Increasing Access to Special Collections

Ricky Erway, OCLC Research
LIBER Annual Conference, Barcelona
29 June 2011
OCLC Research: Mobilizing Unique Materials

- Sharing special collections
- Rapid Capture
- Capture and Release
- Scan and Deliver
- Single Search
- Born Digital
- Metadata is the Interface
- Good Terms
- Well-Intentioned Practice
- Taking Our Pulse
- Shifting Gears

Zeets Jones. Light at the End of the Tunnel. 2007. (CC-BY-NC-ND)
Shifting Gears: Gearing Up to Get into the Flow

Ricky Erway and Jennifer Schaffner

Program officers
OCLC Programs and Research

A publication of OCLC Programs and Research
Shifting Gears - Recommendations

- It’s about access
- It’s about quantity
- Selection has already been done
- Iterate
- Programs not projects
- Partnerships
- Adopt archival description practices
- Discovery happens elsewhere
Selecting for digitization

If we could we’d digitize everything en masse

Look for high-demand collections or series

Compromise by digitizing samples to gauge the interest

Sometimes we have to do selections

(Is it worth it?)

Put your efforts into the larger context

Can you get the best of all worlds?
Taking Our Pulse: The OCLC Research Survey of Special Collections and Archives

Taking Our Pulse:
The OCLC Research Survey of Special Collections and Archives

Jackie M. Dooley
Program Officer

Katherine Luce
Research Intern

OCLC Research

OCLC
A publication of OCLC Research
Taking Our Pulse - Key Findings

• Many rare and unique material are undiscoverable
• Tough economy renders “business as usual” impossible
• Staff are having trouble keeping up

• The top three “most challenging issues”
  • Space
  • Born Digital
  • Digitization
Taking Our Pulse - Recommendations

- **User Services**
  - Promote policies that facilitate (rather than inhibit) access

- **Born-Digital Archival Materials**
  - Identify practices of archivists useful for born digital
  - Basic steps for managing born digital

- **Digitization**
  - Scope the digitized rare book corpus
  - Develop models for large scale digitization
Well-intentioned practice for putting digitized collections of unpublished materials online

The primary responsibilities of cultural materials repositories—stewardship and support for research and learning—require us to provide access to materials entrusted to our care. This document establishes a reasonable foundation of practices that increase and significantly improve access to collections of unpublished materials by placing them online for the purpose of furthering research and learning. Although it promotes a well-intentioned, practical approach to identifying and resolving rights issues that is in line with professional and ethical standards, note that this document does not concern itself with what individuals may do with the works. While the document was developed with U.S. law in mind, it is intended to provide a general approach that will resonate in non-U.S. contexts.

If your institution has legal counsel, involve them in adopting this approach; after the approach has been adopted, only seek their advice on specific questions.

Select collections wisely:
- Keep your mission in mind and start with a collection of high research value or high user interest.
- Assess the advantages and risks of relying on for use in the U.S. for public access.
- Some types of materials may warrant extra caution when considering rights issues, such as:
  - Contemporary literary works
  - Collections with sensitive information, such as social security numbers or medical data
  - Materials that are likely to have been created with commercial intent (because they are more likely to have economic value)
  - Very recent materials that were not intended to be made public.
- If research value is high and risk is high, consider compromises, such as making a sensitive series accessible on-site only, if a suitable time has passed.

Use archival approaches to make decisions:
- Get pre-approval and co-curation records for permissions, rights, or restrictions.
- Assess rights and privacy issues at the appropriate level, even when the collection is online.
- Attempt to contact and get permission from the rights holder. If there's an identifiable rights holder at the level.
- Include what you know about the rights status in the description of the collection, including if the collection is in the public domain, if the institution holds the rights, or if the rights holder has given the institution permission to place the digitized collection online.
- Document your processes, findings, and decisions and share them with your profession community.

