Seasons of Change:
Digital Preservation in an Ever-Changing Digital Environment

Charleston Conference
November 8, 2018
Our Presenters:

Heather Staines, Hypothesis, is Director of Partnerships, working with publishers and platforms to enable annotation.

Shannon Keller, New York Public Library, Helen Bernstein Librarian for Periodicals and Journals, Collections and Research Services

Bob Boissy, Account Development Director, Springer Nature
NASIG Digital Preservation Task Force

Raise awareness of digital preservation.

Develop some best practices for the industry.

Marketing digital preservation to a broad range of library administrations and publishers.

Chair: Shannon Keller (NYPL)

James Phillpotts (Oxford University Press), Wendy Robertson (University of Iowa), Heather Staines (hypothes.is), Zach Van Stanley (University of Denver)

Board Liaison: Ted Westervelt
Links to Digital Preservation Task Force Documents

- Digital Preservation 101
- Guide to The Keeper’s Registry
- Talking Points and Questions to ask Publishers about Digital Preservation
What is Digital Preservation?

“Digital preservation refers to the series of managed activities necessary to ensure continued access to digital materials for as long as necessary” (definition from the Digital Preservation Coalition). It is a suite of services and an ongoing process.
LOCKSS (Lots of Copies, Keep Stuff Safe)

A distributed network to preserve content as it appears on the publisher’s site, with regular integrity checks of the data.

Global LOCKSS Network (GLN) preserves content:

- Available online (open access and subscription-only)
- No cost, but limited space.

Private LOCKSS Network (PLN)

- Examples: MetaArchive, PKP
- One type of software or a specific geographic area.
- Any digital content, not only books and journals.
CLOCKSS (Controlled LOCKSS Network)

CLOCKSS is:

- A **dark archive** of scholarly publications
- Preserved using LOCKSS
- A “trigger event” (content not available from any publisher) makes content freely available with a **creative commons license**.
- Publisher cost based on annual journal or ebook revenue.
- Libraries may join as a supporting members
- Membership is not required to access triggered content.
CLOCKSS for Library Publishers

- Digital Commons: Bepress Agreement in place. (32 publishers)
- OJS: No Agreement is necessary. (24 Publishers)
- Plug-ins for both platforms, so minimal set-up.
- Harvesting approach = publishers do not need to do work at their end.
- Other:
  - IUScholarworks from Indiana University Libraries,
  - Divided Society from Linen Hall Library (in Ireland)
  - Utrecht University Library Open Access Journals.
Portico

Portico is a community supported dark archive to ensure electronic content remains accessible for the long term.

- Primary access is a “trigger event.”
- Portico makes content available for use.
- Open content stays open.
- Cost to publishers is based on annual journal or ebook revenue.
- A library that publishes may choose to join both as a library and publisher.
Portico: Info for Library Publishers

- OJS: export plugins available for OJS 3.0 and 2.4.1+ (Currently from Github but soon PKP 3.1 plugin, so available through the OJS UI.)
- Digital Commons included in Portico
- Other systems: Portico will work with library/publisher to get the package how they need it to be; Portico can accept the metadata in a spreadsheet (template), then transform to XML
Consortial Trusted Digital Repositories

Some libraries have partnered to create a Trusted Digital Repository (TDR) meeting specific international standards (ISO 16363).

Currently six have been certified (including CLOCKSS and Portico):

- Canadiana.org
- Chronopolis
- Hathitrust
- Scholars Portal (a consortia of 21 university libraries in Ontario, Canada)
Other tips and resources:

Keepers Registry

Index of journals preserved by one or more archiving agencies.

- Librarians can check that titles in their collection have been preserved.
- As a publisher, heck that your titles are correctly listed with preservation coverage as you expect.

For additional information, see our Guide to the Keepers Registry.
Other tips and resources (continued):

Library of Congress’ Recommended Formats Statement

Provides good guidance for publications in all formats to make sure they will last through time.

Digital Preservation Handbook (by Digital Preservation Coalition)

An internationally authoritative and practical guide for:

- Managing digital resources over time
- Issues in sustaining access to them.
Digital Preservation: Library Perspectives

- Access, access, access
- Keepers Registry Analysis
- Unique at NYPL
Digital Preservation: Library Perspectives

Member Services
Welcome, Shannon Regan.
Our Member Services area provides additional tools for discovering reporting on the archival status of e-journal content.

Title List Comparison
Submit a new file of ISSN or view results
Discover the archival coverage of a list of serial titles important to you. Upload a file of ISSN and receive a report listing those being archived and those "at risk".
These reports have been used to assist with library decisions around print rationalisation and subscription management.

Machine to Machine Interfaces
Access login details and documentation
We provide APIs onto the Keepers Registry data.

