

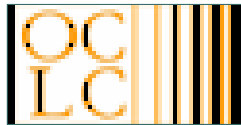


Hebrew Cataloging in OCLC

41st AJL Conference

Cambridge, MA

June 20, 2006



Jasmin Nof
Judaica and Hebraica Cataloger
University of Maryland

Hello, my name is Jasmin Nof and I'm the Judaica and Hebraica cataloger at the University of Maryland. Today I'll be sharing with you some of my experiences with the Hebrew cataloging features of OCLC.




Outline

- History
- Preliminary steps
- Macro
- Bi-directional text
 - Input
 - Display
- Conclusions for the future

We'll start with a brief history of vernacular cataloging in OCLC; then I'll share with you some of the preliminary considerations that had to be taken into account when deciding whether or not to add vernacular to OCLC records and to Maryland's catalog; next, I'll show you some examples of Hebrew text created by the Macro that translates from the standard romanization to Hebrew characters; I'll also describe the challenges of bi-directional text input and reading. Finally I'll summarize, and yield the floor to our next speaker.

To begin, let us briefly consider OCLC's history of vernacular cataloging. The Chinese/Japanese/Korean languages were the first non-Roman languages for which the vernacular characters could be added to records. This capability was introduced in May 1986. Last year, almost twenty years later, July 2005, the ability to include Hebrew characters was included in OCLC's Connexion Client Version 1.4 release. With the approval of the University of Maryland library administration and Jewish Studies faculty, I began inputting parallel fields to Hebrew and Yiddish bibliographic records. The following are some challenges I've encountered along the way.



Preliminary steps: Introduction

catalogusmai sign in | end ses

[basic search](#) | [advanced search](#) | [multi-campus search](#) | [recent searches](#) | [basket](#) | [results](#) | [preferences](#)
[CATALOG](#) | [JOURNALS](#) | [COURSE RESERVES](#) | [choose campus](#)

▶ You are here: [Catalog - University of Maryland, College Park](#) [library home](#)

Alphabetical list of results: Titles [prev pg](#) [next pg](#)

No. of Recs	Entry	
1	Russia land of the tsars	
1	א"א : היה או לא היה	
1	א.ב. קאהאן ביבליאגראפיע : מיט ביאגראפישע דאטעס / צוזאמענגעשטעלט פון יעפים ישורין	
1	אבי ואולה"י	
1	אברהם ב. יהושע	
1	אגדה למעשה : עיונים בסוגיות משפחה, חברה ועבודת השם	
1	אגמים	
1	אגרות אה"ח	
1	אגרת למנהיגי יהודי-ציוני ישראל	
1	אדום צהוב שחור	

[prev pg](#) [next pg](#)

First, there is nothing more exciting than seeing Hebrew displayed in the library's OPAC! This is a listing of titles found just by browsing on the letter "alef". Of course, before I could input any Hebrew, the language needed to be added to computer in the Control Panel's "Regional and Languages" options. Then, lacking a Hebrew/English keyboard, I found an image of the Hebrew keyboard map online and printed it out for reference. Soon I was able to touch type in Hebrew. The amount of time needed to input the characters was one of the early factors considered when deciding whether or not I should add Hebrew text to records. On average, it takes me about 3 minutes to add parallel information for the 100, 245 and 260 fields.

**Preliminary steps:
Placement of parallel fields, cont.**

Library of Congress >> Go to Library of Congress Authorities

LIBRARY OF CONGRESS ONLINE CATALOG

Help | New Search | Search History | Headings List | Titles List | Request an Item | Account Status | Other Databases | Start Over

040 __ |a DLC-R |c DLC-R
 042 __ |a pcc
 043 __ |a a-is--
 050 00 |a DS109.86.B37 |b A3 2004
 066 __ |c (2)
 100 1 |6 880-01 |a Baram, Haim.
 245 10 |6 880-02 |a Adom, tsahov, shahor / |c Hayim Bar'am.
 246 1_ |i Title on t.p. verso: |a Red yellow black
 260 __ |6 880-03 |a Tel-Aviv : |b Sifriyat Ma'ariv, |c c2004.
 300 __ |a 134 p. ; |c 21 cm.
 600 10 |a Baram, Haim |x Childhood and youth.
 651 _0 |a Jerusalem |x History.
 650 _0 |a Children |z Jerusalem |v Biography.
 880 __ |6 100-01/2r |a ברעם, חיים /
 880 10 |6 245-02/2r |a אדום, צהוב, שחור / |c חיים ברעם.
 880 __ |6 260-03/2r |a תל-אביב : |b ספרית מעריב, |c c2004.

