Building a Social Compact for Preserving E-Journals

Anne R. Kenney
NASIG, Washington, DC
May 2015
Why Worry?

- Digital-first ecology disrupts traditional roles and responsibilities
- Inadequate protection of digital content
- Uncertainty
- Relationship differences inhibit collaborative approaches
Graph 6

*** Ongoing Resource Expenditures vs. Total Materials Expenditures, 1993-2012
Yearly Increases in Average Expenditures

***Formerly Electronic Resource Expenditures. After 2011-12, figure represents increase in Average Ongoing Resource Expenditures.
"What percentage of your library’s materials budget is spent on the following items?" Average of percentages estimated by respondents, by institution type.

Ithaka S&R US Library Survey 2013
FIGURE 25
Averages of budget estimates and predictions. (2010 estimate and 2015 prediction data are from the 2010 survey, while 2013 estimate and 2018 prediction data are from the 2013 survey.)*
“How important to you is it that your college or university library provides each of the functions below or serves in the capacity listed below?” Percent age of respondents rating each as "very important," by institution type.
Digital preservation represents one of the grand challenges facing higher education. Yet... the responsibility for preservation is diffuse and the responsible parties have been slow to identify and invest in the necessary infrastructure...the digital portion of the scholarly record is increasingly at risk and solutions may require unique arrangements within the academy for sharing preservation responsibility.

Adapted from "Urgent Action Needed to Preserve Scholarly Electronic Journals," Don Waters et al, 10/2005
E-Journal Archiving Metes and Bounds: A Survey of the Landscape

by Anne R. Kenney, Richard Entlich, Peter B. Hirtle, Nancy Y. McGovern, and Ellie L. Buckley

4 Library Issues

- Sense of urgency
- Need for trusted independent archiving
- Content coverage and access conditions
- Resource commitment and competing priorities
- Need for collective response
Identifying Programs

- Not-for-profit program independent of publisher
- Commitment to archiving scholarly peer-reviewed e-journals
- Formal arrangements with publishers
- Program in place and sustainable
- Beneficial to many academic libraries
Original Group of 12

- CISTI (Canada)
- CLOCKSS
- OCLC ECO
- OhioLINK EJC
- KB e-Depot (Holland)
- kopal/DDB (Germany)
- LANL-RL
- LOCKSS Alliance
- Ontario Scholars Portal (Canada)
- PANDORA (Australia)
- Portico
- PubMed Central
Core Principles for Digital Preservation Repositories (2007)

1. The repository commits to continuing maintenance of digital objects for identified community/communities.

2. Demonstrates organizational fitness (including financial, staffing, and processes) to fulfill its commitment.

3. Acquires and maintains requisite contractual and legal rights and fulfills responsibilities.

4. Has an effective and efficient policy framework.

5. Acquires and ingests digital objects based upon stated criteria that correspond to its commitments and capabilities.
Core Principles for Digital Preservation Repositories (2007)

6. Maintains/ensures the integrity, authenticity and usability of digital objects it holds over time.

7. Creates and maintains requisite metadata about actions taken on digital objects during preservation as well as about the relevant production, access support, and usage process contexts before preservation.

8. Fulfills requisite dissemination requirements.

9. Has a strategic program for preservation planning and action.

10. Has technical infrastructure adequate to continuing maintenance and security of its digital objects.
Certified Archives for E-Journals

- Portico (2010)
- Scholars Portal (Ontario) (2013)
- CLOCKSS (2014)

Center for Research Libraries
www.crl.edu/archiving-preservation/digital-archives/certification-assessment
ARL Members Supporting Certified Archives

- CLOCKSS 21.6% (27)
- Portico 75.2%% (94)
- Scholars Portal 6.4%(8), but 50% Canadian ARLs
- LOCKSS program—more than half of ARL are or have participated
- In survey, 43.2% participated in 1 program; 27% in 2 programs; 27% in 3 programs
Content and Coverage

- Coverage continues to grow
  - CLOCKSS 202 publishers
  - Global LOCKSS 530 publishers
  - Portico 255 publishers
  - Scholars Portal 33 publishers loaded; 8 to be loaded

- Coverage uneven
- Difficult to create definitive list
- Major publishers well represented
- Significant duplication across services
- Redundancy vs greater content coverage
Publisher Overlap (2006)

No. of e-journal archiving initiatives

- 2 or more: 37
- 3 or more: 17
- 4 or more: 14
- 5 or more: 8
- 7 or more: 6
- 9 or more: 2

No. of publishers involved:
- 3 or more: 17
- 4 or more: 14
- 5 or more: 8
- 7 or more: 6
- 9 or more: 2
Keepers Registry Statistics (2015)

Serials with volumes reported as 'ingested and archived' by at least one Keeper: 27,463

Serials with volumes reported as 'ingested and archived' by three or more: 9,785
| 53. Journal of human genetics | 1435-232X (Online); 1434-5161 (Print) | Nature Publishing Group | In progress: v. 45-53, 55-59 | CLOCKSS Archive |
| Springer-Verlag | Preserved: v. 54 | CLOCKSS Archive |
| Springer Japan | Preserved: v. 43-52 | e-Depot |
| Springer Science+Business Media | In progress: Not yet processed. | Global LOCKSS Network |
| Nature Publishing Group | In progress: - | Portico |
| Springer | Preserved: v. 43-53 | Scholars Portal |
Inadequate Preservation Coverage (2010 JISC Study)

Which NESLi2 and NESLi2 SMP publishers are participating in e-journal archiving programmes?