Provide clear and written policy statements and declarations to users of online collections:
- Adhere to a liberal take-down policy, such as: “These digitized collections are accessible for purposes of education and research. We've indicated what we know about copyright and rights of privacy, publicity, or trademark. Due to the nature of archival collections, we are not always able to identify this information. We are expert to keep from any rights owners, so that we may obtain accurate information. Upon request, we will remove material from public view if we receive a rights issue.
- Use an appropriate Disclaimer at the institutional level, such as: [institution] makes digital versions of collections accessible in the following situations:
  - They are in the public domain
  - The rights are owned by the institution
  - [Institution] has permission to make them accessible
  - We make them accessible for education and research purposes
  - There are no known restrictions on use
- To learn what your responsibilities are if you choose to use the materials, go to: [link]

Preventively, work with donors:
- Identify possible intellectual property issues and get relevant contact information.
- Ask donors to state any privacy concerns and identify sensitive materials that may be in the collection.
- Suggest that donors transfer copyright to the institution or license their work under a Creative Commons CCO License.
- Include statements in your collection policies and in your deeds of gift or transfer documents that:
  - ensure that no restrictions are placed on content that is already in the public domain;
  - guarantee to disclaim the materials for unrestricted access when donors retain the rights;
  - and guard against limitations or restrictions on future rights.

For more about this document, see: [link] Revised 05/21/2010
Well-Intentioned Practice - Main Points

- Select collections wisely
- Use archival approaches to make decisions
- Provide take-down policy statements and disclaimers
- Prospectively, work with donors
Good Terms: Improving Commercial-Noncommercial Partnerships for Mass Digitization

D-Lib Magazine
November/December 2007
Volume 13 Number 11/12
ISSN 1082-9873

Good Terms - Improving Commercial-Noncommercial Partnerships for Mass Digitization

A Report Prepared by Intelligent Television
for RLG Programs, OCLC Programs and Research

Peter S. Kaufman and Jeff DRose
Intelligent Television

Executive Summary

In 2007, OCLC Programs and Research engaged Intelligent Television to study the partnership agreements between cultural institutions and for-profit companies for the mass digitization of books and other media. This report presents the findings of that study.

Libraries have been digitizing portions of their collections for more than twenty years, but recent opportunities to work with private partners, such as Google, Microsoft, and others, on mass digitization has opened up possibilities that were unimaginable just a few years ago. Private funding, commercially developed technology, and market-oriented solutions together may generate larger aggregations of digitized books far sooner than the library community had dreamed possible. There are many efforts underway to assess various aspects of these partnerships; this paper focuses on the terms in mass digitization agreements that affect research-community-centered outcomes.

The libraries and other cultural institutions that private companies first approached saw significant potential in these overtures; they were elated in seeing that near-term local needs were met. Only when it became clear that a significant number of these partnerships were underway did the library community as a whole begin to think about the overall impact of these business relationships on the future of scholarship.

When we fantasize about that future, we imagine a single way to search all digitized books, journals, and other media, a combined index of all the full texts that will enable research that is otherwise
Good Terms - Recommendations

Negotiate for

- Limited confidentiality
- Complete deliverables
- More open access allowing the development of new applications
- Less restricted distribution
- Responsible treatment of usage data
  and
- Limited duration and survivability
Rapid Capture: Faster Throughput in Digitization of Special Collections

Rapid Capture:
Faster Throughput in Digitization of Special Collections

Ricky Erway

Senior Program Officer
OCLC Research

OCLC
A publication of OCLC Research
University of California, Berkeley and University of the Pacific
Archives of Traditional Music, Indiana University
Minneapolis, Jan. 7/44

One of the lightest hockey schedules ever drawn by a University of Minnesota team will get under way January 7-8 when the Gophers face the fast-treading, experienced St. James Club of Winnipeg at the Minneapolis Arena.

At present the St. James series is the only set of games definitely scheduled for Coach Harvey Armstrong and his pucksters. Dates and arrangements have not been completed as yet for a meeting with the Canadian Navy team of Port Arthur, however Coach Armstrong feels confident that the dates of January 14-15 will be approved by Canadian immigration authorities and military officials which permit the passage of Canadian teams across the border, thus enabling the Port Arthur team to keep its appointment in Minneapolis for their encounters.

Minnesota's old rival in the ice sport, the University of Michigan, may cancel its game set for the area the later part of January due to manpower problems facing the Wolverines this year. Most of the Michigan puck candidates are naval reserves and as such are under the forty-eight hour rule which forbids navy men to be away from their bases in excess of the allotted time. 

On word received from Michigan athletic officials they claim it may prove impossible for them to send a team on the road for a two day stand and still abide by navy orders.

