Managing a Digitization Project: Arcadia 4, Old Yale Persian

Nehru Memorial Museum and Library
New Delhi, India
6 November 2013

Sarah Calhoun
Librarian for South Asian Studies
Yale University Library
Yale’s South Asia Collection

- About 100,000 titles cataloged online
- 22 Yale Library locations
- Languages:
  - 75% English (75,000 titles)
  - 7% Hindi (7,000 titles)
  - 3% each of Sanskrit, Tamil, Tibetan
  - About 1% each of Urdu, Bengali, Arabic, Persian, Gujarati, Pali, and Prakrit
- Two call number ranges
Call number ranges

- Yale Classification System ("Old Yale")
  - Ca. 1700s – ca. 1950
  - Similar to Library of Congress class, but unique to Yale
  - Not all holdings fully updated in Yale’s OPAC
  - Holdings don’t show properly in OCLC/WorldCat

- Library of Congress Classification System
  - Ca. 1950 – present
  - National standard for large research libraries
  - Compare collections via OCLC/WorldCat
Technology has disrupted my job

- Service model vs collections model
- Facilitate easy and seamless digital access to previously hidden collections
- Lower barriers to entry for faculty/students
- There is too much technology for me to know – I need to work with (and trust) my colleagues in other departments to guide my decisions about collections
- Yet again: Best practices are a moving target
My questions after assessment

- What do we have that is currently unfindable or hidden from our patrons?
- How can I make our existing collection easier to explore and use?
- How can I empower our patrons to find things on their own, without needing me as an intermediary?
Opportunities for Growth

Conclusion:

Work needs to be done!
“Every book its reader”

- No way to discretely describe “the collection”
- Many “difficult languages” are under-cataloged (i.e. Persian, Bengali, Assamese)
- Pali and Prakrit books often classed under the writing system, and not the language
- Many records not in OCLC, and using out-dated bibliographic standards
Conceiving of a project

- Cataloging [inventory control]
- Conservation [“relic”]
- Digitization [access]

- Open up one hidden collection of Persian material from the Old Yale Class, instead of adding new accession
- Succeed small instead of fail big
  - Find a representative group of items
  - Check to make sure workflow is sound
  - If sample batch works, then expand scope
- Collection website not live, but info elsewhere...
Steps for managing a digital project

1. Select items, write proposal, consult stakeholders
2. Find funding
3. Team meeting
4. Cataloging
5. Conservation (1)
6. Digitization
7. QC and Ingest
8. Conservation (2)
9. Outreach
Step 1: Selection, South Asia Librarian
Step 1: Selection

- I first assessed the state of the old Yale Class, Persian philology section (a “difficult” area)
- Responsibility falls between South Asia and Near East collections.
- Includes religious texts, literary works, translations, and grammars.
- Small, closed collection.
Step 1: Selection, South Asia Librarian

We did a "shelf read" to compare 200 books on the shelf to our catalog.

About 30% of the Persian philology books were undercataloged and crumbling.
In the end, I selected 20 Persian philology books for this project. All of them were...
Step 1: Selection, South Asia Librarian

- Out of copyright
- In poor condition
- Useful for faculty and students
- Rare
- Needed cataloging
We took a conservative look at copyright.

There were some books that we couldn’t select because of copyright ambiguity.

**Title:** “Mir Ghulam Ali Mail Tattavi”  
**Author:** Kullyat-i-Mail  
**Published:** Karachi [Pakistan], 1959  
**Language:** Persian  
**Call number:** Frh57 M282 K8 1959
Step 2: Finding Funding

YALE LIBRARY: INNOVATIVE SOLUTIONS FOR LIBRARIES WORLDWIDE

June 2013

The team has created innovative cataloguing and software solutions applicable to libraries worldwide. For example, cataloguing Yale’s African holdings resulted in transcriptions of several non-Latin alphabet African languages. These have since been approved by the Library of Congress.

Many of the collections made accessible through the project have a unique historical value. A good example is a previously unknown archival collection of the Jewish community of North Africa.

The documents are written in Judeo-Arabic and Hakifya, a merging of Hebrew and Spanish.

The materials have now been catalogued and described by Professor Mohe Bar Asher of the Hebrew University.

Professor Bar Asher’s father worked as a scribe in Morocco and wrote several of the manuscripts in Yale’s collection.
Step 2: Finding Funding

I received a library-managed grant to catalog, conserve, and digitize these 20 books.
This meeting allowed us to clarify the sequence of events.
It also answered questions about each others’ workflows.
Step 4: Cataloging

Abdelahad Hannawi
Near East Cataloger

Patricia Thurston
Head of Specialty Cataloging
Step 4: Cataloging

- Catalog records were completed, with original script when appropriate.
- Call numbers were updated to the Library of Congress classification system.
Sample updated record (Orbis)

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform Title:</td>
<td>[Poems. 1863]</td>
</tr>
<tr>
<td>Title:</td>
<td>Hāzā dīwān-i Mawlānā 'Abd al-Rahmān. هذَا دیوان مولانا عبد الرحمن.</td>
</tr>
<tr>
<td>Published/Created:</td>
<td>[Bombay: s.n.], 1280 [1863]</td>
</tr>
<tr>
<td>Physical Description:</td>
<td>1 v. (unpaged); 25 cm.</td>
</tr>
</tbody>
</table>

