Non-Roman & Ideographic Cataloging:
Cataloging Examples

Lisa Robinson, Michigan State University
How to Catalog What You Can’t Read?

- Catalog in-house
  - Need language expertise & computer set-up
- Outsourced cataloging
  - Need a budget, but get good quality records
- Record loads
  - You get what you get!
- Use copy
  - How to search & match titles?

MSU is following all these methods to get non-Roman & ideographic materials cataloged.

We hired a full-time Chinese cataloger due to the volume of materials acquired in Chinese. She also adds Chinese language information (880 fields) to existing records. But there needs to be a sufficient volume of work to justify hiring a cataloger.

We outsourced cataloging of 48 Hebrew, Yiddish, & Arabic books to Backstage Works. Very pleased with the results. We had no budget for outsourced cataloging, but the donor of the materials also gave some money to pay for some cataloging. We spend $1100 for 48 titles (40 monographs, 10 serials) = $23/title. Records included non-Roman scripts in 880 fields.

We load records for non-Roman materials from 2 sources: MARCNow for Russian and Hathi Trust. MARCNow provides good records and we have 2 catalogers who can read Russian to match the records to books & make edits as needed. HathiTrust load includes Arabic records, which have a variety of coding problems.
I searched Hebrew books using just the ISBN and “matched” copy using pagination & size. Deborah then verified that the title, edition, publisher, pub year matched.

I also found copy for 30 Arabic books searching on ISBN and “matching” on extent, size, and partially transliterated title. But now that Deborah is learning Arabic, I can ask her to verify that the title matches the record. ;->

We hired a student with Hebrew language skills to search SkyRiver & WorldCat for copy. A cataloger chose the best record from the copy found by the student. Deborah resolved questions about edition vs. printings, etc.
In the Catalog

- Non-Roman scripts and diacritics can present problems in the catalog
  - Display of diacritics
  - Coding of characters
  - Linked fields invalid, missing
  - Display of right-to-left scripts
This record is from a Hathi Trust record load. This problem doesn’t happen in all records with underdot characters. We have a call into III.
Some diacritic marks, such as the alif & ayn, have been put in records with codes that are not recognized by the system. They show up in Sierra and the public catalog as a series of letters & numbers.

These 2 codes are for the alif & ayn respectively.
You can find these records by doing a keyword search on “U+02be” and on “U+1e9a”

There were 3000 of the first error and 956 of the second error in our catalog as of Sunday!
The problem was the alif in **Mavā‘id, Qur’ān** and **Bamba’ī**

I fixed this problem by pasting in the correct word from another library’s public catalog. We would like to fix all of these with a global update, but have not (yet?) been able to create a review file in Create Lists.

The recent enhancement suggestion to be able to create a review file from a search done in Sierra/Millennium would be helpful.
Coding of Diacritics: Public Display Fixed

\[
\begin{tabular}{|l|l|}
\hline
Publisher & Bambalī, 1967. \\
Citation & No citation available. See Citation Help. \\
\hline
\end{tabular}
\]

**Available online:**
- Connect to Hathitrust online resource - All users

**LOCATION**
- HATHETRUST COLLECTION
- Online
- ELECTRONIC

**Description**
- 368 p. 28 cm.

**Note**
- Record based on print version record.

**System Req**
- Available via the World Wide Web

**Uniform Title**
- Qur'an
This IS NOT valid MARC! But haven’t tried to fix them since they displays fine in public catalog and do not appear to be a problem for indexing. But I need someone with language expertise to really test whether indexing is working.
Any mistake in the coding of 880 fields means they don’t display. But many catalogers have trouble checking them.

We used to strip out 880 fields from copy. Now our Chinese cataloger is adding them to Chinese language records.

The HathiTrust Arabic records all come from Princeton University Arabic Collection. Princeton has enhanced its local records with Arabic script, but not the necessarily the OCLC master records. We could still add the Arabic 880 fields by copying from the Princeton public catalog. But I have very few staff who are willing to work with non-Roman scripts.
This is a typical record with multiple 880 fields linked to variable fields.
This is the public catalog view of the previous record. I have added the yellow highlighting to show where the Arabic & Hebrew scripts display.
Sometimes table of contents 880 fields show up only in the vernacular. In our catalog, the 880 field doesn’t display until it is linked to a variable field.

In this example, I added a 505 note “Table of Content only in Hebrew” by hand. But we’re thinking about a project to add a 505 note to such records in an automated way. 505 _ _ Table of Contents only in vernacular.
In SkyRiver, the linked 880 fields are shown by a bracket, rather than the raw coding. This is a cleaner display, but copy catalogers don’t notice when the link is broken.
We have not established consistent practice for the display of right-to-left scripts in the public catalog. The script is always written right to left, but the text block can be displayed either on the left or right hand side of the screen.

This placement is controlled by a code in the 880 field: /r. We get records both with and without this code. No consensus in public services about which placement is best.
Thank You!