All this, however, has not discouraged Coach Armstrong for he has his squad going through daily drills for the past seven weeks in hope that additional games will be added to the schedule in the near future.

This year's schedule is quite a contrast to the fifteen game campaign of last year in which Minnesota won ten, lost four and tied one, winning
Sharing Special Collections

Advisory group made up of special collections and ILL staff

• Streamlining workflows for loan requests
• Model policy for lending special collections

Capture and Release

“Capture and Release”: Digital Cameras in the Reading Room

Lisa Miller, Steven K. Galbraith, and the RLG Partnership Working Group on Streamlining Photography and Scanning

Helti. mladica / hucho hucho - catch and release. 2008. (CC-BY-NC-ND)
Capture and Release - Benefits

- Digital cameras are gentler on collection materials
- Digital cameras increase researcher satisfaction
- Digital cameras reduce repository workload
- Digital cameras enhance security and save reading room checkout time
- Digital cameras save paper and photocopy toner
- Digital cameras reduce liability for copyright infringement
- Repositories stay current and resolve an ongoing issue
- Digital cameras facilitate use

<table>
<thead>
<tr>
<th>Facet</th>
<th>Shutter-bug</th>
<th>Exposed</th>
<th>Camera-shy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional photocopying</td>
<td>Self-service</td>
<td>Self-service after staff review</td>
<td>All copying done by staff</td>
</tr>
<tr>
<td>(possible baseline for digital camera policy)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equipment</td>
<td>No flash, no lights</td>
<td>Patron’s camera limits on supporting equipment (copy stands, tripods, cords, etc.)—some pieces allowed, others not</td>
<td>Repository’s camera only (and possibly other equipment supplied by repository)</td>
</tr>
<tr>
<td></td>
<td>Allow flatbed scanners</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Allow and/or provide copy stand, tripod, extension cords, stepstool, etc.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Repository supplies camera or self-service overhead book scanner in addition to allowing patron’s camera</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Photography space</td>
<td>In reading room at any station</td>
<td>In reading room at designated stations, usually close to reference desk</td>
<td>Separate room</td>
</tr>
<tr>
<td>Photography rules</td>
<td>No standing on tables or chairs</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>No rearrangement of furniture</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>No materials on floor</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Remain behind table, facing forward at all times</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Set camera to “mute”</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Do not disturb others</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>No photographs of reading room, staff, or patrons</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appointments</td>
<td>Appointment not required</td>
<td>Appointment made during visit</td>
<td>Appointment made in writing in advance</td>
</tr>
<tr>
<td>Staff review of collection materials</td>
<td>Part of standard staff surveillance of patrons in reading room</td>
<td>Patron must verbally notify reference attendant each time camera is used and show attendant the materials being photographed</td>
<td>Patron must formally indicate and curatorial staff formally review all materials</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Camera stays in locker until approval is given</td>
<td>Same day approval may be given</td>
</tr>
</tbody>
</table>


Scan and Deliver: Managing User-Initiated Digitization in Special Collections and Archives

Scan and Deliver:
Managing User-initiated Digitization in Special Collections and Archives

Jennifer Schaffner
OCLC Research

Francine Snyder
Solomon R. Guggenheim Museum

Shannon Supple
University of California at Berkeley

OCLC
Scan and Deliver - tiered workflow

A. Review
- How do users request digital reproductions?
- How are requests for copies approved?
- How many people need to approve?
- Have the items already been digitized?
- How do you manage copyright & legal issues?

B. Decide
- Will you keep the images?
- Is it useful for you to keep the images?
- Do you have sufficient resources to keep them?
- What resolution is needed for scanning?
- Is the request for high-resolution?
- What metadata will you create?
- Will you scan the whole volume?

