



## **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.

## Searching for Copy You Can't Read

- **Search the info you have**
  - ISBN, call number, partially transliterated title
  - May need assistance to verify that copy matches book
- **Hire students with language skills to search in non-Roman scripts**
  - Catalogers still needed to evaluate copy



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

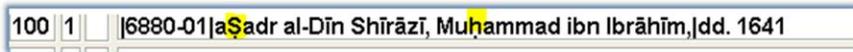
- **Non-Roman scripts and diacritics can present problems in the catalog**
  - Display of diacritics
  - Coding of characters
  - Linked fields invalid, missing
  - Display of right-to-left scripts

## Display of Diacritics

- Underdot displayed as blank box in Sierra browse results



- But OK in record editor & public view



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.

# Coding of Diacritics: Problem

The screenshot shows the Sierra library catalog interface. At the top, the title "Coding of Diacritics: Problem" is displayed. Below it, the Sierra logo and navigation menu are visible. The main content area shows a record for a book with the following details:

- Record ID: d8125703x
- LANG: ara Arabic
- CAT DATE: 12-09-2010
- SUPPRESSION: NONE
- SKIP: 0
- BIB LVL: m MONOGRAPH
- COUNTRY: India
- LOCATION: ht HATHITRUST COLLECTION
- MAT TYPE: m ELECTRONIC RESOURCE

The MARC Leader is: ##### n a m a 2 2 ##### 1 4 5 0 0

The record contains several fields with call numbers and titles. Two blue callout boxes highlight specific diacritic codes:

- A callout box labeled "\U+02be\" points to the code "U+02be" in the title field.
- A callout box labeled "\U+1e9a\" points to the code "U+1e9a" in the title field.

The title field contains the following text: "Taj al-tarajim, h[electronic resource] b mawsum bi MawāU+02beid al-raḥmān fi tarjumat al-QurU+02beān [Tarjumah-i, Muḥammad ibn Ḥusayn mulaqqab bi-Jamal al-Dīn, wa marʿuf bi-Aqā Jamāl, wa mawḥf bi Muḥaqiq-i Khūnsārī, hasab al-farmān-i Shāhanshāh-i Irān, Nādir Shāh-i Afshār. BambU+1e9a]c1867.

The record also includes fields for the number of pages (368 p.), the format (p.c28 cm.), and the availability (Available via the World Wide Web).

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.

## Coding of Diacritics: Public Display Problem

- Bad code displays to public
- Bad code affects searching

RECORD # b8125703 Entire Catalog Search  
 Limit search to available items

**Title** Tāj al-tarājim, [electronic resource] mawsum bi Mawā'id al-rahmān fi tarjumat al-Quṭub al-**U+02be**ār [Tarjumah-i, Muḥammad ibn Ḥusayn mūlaqqab bi-Jamāl al-Dīn, va ma'rūf bi-Aqā Jamāl, va mawṣūf bi Muḥaqqiq-i Khūnsāri, hasab al-farmān-i Shāhanshāh-i Irān, Nādir Shāh-i Afshār.]

**Publisher** Bambā**U+1e9a**ī, 1867.

**Citation** No citation available. See [Citation Help](#).

**Available online:**  
[Connect to HathiTrust online resource - All users](#)

LOCATION	CALL # (Click to browse similar items)	STATUS
HATHITRUST COLLECTION	Online	ELECTRONIC

IUG 2013 San Francisco

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!

## Coding of Diacritics: Fixed

Sierra - MSU Libraries - Staff: Lisa Robinson - b8125703x  
FILE Edit View Go Tools Admin Reports Help  
sierra FUNCTION Catalog  
RECORD # b8125703x Search Local  
Back to Browser  
b8125703x Last Updated: 04-19-2013 Created: 12-09-2010 Revisions: 5  
LANG ara Arabic CAT DATE 12-09-2010 SUPPRESSION NONE  
SKIP 0 BIB LVL m MONOGRAPH COUNTRY India  
LOCATION ht HATHTRUST COLLECTION MAT TYPE m ELECTRONIC RESOURCE  
MARC Leader ##### n a m a 2 2 ##### 1 4 5 0 0  
o 001 | hth008435122  
y 005 | m d  
y 007 | c r b n -- a u a u a  
y 008 | 010829 s 1867 | | s 0 0 0 0 ara d  
y 040 | ZCUjzZCU  
y 049 | QEMP  
u 130 0 | Qur'an  
t 245 1 0 | Tāj al-tarājīm [electronic resource] b mawsum bi Mawā'id al-ra'mān fi tarjumat al-Qur'ān i/c  
| [Tarjumah-i, Mu'ammad ibn 'usayn mulaqqab bi-Jamāl al-Dīn, va ma'yūf bi-Aqā Jamāl, va maw'ūf bi Mu'aqqiq-i Khūnsārī, hasab al-farmān-i Shāhanshāh-i Irān, Nādir  
| Shāh-i Afshār.  
p 260 | Bamba'i, c1867.  
r 300 | 368 p.; 28 cm.  
n 500 | Record based on print version record.  
n 538 | Available via the World Wide Web  
y 856 4 0 | uhttp://catalog.hathitrust.org/Record/008435122 Connect to HathiTrust online resource - All users  
Edit Mode (OVR)