<table>
<thead>
<tr>
<th></th>
<th>LOCKSS</th>
<th>CLOCKSS</th>
<th>Portico</th>
<th>Post-cancellation access route?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participating?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>17</td>
<td>23</td>
<td>23</td>
<td>9</td>
</tr>
<tr>
<td>Total as a percentage of total NESLi2 and NESLi2 SMP publishers</td>
<td>40.46%</td>
<td>54.76%</td>
<td>54.76%</td>
<td>21.43%</td>
</tr>
</tbody>
</table>
Inadequate Preservation Coverage (2011 2CUL Study)

- Cornell/Columbia (2CUL) Study, 2011
  - Only 13-16% of e-journal titles preserved by LOCKSS and/or Portico
  - No formal process for identifying e-journals for preservation consideration

- Scope of problem

  - 200,000 e-serial titles in the catalog
  - Ca. 80,000 currently published
  - Ca. 25,000 preserved
  - Ca. 20,000 newsletters, trade, etc.
    - (EMIS Emerging Markets, Factiva, etc.)
  - Ca. 8,000 CJK
    - (China Academic Journals, DBPIA, etc.)
  - Target Pool: 25,000-30,000 titles
Inadequate Preservation Coverage

- 2CUL Keepers Registry Study, 2012
  - Only 23-27% of e-journals with ISSNs preserved by any of the preservation programs
  - Considerable overlap in preservation from one program to the next

<table>
<thead>
<tr>
<th></th>
<th>CLOCKSS</th>
<th>e-Depot</th>
<th>LOCKSS</th>
<th>Portico</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLOCKSS</td>
<td>7928</td>
<td>63%</td>
<td>54%</td>
<td>88%</td>
</tr>
<tr>
<td>e-Depot</td>
<td>72%</td>
<td>7022</td>
<td>52%</td>
<td>84%</td>
</tr>
<tr>
<td>LOCKSS</td>
<td>66%</td>
<td>56%</td>
<td>6540</td>
<td>72%</td>
</tr>
<tr>
<td>Portico</td>
<td>65%</td>
<td>55%</td>
<td>44%</td>
<td>10737</td>
</tr>
</tbody>
</table>
Vulnerable Content

Third party publications
  ◦ e.g., Harvard Theological Review, via Cambridge

Aggregators
  ◦ e.g., Children’s Voice, via Proquest

Small publishers
  ◦ e.g., Asian Highlands Perspectives (plateauculture.org)

Open access
  ◦ e.g., Teiresias (McGill Classics Dept)

Historical
  ◦ E.g., Bulletin du comite des forges (in Making of the Modern World)
Coordinating Repository Programs

- Ensuring content coverage and redundancy
- Creating common finding aids, access mechanisms, registry services and reduced costs
- Exchanging information, strategies, software, documentation
- Providing formal succession plans
# Network Indicators

<table>
<thead>
<tr>
<th>Archiving Activity</th>
<th>CSI</th>
<th>ECO</th>
<th>EJC</th>
<th>KB</th>
<th>KOP</th>
<th>LA</th>
<th>LANL</th>
<th>NLA</th>
<th>OSP</th>
<th>PMC</th>
<th>PORT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exchange ideas and strategies</td>
<td>P</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Share planning documents</td>
<td></td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Share software</td>
<td></td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Coordinate content selection</td>
<td></td>
<td></td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Reciprocal archiving/off-site storage/mirroring</td>
<td>●</td>
<td></td>
<td>P</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>P</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Secondary archiving responsibility</td>
<td></td>
<td></td>
<td></td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>Shared facilities/resources</td>
<td>●</td>
<td></td>
<td></td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td>●</td>
</tr>
<tr>
<td>Other</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Conclusions

- All of the scholarly record must be preserved
- Coverage is uneven and insufficiently documented
- Much content is at risk
- No party can do this alone
- Preservation will require a social compact of all parties
- Viable preservation options are available and need support

- No single archiving program will meet all needs
- Programs need greater transparency and coordination
- Access should be reassessed
  - Standard means for recording access rights for preserved titles
  - Legislation needed to cover legal deposit
  - License terms inadequate
  - Moving wall vs trigger
Recommendations for Libraries

- Acknowledge this is a shared responsibility
- Press publishers and libraries to enter archiving relationships
- Share information about commitments and decisions
- Participate in at least one initiative
- Move beyond a registry of archived publications to identify at risk materials critical to scholarship
- Press existing preservation programs to meet their needs
- Lobby programs to participate in networks
Recommendations for Publishers

- Enter into relationships with one or more e-journal archiving programs
- Provide adequate information and data to archivers on coverage
- Extend liberal archiving rights in licensing agreements and to aggregators
- Reconsider access restrictions
Recommendations for e-Journal Archiving Programs

- Expand coverage to include vulnerable materials
- Present evidence of minimal level of services for long-term, well managed collections (open to audit, certified)
- Be overt and explicit about what is archived
- Assure appropriate property rights
- Negotiate with regard to eventual placement in the public domain
- Form a network of mutual support and interdependence
What Can NASIG Do?

- Raise awareness of the issue
- Endorse power of collective action
- Consider set of principles *and* actions
- Foment international cooperation
- Identify other parties, e.g., role for funding bodies in research, development and requirements
Thank you!
ark3@cornell.edu