C. Scan
- Who digitizes and describes materials?
- Will you perform quality control?

D. Deliver
- How do you deliver digital copies to the user?

<table>
<thead>
<tr>
<th>REVIEW</th>
<th>INSIDE TRACK</th>
<th>MIDDLE TRACK</th>
<th>OUTSIDE TRACK</th>
</tr>
</thead>
<tbody>
<tr>
<td>REQUEST</td>
<td>Verbal request</td>
<td>Simple form</td>
<td>Detailed forms</td>
</tr>
<tr>
<td>APPROVE</td>
<td>One staff member</td>
<td>Two staff</td>
<td>Three or more staff</td>
</tr>
<tr>
<td>SEARCH FOR EXISTING SURROGATE</td>
<td>Don’t take the time</td>
<td>Quick check</td>
<td>Search: Google Books, EEBO &amp; ECCO, HathiTrust, Etc…</td>
</tr>
<tr>
<td>RIGHTS</td>
<td>User’s responsibility</td>
<td>User’s responsibility</td>
<td>Search, control, &amp; monitor thoroughly</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DECIDE</th>
<th>INSIDE TRACK</th>
<th>MIDDLE TRACK</th>
<th>OUTSIDE TRACK</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAFF</td>
<td>One staff member</td>
<td>Two staff</td>
<td>Multiple specialists: Curator, Conservator, Metadata creator, Image professional, Etc…</td>
</tr>
<tr>
<td>RESOLUTION</td>
<td>Easiest option</td>
<td>Middling to high quality</td>
<td>High quality</td>
</tr>
<tr>
<td>METADATA</td>
<td>None</td>
<td>Filename and keyword or two</td>
<td>Metadata &amp; citation (choose standards)</td>
</tr>
<tr>
<td>WHOLE / PART / ITEM</td>
<td>Will scan whole unit if it takes up to “x” minutes</td>
<td>Entire chapter or series</td>
<td>Entire volume or collection (large-scale digitization)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SCAN</th>
<th>INSIDE TRACK</th>
<th>MIDDLE TRACK</th>
<th>OUTSIDE TRACK</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAFF</td>
<td>One staff member</td>
<td>Two staff</td>
<td>Multiple staff</td>
</tr>
<tr>
<td>QUALITY CONTROL</td>
<td>Trust your staff!</td>
<td>Quick review</td>
<td>Thorough review by multiple specialists (color-correction, etc.)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DELIVER</th>
<th>INSIDE TRACK</th>
<th>MIDDLE TRACK</th>
<th>OUTSIDE TRACK</th>
</tr>
</thead>
<tbody>
<tr>
<td>DELIVERY METHOD</td>
<td>Deliver forthwith to user</td>
<td>Keep in local files and Web site</td>
<td>Digital library, Digital repository, DAMS, Consortia, Aggregations</td>
</tr>
</tbody>
</table>

Figure 1. Tiered Workflow for User-Initiated Digitization
Single Search

The Quest for the Holy Grail:
Single Search across an Institution’s Collections

Leah Prescott
Digital Projects Coordinator
Getty Research Institute

Ricky Erway
Senior Program Officer
OCLC Research

A publication of OCLC Research
Single Search - Implementers Share Experiences

Victoria & Albert Museum
Yale Center for British Art
Yale University
University of Calgary
Minnesota Historical Society
The Magnes Collection of Jewish Art and Life
Getty Research Institute
Wellcome Trust
Smithsonian Institution

Motivation
Organizational structure
Technical approaches
Crosswalks

Gathering data
Vocabularies
Managing digital assets
Access
Born Digital

Definition

Archival practices

First steps
Types of Born-Digital Materials

- Digital photographs
- Digital documents
- Harvested Web content
- Digital manuscripts
- Electronic records

- Static data sets
- Dynamic data
- Digital art
- Digital media publications
The Metadata is the Interface: Better Description for Better Discovery

The Metadata is the Interface
Better Description for Better Discovery of Archives and Special Collections,
Synthesized from User Studies

Jennifer Schaffner

Program Officer
OCLC Research

A publication of OCLC Research
Getting the most out of metadata

• Reduce the amount of effort we put into metadata, while improving discovery
• Do high-level description
• Iterate
• Describe the “aboutness”
• Encourage tagging
• Involve researchers
• Be where the users are
There is light at the end of the tunnel...

Eric Jones. Lon Peris cycle track seen from the other end of the tunnel. 2007. (CC-BY-SA)
Links and contact info

• Well-Intentioned Practice http://www.oclc.org/research/activities/rights/practice.pdf
• Good Terms  http://dlib.org/dlib/november07/kaufman/11kaufman.html
• Sharing Special Collections http://www.oclc.org/research/activities/sharing/default.htm
• Born Digital Special Collections http://www.oclc.org/research/activities/borndigital

oclc.org/research/publications
erwayr@oclc.org