STRATEGIES FOR EXPANDING E-JOURNAL PRESERVATION

Mellon Grant to Cornell and Columbia University Libraries (2CUL) November 2013 – April 2015

Oya Y. Rieger, Cornell University Library
Bob Wolven, Columbia University Libraries
HOW DID WE GET HERE?

- 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
# E-JOURNAL PRESERVATION OVERLAP

<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>
DON’T JUST TALK ABOUT IT
DO SOMETHING!
Only surface understanding of the preservation strategy and its implications at Columbia and Cornell

No formal process in place for identification of e-journals for preservation consideration

Lack of organizational leadership to bring together related parties from collections, IT, and scholarly communication teams
PROJECT OBJECTIVES

- Analyze the corpus of unpreserved e-journals
- Set criteria to establish preservation priorities
- Develop, test, apply, share methods
- Develop context-specific license language
- Promote models for distributed action
- Get more titles preserved!
PUTTING PLAN INTO ACTION

- Principle Investigators
  - Oya Rieger (Cornell), Bob Wolven (Columbia)

- Project Lead
  - Joyce McDonough, Director, Continuing and Electronic Resource Management, Columbia

- Project Staff
  - E-Journal Preservation Librarian (14 months)
Help us set priorities & validate methods

2CUL
- Columbia: Janet Gertz (Preservation), Breck Witte (Technology)
- Cornell: Bill Kara (Serials), Kizer Walker (Collection Development)

BorrowDirect
- Collection Development AULs
INFORMAL PARTNERS

- Preservation agencies: CLOCKSS, LOCKSS, Portico
- Preservation registry: Keepers Registry, OCLC
- Interested parties: CRL, ALA, Duke, MIT ...
SORTING THINGS OUT
“PRESERVATION”
A PRAGMATIC APPROACH

- Preserved with access
- Preserved with minimal access
- Protected
- In Process
- Not Preserved
EXTENT OF PRESERVATION
INITIAL IDEAS

- Ongoing preservation (or ‘complete’ for discontinued/defunct journals)
- Substantially preserved with gaps
- Small portion preserved
- Not preserved
SETTING PRIORITIES
INITIAL IDEAS FOR SELECTION CRITERIA

- Importance:
  - Scholarly import
  - Widely collected
  - Formerly collected in print
  - Cost

- Vulnerability:
  - Born digital
  - Single-title publisher
  - Freely available on the web
SEEING THE TREES: A SENSE OF SCALE

- 200,000 e-seria 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
SETTING PRIORITIES: CATEGORIES

- 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
  - Bulletin du comite des forges (in Making of the Modern World)
NEXT STEPS: TESTING THE WATERS

- Validate priorities
- Continue analyzing full data set
- Work with distributors & society publishers to facilitate agreements
- Work with academic publishers & libraries on enabling strategies
- Test expansion of LOCKSS & web archiving for small publishers - freely available content
BUILDING A SUSTAINABLE APPROACH

- Model license language covering specifics (preservation methods, third-party publications, transfer titles, timing)
- Standard means of recording access rights for preserved titles
- Aggregate information about distributed preservation activities
TEST CASE: COLUMBIA AS PUBLISHER

- Ca. 50 e-journals:
  - HathiTrust (7)
  - Available through commercial aggregators (20)
  - Published by the Libraries (6)
- Other:
  - Some in our web archive
  - Some with own domain (helvidius.org)
  - Intermarium (a video journal)
DISCUSSION QUESTIONS

- What should be our criteria for setting preservation priorities?
- Any practical taxonomies for determining preservation level?
- How can we further engage our community to build awareness and encourage action?
- How do you see your organization participating?
- Any cautionary words or advice?