The problem was the alif in **Mavā'id**, **Qur'an** and **Bamba'i**

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

RECORD # b8125703  Limit

**Mavā'id** **Qur'an**

**Title** Tāj al-tarājim, [electronic resource] mawsum bi Mavā'id al-raḥmān fi tarjumat al-Qur'ān / [Tarjumah-i, Muḥammad ibn Husayn mullaqqab bi-Jamāl al-Dīn, va ma'rūf bi-Aqā Jamāl, va mawṣūf bi Muḥaqqiq-i Khūnsāri, hasab al-farmān-i Shāhanshāh-i Irān, Nādir Shāh-i Afshār.

**Publisher** Bamba'i, 1867.

**Citation** No citation available. See [Citation Help](#).

**Available online:**  
[Connect to HathiTrust online resource - All users](#)

**Bamba'i**

LOCATION	# (Click to browse similar items)	STATUS
HATHITRUST 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](#)

IUG 2013 San Francisco

## Linked Fields: Invalid MARC Coding

- Some records have doubled Variable fields instead of linked Variable-880 fields
- 880 coding missing

t	245	0	0	603 a [Calligraphic specimen, f17th century? .] h [electronic resource]
t	245	0	0	603 a [نموذج من فن الخط العربي] f:القرن 17م .] h [electronic resource]

- Looks OK in public catalog
- Appears to index OK

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.

## Linked 880 Fields

- **Missing codes or mistakes means 880 fields won't display**
- **Copy catalogers find 880 field coding hard to check**
- **Adding 880 fields to Chinese records**



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.

## Record with 880 fields

The screenshot shows a Sierra library catalog record for 'Irwin T. and Shirley Holtzman'. The record is identified by the call number b998651x. The interface includes a search bar with the title 'Irwin T. and Shirley Holtzman' and a 'Search' button. The record details are as follows:

Field	Value
b998651x	Last Updated: 04-09-2013 Created: 04-06-2013 Revisions: 2
y 003	SKY
y 005	20130404230736.0
y 008	130404 s 1976 is 0 0 0 1 ara d
y 040	UtOrBLWjctUtOrBLW
y 041 1	arajhheb
y 049	EEMJ
y 050 4	PJ7802jb.B23 1976
c 090	PX]aP.J7802jb.B23 1976
y 090	PX]aP.J7802jb.B23 1976
a 100 1	[6880-01]a'bbāf, Mahmūdjd1935-
t 245 1 0	[6880-02]a'wdat 'Alf Bābā : [bmasrahyah ]cMahmūd 'Abbāf.
u 246 1 5	[6880-03]a'suvo shel 'Ali Baba
e 250	[6880-04]a'-'Tab'ah 2.
p 260	[6880-05]a'su'afat, al-Quds : [bMatba'at al-Sharq al-ta'wuniyah ]c1976.
r 300	48 p. :c21 cm.
n 500	MSU: Gift of Irwin T. and Shirley Holtzman.
n 500	Play.
s 830 0	Irwin T. and Shirley Holtzman Israeli Literature Collection.]5MIEM
y 880 1	[6100-01]k'جوليس منصور.]d1935-
y 880 1 0	[6245-02]k'محور جوليس منصور:] حواد علي بابا
y 880 1 5	[6246-03]l'عبدالرحمن بن يوسف بن عبد الوهاب
y 880	[6250-04]l'الطريقه
y 880	[6260-05]l'سيرة النبي صلى الله عليه وسلم:] حصاد التاريخ



This is a typical record with multiple 880 fields linked to variable fields.

## Record with 880 fields – Public View

<b>Author</b>	Abbāsi, Mahmūd, 1935-	محمود عبّاسي، 1935-
<b>Title</b>	'Awdat 'Alī Bābā : masrahiyah / Mahmūd 'Abbāsi.	أفودة علي بابا 2 مسرحية / محمود عبّاسي.
<b>Publisher</b>	Shu'afat, al-Quds : Matba'at al-Sharq al-ta'i'awuniyah, 1976.	شفاط القدس : مطبعة الشرق التاوينية، 1976.
<b>Citation</b>	No citation available. See <a href="#">Citation Help</a> .	
<b>LOCATION</b>	<b>CALL # (Click to browse similar items)</b>	<b>STATUS</b>
MSU SPEC COLL RARE BOOKS	XX PJ7802 .B23 .1976	LIB USE ONLY
<b>Edition</b>	al-Tab'ah 2.	الطبعة 2.
<b>Description</b>	48 p. ; 21 cm.	
<b>Series</b>	Irwin T. and Shirley Holtzman Israeli Literature Collection.	
<b>Local Note</b>	MSU: Gift of Irwin T. and Shirley Holtzman.	
<b>Note</b>	Play.	
<b>Other Title</b>	Shuvo shel 'Alī Baba	
	שׁוּבוֹ שֶׁל אַלִּי בָּבָא	

Permanent record link: <http://catalog.lib.msu.edu/record=b9898651~S39a>

This is the public catalog view of the previous record. I have added the yellow highlighting to show where the Arabic & Hebrew scripts display.

