RLG Shared Print Update

Constance Malpas
Dennis Massie

ALA Midwinter
Denver, Colorado
26 January 2009
Overview

- RLG ‘Prospective’ Journals Preservation Project – status report
- De-duplication of print journal back files – working group update
- Network disclosure of print archiving commitments – MARC 583 proposal
Journals Preservation Project: Context

- Growing evidence that traditional print repositories are reassessing print preservation mandate
  - Ithaka surveys of faculty / librarian perceptions (2003-)
  - SCONUL survey of library leaders (2008)
- OCLC analyses suggest duplication in system-wide library holdings is relatively limited
  - Google 5 study (2006)
  - Global Resources Report (2007)
  - Unique print books in ARL libraries (2008)
- Print preservation risk models
  - C. Yano/Ithaka research report (2008)
RLG Journals Preservation Project

- **Goal**: Develop a shared methodology and operational workflows for identifying and managing at-risk serial publications *as a class*
- **Strategy**: Encourage libraries with existing collection strengths to assess print serials holdings and acquisitions with an eye to system-wide preservation requirements
- **Scope**: Active, refereed scholarly journals in the humanities with print-only distribution channels and <50 holding libraries
- **Method**: Assign sample titles for institutional review to assess condition, coverage, scholarly value and commitment to retain/acquire indefinitely
Current Status

- 230+ titles identified for inclusion based on WorldCat holdings and bibliographic data
  - 42% English language publications
  - Median holdings per title = 24
  - Median holdings per title in pilot group = 2
  - Average age of publication = 27 years

- Title list made available as WorldCat list
  - Reference-able data
  - Lives ‘in the network’

- Review template successfully deployed
  - Volume-level validation of coverage and condition
  - Institutional retention assessments
<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>OCLC #</td>
<td>&quot;Title&quot;</td>
<td>&quot;Author&quot;</td>
<td>&quot;Language&quot;</td>
<td>&quot;Publisher&quot;</td>
<td>&quot;Notes&quot;</td>
</tr>
<tr>
<td>2</td>
<td>2240614</td>
<td>Tareas.</td>
<td>Spanish</td>
<td>Paramarī</td>
<td>Arizona</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>3575530</td>
<td>Ars Hungarica.</td>
<td>Magyar Tudor Hungarian</td>
<td>Budapest,</td>
<td>Getty</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>1754022</td>
<td>Tül%riyatif mecmuas%A+</td>
<td>İ%stanbul U: Turkish</td>
<td>İ%stanbul, NYU</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>44685116</td>
<td>Socie% tel% s &amp; repre% sentations.</td>
<td>CREDEHES</td>
<td>French</td>
<td>Paris : CR</td>
<td>NYU</td>
</tr>
<tr>
<td>21</td>
<td>4948356</td>
<td>Parmassos.</td>
<td>Philologikos</td>
<td>Greek, Mo</td>
<td>Athens P</td>
<td>NYU</td>
</tr>
</tbody>
</table>
Archiving long-tail resources requires cooperation

Avrea Parma: rivista quadrimestrale di storia, letteratura e arte.

- Limited aggregate holdings
- Limited holdings within group

Began 1912...
- Getty holds recent issues only
- Arizona holds single volume
Preliminary Findings

- Approx. 40% of titles reviewed are held in their entirety by the assigned institution
  - Title-level retention claim could be made with relative ease; archiving investment is minimal, prospective
  - Pilot project could elevate 90 or more at-risk serial titles in initial sample to ‘archived’ status within months

- Approx. 30% of titles are <50% complete at assigned institution
  - Will likely require reassignment to other archiving partner; local gap-filling for these titles is last resort
  - Broadening participation base will increase number of titles susceptible to near term archiving commitment
Preliminary Findings (cont.)

- Local burden for (light-weight) validation is relatively modest, approx. 15 minutes/title
  - Low overhead makes multiple validations feasible

- Pilot participants are prepared to declare archiving commitment for selected titles
  - Global retention / access policy for locally archived titles

- Relatively low cost of titles makes ongoing acquisitions a low-risk proposition
  - @ $50/title, prospective archiving is sustainable
What’s Next?

- Scale up review activity to ‘production’ level
- Develop method for re-assignment of titles within group
- Assess cost/benefit of secondary evaluation for titles held at >50% completion rate
- Assess cost/benefit of secondary evaluation for titles held at <50% -- will likely entail outreach to libraries outside pilot group

Target: Complete review of initial 230 title sample by April/May
Compile data on direct costs of validation and continuing subscriptions
Food for Thought

10,000 peer-reviewed journals in print-only form...

- If 40% held in entirety by at least one institution, can ‘rescue’ 4000 journal titles at minimal cost and effort;

- For remaining 60%, need to assess value of retrospective archiving efforts, which are likely to be labor intensive if complete coverage is sought;

- Until more is known about aggregate demand for these resources, cooperative effort in prospective archiving provides appropriate balance of risk/value.
RLG Journals Preservation Project

- Participating Libraries
  - Getty Research Library
  - New York University
  - Penn State University
  - Stanford University
  - University of Arizona
  - UCLA
  - University of Michigan
  - University of Pennsylvania

...and the California Digital Library

Questions, Comments?
Willing to Perform Secondary Validation?
MARC 583 for Distributed Print Archiving

- 583 defined for use in disclosing local preservation action: repair, reformatting, etc

- As part of CCMT project (2007-’08), guidelines developed for use in declaring print archiving commitments as part of batch processing

