Lippincott, Library Publishing Coalition (LPC)
- Alex Mendonça, SciELO
- Susan Murray, African Journals Online (AJOL)
- Victoria Reich, LOCKSS
- Heather Joseph, SPARC
Launch

**Summer 2015:** The PKP PLN will be ready to extend preservation service to any interested OJS journal anywhere in the world.