## Record with Table of Contents in 880 field

<b>Author</b>	Appefeld, Aron.		אפפלד, אהרן
<b>Title</b>	Tor ha-pe'ot / Aharon Appefeld.		תור הפלאות / אהרן אפפלד
<b>Publisher</b>	[Tel Aviv, Israel] : ha-Kibuts ha-me'uhad, 738 [1978]		[תל-אביב] : הקיבוץ המאוחד, תשל"ח [1978]
<b>Citation</b>	No citation available. See <a href="#">Citation Help</a> .		
<b>LOCATION</b>	<b>CALL # (Click to browse similar items)</b>	<b>STATUS</b>	
MSU MAIN LIBRARY	PJ5054.A755 T6 1978 c.2	AVAILABLE	
MSU SPEC COLL RARE BOOKS	PJ5054.A755 T6 1978	LIB USE ONLY	
<b>Description</b>	176 p. : 21 cm.		
<b>Series</b>	Irwin T. and Shirley Holtzman Israeli Literature Collection.		
<b>Note</b>	A novella and a short story.		
<b>Contents</b>	Table of contents only in Hebrew.		
<b>Local Note</b>	MSU: Copy 1 gift of Irwin T. and Shirley Holtzman.		
<b>Subject</b>	Novels -- Austria -- Fiction.		
<b>Genre</b>	Israeli literature.		
<b>Add Title</b>	Age of wonders		
<b>ISBN/ISSN</b>	sky208294100		



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.

# Linked 880 Fields in SkyRiver

The screenshot shows the SkyRiver catalog interface. At the top, there is a search bar with the text 'sky253617128'. Below the search bar, there are several buttons: 'Back to results', 'Previous 19 of 19', 'Export', 'Verify Headings', 'Find', 'Check Spelling', 'Search Authorities', 'Edit 33X', 'Derive', 'RDA Toolkit', 'Print Record', and 'Update Master Record'. The main content area displays a record for 'sky235054639' with various fields and a list of linked 880 fields. The linked fields are shown with a bracketed notation, such as '020 [ 9784061818835 (pbk.) ]'. The interface is powered by SkyRiver.

Field	Value
020	[ 9784061818835 (pbk.) ]
020	[ 406181883X (pbk.) : jc Y1456 ]
040	[ SFR :jc SFR :jd LPU :jd OCLCQ :jd JTE ]
049	[ EEMR ]
066	[ :jc \$1 ]
092	[ Fic :jb Kyogoku ]
100	[ 1 Kyogoku, Natsuhiko, :jd 1963- ]
100	[ 1 京極夏彦, :jd 1963- ]
245	[ 1 0 Tesso no ori / :jc Kyogoku Natsuhiko. ]
245	[ 1 0 鉄鼠の檻 / :jc 京極夏彦. ]
260	[ Tokyo : :jb Kodansha, :jc 1996. ]
260	[ 東京 : :jb 講談社, :jc 1996. ]
300	[ 826 p. : :jb ill. ; :jc 18 cm. ]
490	[ 1 Kodansha noberusu ]
490	[ 1 講談社ノベルス ]
504	[ Includes bibliographical references (p. 826). ]
655	[ 7 Mystery fiction. :j2 gsaif ]

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.

## Right-to-Left Script Display

Result page: ◀ PREVIOUS NEXT ▶

**On Left**

<b>Author</b>	Sadr al-Dīn Shīrāzī, Muhammad ibn Ibrāhīm, d. 1641 صدر الدين شيرازي، محمد بن إبراهيم
<b>Title</b>	[Asrār al-āyāt wa-anwār al-bayyināt ... etc., 193-?] [electronic resource]. [أسرار الآيات وأنوار البينات ... الخ 193-?].
<b>Publisher</b>	[193-?].
<b>Citation</b>	No citation available. See <a href="#">Citation Help</a> .

**On Right**

<b>Author</b>	'Abbāsī, Mahmūd, 1935-	عباسي، محمود، 1935-
<b>Title</b>	'Awdat 'Alī Bābā : masraḥīyah / Maḥmūd 'Abbāsī.	عودة علي بابا : مسرحية / محمود عباسي.
<b>Publisher</b>	Shu'afat, al-Quds : Maṭba'at al-Sharq al-ta'āwunīyah, 1976.	شعافط القدس : مطبعة الشرق التعاونية، 1976.

---

IUG 2013 San Francisco

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!**