- Guidelines tabled after CCMT pilot terminated, December 2008

- Revived as part of RLG Journals Preservation project
MARC 583 tag usage in WorldCat

- Approximately 400K tags, representing one million holdings
- Reflects wide range of local actions or intents
  - Microfilmed, digitized, will not conserve, selected for Web archiving, ISSN registration ...
- Used in Registry of Digital Masters to record preservation digitization actions
- Not retained in all batch loads ("local data")
583 and Cooperative Print Management

- Enables **effective network disclosure** of institutional archiving commitments

- **Raises visibility and value** of extant print archives

- Increases **incentives for participation** in distributed archiving arrangements

- Supports **coordinated management of print resources** across regional, national boundaries

- Represents **existing infrastructure that can be deployed in the short term**
583 for Monographic Print Archiving

Draft guidelines (2008) **required disclosure** of
- Level of retention commitment (dark, dim, light)
- Date of commitment

...**recommended disclosure** of
- Condition assessment (has/not been reviewed)
- Item condition
- Date of assessment

And **posited role for national libraries** as ‘dark’ archives, collections of last resort
Preliminary Consultation – Jan 12 2009

- Representatives from LoC, NAL, NLM
- CONSER liaisons from institutions participating in RLG Journals Preservation Project
- UC Shared Print Management team

Questions:
- What are the key *operational obstacles* to disclosing title-level retention commitments for print serial holdings in the MARC 583 tag of master bibliographic record?
- Feasible to apply *class-based retention commitments* to materials in your collections?
- Currently managing *title-level condition or retention information* outside of MARC bibliographic or holdings record? Cost/benefit of disclosing this data?
Additional Consultations at ALA Midwinter

- ALCTS CDS Chief Collection Development Officer
  - Potential value in cooperative print management

- Preservation Administrators Interest Group
  - Role in cooperative preservation initiatives

- CONSER At Large
  - Technical considerations
Proposed modifications (Allgood)

- 583 for print archiving action:
  - $3 Specification of copy (optional)
  - $a Level of archiving commitment (required)
    - Dark archive copy – permanent retention, surrogate access
    - National – permanent retention, on-site/surrogate access
    - Light archive – permanent retention, circulating copy
  - $c Time/Date of action (required) – date archived
  - $l Status (optional) – access restrictions if any
  - $m Dates and/or sequential designation of issues (required) – range-specific commitment
  - $u URL link (optional) – add’l documentation as needed
  - $5 Institution to which field applies (NR) (required)
  - 042 used to identify records as part of national trust

Text in red represents proposed revision to draft guidelines
Sample Record

>010 86647321
>040 DLC $c DLC
>012 3 $i 8606
>022 0262-558X
>042 lc $a ccmt
>043 e-uk---
>050 00 N6768 $b .N48
>082 00 709/.42 $2 19
>090 $b
>049 OCLC
>245 00 New Arcadians' journal.
>260 Bradford, West Yorkshire : $s
>300 v. : $b ill. ; $c 21 cm.
>500 Description based on: No. 1
>583 1 $a National copy $c 2008-01-13 $m Vol. 1, no. 1 (1981)-$2 ccmt $5 CSstoGRI
>650 0 Art, British $v 20th century $v Periodicals.
>850 0
>029 1

Identifies record as part of print archive ‘registry’

Title subject to permanent retention; on-site or surrogate access only

Effective date of archiving commitment

Terminology conforms to CCMT data dictionary

Institutional identity of archive

Range-specific archive commitment
What’s Next?

- Further consultations
  - Collection managers – How and when will data be used?
  - Technical services – Workflow Integration
  - Preservation administrators – Condition terms

- Possible MARBI proposal – use of 583$m to disclose range-specific commitments

- Draft CCMT thesaurus for submission to MARC Standards Office to establish CCMT as authentication code and source of terms.

- Coordination with OCLC product management – batch loads, WorldCat quality assurance

**Target:** Approvals in place for May 2009 MARC Update to WorldCat
Testbed data from RLG Journals Project and external sources
Special Thanks...

*Everett Allgood, NYU
Diane Boehr, NLM
Robert Bremer, OCLC
Margaret Byrnes, NLM
Laura Hartmann, NLM
Chris Cole, NAL
Leighann Ayers, Michigan
Valerie Bross, UCLA
Jeanne Drewes, LoC

Ann Fath, Getty
*Les Hawkins, LoC
Dianne McCutcheon, NLM
Jake Nadal, UCLA
Glenn Patton, OCLC
*Christopher Walker, Penn State
Tina Shrader, NAL
Shana McDanold, Penn
Emily Stambaugh, CDL

Questions, Comments?
Technical Expertise to Contribute?
Questions

- Which sources of extant data will deliver greatest impact?
  - JSTOR archives (dim, dark?) – MSU, CRL, FCLD, etc.
  - UK Research Reserve Retention Register
  - Regional last/single copy archives – VALE, PALCI, etc.

- Is there value in recording digital archiving commitments?
  - Portico
  - LOCKKS
  - HATHI
Stay Informed . . .

- RLG Future of Collections email discussion list
- RLG Shared Print Facebook Group
- OCLC Research Web site
  www.oclc.org/programs/ourwork/collectivecoll/sharedprint
- RLG Journals Preservation Project sample title list
  www.worldcat.org/profiles/cmalpas/lists/346407

Constance Malpas
malpasc@oclc.org

Dennis Massie
massied@oclc.org