User Profile
View and update your details
Update your email address, password and other details.
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):

- Scholars Portal (06 Jun 2018)
- British Library (04 Jun 2018)
- HathiTrust (04 Jun 2018)
- PKP Preservation Network (02 Jun 2018)
- Archaeology Data Service (24 May 2018)
- Global LOCKSS Network (23 May 2018)
- CLOCKSS Archive (22 May 2018)
- Cariniana Network (21 May 2018)
- Library of Congress (07 May 2018)
- Portico (03 May 2018)
- National Science Library, Chinese Academy of Sciences (26 Mar 2018)
- Swiss National Library (12 Mar 2018)
- e-Depot (20 Dec 2017)

The most up-to-date information may be available on the Keeper’s website.
Digital Preservation: Library Perspectives
In your opinion, with whom should the primary responsibility of preserving digital scholarly content lie?
Hot Takes from the NASIG Digital Preservation Task Force Survey

Ranking the barriers to expanding digital preservation initiatives at respondent’s institution:
(1=Highest Barrier, 5=Lowest Barrier)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>TOTAL</th>
<th>SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Institutional priorities</td>
<td>26.47%</td>
<td>21.57%</td>
<td>19.12%</td>
<td>18.63%</td>
<td>14.22%</td>
<td>204</td>
<td>3.27</td>
</tr>
<tr>
<td></td>
<td>54</td>
<td>44</td>
<td>39</td>
<td>38</td>
<td>29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>17.07%</td>
<td>30.24%</td>
<td>26.34%</td>
<td>13.66%</td>
<td>12.68%</td>
<td>205</td>
<td>3.25</td>
</tr>
<tr>
<td></td>
<td>35</td>
<td>62</td>
<td>54</td>
<td>28</td>
<td>26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Money</td>
<td>40.48%</td>
<td>27.62%</td>
<td>17.62%</td>
<td>6.19%</td>
<td>8.10%</td>
<td>210</td>
<td>3.86</td>
</tr>
<tr>
<td></td>
<td>85</td>
<td>58</td>
<td>37</td>
<td>13</td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technical complexity</td>
<td>3.85%</td>
<td>10.58%</td>
<td>25.96%</td>
<td>43.75%</td>
<td>15.87%</td>
<td>208</td>
<td>2.43</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>22</td>
<td>54</td>
<td>91</td>
<td>33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unclear understanding of</td>
<td>13.68%</td>
<td>10.85%</td>
<td>12.26%</td>
<td>16.51%</td>
<td>46.70%</td>
<td>212</td>
<td>2.28</td>
</tr>
<tr>
<td>responsibilities for</td>
<td>29</td>
<td>23</td>
<td>26</td>
<td>35</td>
<td>99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>digital preservation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
To update the background image:

• Go to 'View'
• Select 'Slide Master'
• Select the page with the image
• Right click on the image and select 'Change picture'
• Navigate to the location with the new image
• Select 'insert'

Please note the new image needs to be at least 19cm x 27.5cm to fit the area. If the image does not fit you will need to manually manipulate the image to fit.
First, Breaking News

- Springer Nature has made an agreement with CLOCKSS to preserve our eBooks
- This complements our existing arrangement with Portico
- As the world’s largest scholarly book publisher, this is another important preservation milestone for everyone’s sake >250,000 titles
The Debate

Pre-Print Archives
Primary publishers would prefer not to spend time and money preserving pre-prints

Why?
Why spend money on content that is not peer reviewed?
It is another copy of a document to get in the way of the version of record
There are lots of messy duplicates in pre-print servers
Pre-print servers are to meet immediate needs, not preservation needs

Digital preservation experts, librarians and other stakeholders want to make the effort

Why?
There are significant numbers of pre-prints that are never otherwise published
Especially in math and high-energy physics
It is best to be comprehensive
The FAIR Principles and Preservation

**Findable**: In the sense that it ought to be easy to see that any scholarly content has been faithfully preserved

**Findable**: In the sense that triggered/released or otherwise open content is easily discoverable

**Accessible**: In the sense that content that has been triggered for public release is accessible, and this means on a platform that is durable and optimized for discovery

**Interoperable**: In the sense that preserved content is always converted to new formats well before old formats are obsolete – this strikes me as the heart of digital preservation

**Interoperable**: Right now we are tied to Adobe

**Reusable**: Preservation of data sets in the scholarly record is what comes to mind here, together with really excellent description of what the data is, how it was collected, etc.

Example of really old data – that medium is called paper
Issues

Preserving content that has not been properly organized in the first place
Take the NISO Presentation and Identification of e-journals standard as an example
  Content needs to be stored under the original publication title
Is it the job of preservationists to clean and organize the scholarly record before preservation occurs?
Is it the job of preservationists to clean and organize the record that has already been preserved?
One day the whole scholarly record will be open access
  Will that much FAIRness lead to laxity on the preservation side?

Recently visited library with an access problem
Thank you

Robert W. Boissy, Director of Account Development (technical and promotional services to libraries)

Robert.Boissy@SpringerNature.Com

+1-781-244-7918