**Yale Holdings**

<table>
<thead>
<tr>
<th>Location:</th>
<th>SML, Near East Collection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Call Number:</td>
<td>PK6818 R3 A17 1863</td>
</tr>
<tr>
<td>Status:</td>
<td>- Unavailable. Use Borrow Direct or Interlibrary Loan.</td>
</tr>
</tbody>
</table>

**Notes:**
- Lithographed.
- Poems.

**Variant and related titles:**
- Dīwān-i Mawlānā 'Abd al-Rahmān
- Dīvān-i Mawlānā 'Abd al-Rahmān
- دیوان مولانا عبد الرحمن

**Format:** Book

Link to this page: [http://hdl.handle.net/10079/bibid/7653827](http://hdl.handle.net/10079/bibid/7653827)

1. Original script
2. Full level cataloging with authority control
3. LC Call number
Sample updated record (Orbis)

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform Title:</td>
<td>[Poems. 1863]</td>
</tr>
<tr>
<td>Title:</td>
<td>Hāzā diwān-i Mawlānā ‘Abd al-Rahmān. هذا ديوان مولانا عبد الرحمن.</td>
</tr>
<tr>
<td>Published/Created:</td>
<td>[Bombay : s.n]. 1280 [1863]</td>
</tr>
<tr>
<td>Physical Description:</td>
<td>1 v. (unpaged) ; 25 cm.</td>
</tr>
</tbody>
</table>

Yale Holdings

<table>
<thead>
<tr>
<th>Location:</th>
<th>SML, Near East Collection Where is this? Request recall or delivery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Call Number:</td>
<td>PK6818 R3 A17 1863</td>
</tr>
<tr>
<td>Status:</td>
<td>- Unavailable. Use Borrow Direct or Interlibrary Loan.</td>
</tr>
</tbody>
</table>

Notes: Lithographed. Poems.

Variant and related titles: Divān-i Mawlānā ‘Abd al-Rahmān Divān-i Mawlānā ‘Abd al-Rahmān ديوان مولانا عبد الرحمن

Format: Book

Link to this page: http://hdl.handle.net/10079/bibid/7653827

1. Original script
2. Full level cataloging
3. LC Call number
4. …but still some ambiguity!
Sample updated record (OCLC)

1. Original script
2. Full level cataloging
3. LC Call number

Plus:
1. Yale’s holdings are displaying
2. Corrected language (Pushto)
Step 5: Conservation, pre-flight
Step 5: Conservation, pre-flight

Conservation staff stabilized books for transport to digitization vendor.

- Boxes
- Disbinding
- Mylar sleeves for individual pages
- Washed some materials to remove tape or other previous repairs

Treated to facilitate safe handling during the digitization process.
Step 5: Conservation, pre-flight

Each page placed in its own individual mylar sleeve (brittle paper)
Step 6: Digitization

Robert Klingenerger
Head of Digital & Reformatting Services
Step 6: Digitization

- **Outsourced to vendor**
  - 19 books were from our normal circulating collection.
  - Sent to external vendor with conservator on staff.
  - Some items disbound.

- **Digitized in-house**
  - One book from Beinecke Rare Book Library.
  - Digitized in their own studio.
  - Not disbound.
Step 7: Digital Quality Control (QC) and Ingest

Robert Klingenberger
Head of Digital & Reformatting Services

Sarah Calhoun
South Asia Librarian

Mike Friscia
Manager of Digital Library & Programming Services
Step 7: Quality Control and Ingest

- We receive files from the vendor on a hard drive
- We inspect the files:
  - Checksum to make sure all bytes there
  - Check RTL/LTR display
  - All pages scanned?
  - All scans clean?
- The files will be ingested into online Fedora repository
Step 7: Quality Control and Ingest

- **Item list:**
  https://www.zotero.org/groups/arcadia4-oldyalepersian

- **More information:**
  http://guides.library.yale.edu/SouthAsiaPersian
Once the books are digitized, they return to Conservation before reshelving.
Step 8: Conservation (again)

When the books return from being digitized, they need to be put back together. This ensures that the object will continue to be usable as far into the future as possible.

- Rebinding (resewing text blocks)
- Reconstructing covers
- Completing final conservation treatment
Step 9: Outreach and Advertisement

http://guides.library.yale.edu/SouthAsiaPersian
My job:
Facilitate easy and seamless digital access to previously hidden collections

Inform local faculty and students
- What the collection is
- How to access it

Create online research guide

Inform non-local scholars

Post on relevant email listservs

Professional societies, like CONSALD, MELA, CRL/SAMP
Twenty Persian philology books, formerly “hidden” in the old Yale class, will now be available to the world. Our patrons will be able to browse the collection online.
Lessons for the future

- Succeed small instead of fail big
- Slow and steady
- Leave enough time for digital QC
- Vet external vendors thoroughly
- Communicate regularly with team
All Yale presentations available online:
http://guides.library.yale.edu/NMMLYale
Yale UNIVERSITY LIBRARY