Adom, tsahov, shahor / Hayim Bar'am
 r: 2005551902
 l: Book (Print, Microform, Electronic, etc.)
 n: [Baram, Haim](#)
 ברעם, חיים
 Adom, tsahov, shahor / Hayim Bar'am.
 אדום, צהוב, שחור / חיים ברעם.
 Tel-Aviv : Sifriyat Ma'ariv, c2004.
 תל-אביב : ספרית מעריב, c2004.
 134 p. ; 21 cm.

Once the green light was given to add the parallel, vernacular fields, we needed to consider how that information would appear in the MARC display of the OPAC.

For example, in its MARC display, the Library of Congress has vernacular information at the bottom of records.

For those unfamiliar with some of the coding, the 880 field is the field used for parallel information, indicating the matching tag in the subfield 6. So for example, this 880 field -- ברעם, חיים -- matches the 100 field -- Baram, Haim. The (2 is notation for Hebrew characters; each character set has its own code, designated in the 066 field.

This inset shows us the default public display for the same title. Here we see that parallel information is displayed together, with the Hebrew right-justified.

Preliminary steps: Placement of parallel fields, cont.

catalogusmai sign in | enc

[basic search](#) | [advanced search](#) | [multi-campus search](#) | [recent searches](#) | [basket](#) | [results](#) | [preferenc](#)
 CATALOG | JOURNALS | COURSE RESERVES | choose campus

▶ You are here: Catalog - University of Maryland, College Park [library home](#)

Full View of Record

[add to my saved items](#) | [add to basket](#) | [save/e-mail](#)
 Choose display: [standard](#) | [named MARC tags](#) | [MARC tags](#)

Record 1 out of 1 [find it](#)

FMT	BK	
042		a pcc
043		a a-is---
049		a UMCP
050 4		a PJ5055.15.A731 b A36 2004
066		c (2
100		6 880-01 a Baram, Haim.
100		6 880-01 a Baram, Haim. a/2)/100-01 6
245 10		6 880-02 a Adom tsahov shahor / c Hayim Bar'am.
245 10		6 880-02 a Adom tsahov shahor / c Hayim Bar'am. a/2)/245-02 6
246 1		i Title on t.p. verso: a Red yellow black
260		6 880-03 a Tel Aviv : b Sifriyat Ma'ariv, c c2004.
260		6 880-03 a Tel Aviv : b Sifriyat Ma'ariv, c c2004. a/2)/260-03 6
300		a 134 p. ; c 21 cm.
4900		6 880-04 a Siporet
4900		6 880-04 a Siporet a/2)/490-04 6
60010		a Baram, Haim x Childhood and youth v Fiction.


author [Baram, Haim.](#)
 ברעם, חיים.

title [Adom tsahov shahor / Hayim Bar'am.](#)
 אדום צהוב שחור / חיים ברעם.

published Tel Aviv : Sifriyat Ma'ariv, c2004.
 תל אביב : ספריית מעריב, c2004.

description 134 p. ; 21 cm.

notes (Siporet)
 (סיפורת)



Maryland's catalog is Ex Libris' Aleph 500, currently version 16.02. In our MARC display, Hebrew fields appear directly after the matching roman field, with duplicate tagging. Whereas in LC's catalog we saw just one 100, 245, etc., with the vernacular in 880's, Maryland has 2 100's, 245's etc., with the vernacular appearing in the 2nd one. This coding is NOT standard for a library catalog, though it is the structure used by OCLC.

This inset shows us the default public display for the same title. Again, the parallel information is displayed together, but the Hebrew is left-justified. Personally, I find this more difficult to read than LC's right justification. I'm not sure if it's an Aleph system setting, or if it's something Maryland can independently control.

The coding structure may seem somewhat irrelevant, especially since the OPAC display is so similar, but it has since caused some unexpected challenges. Maryland has been working with an external vendor for the purpose of creating and maintaining an authority file. Because our non-roman fields are not coded as 880's, as the vendor expected, the initial batch of Hebrew records sent for processing came back with unsatisfactory results such as indicator switches and other character errors. For now, records with non-roman characters have been excluded from the authority control cleanup while we explore ways to solve the problem in our local system. Moral of the story: if your catalog has Hebrew and your library is considering an external vendor for database work, be sure your record structure is standard.



OCLC Record with Macro Conversion

042		pcc
043		a-is—
050	4	JQ1830.A95 #b W45 2005
090		#b
049		UMCP
100	1	וייס, שבח.
100	1	Weiss, Shevah.
245	1 0	שבח וס, דוד ישכול #c / שלטון ובחירות, העשור האחרון #b : בחרתיך ומעסתך
245	1 0	Beḥartikha u-me'astikha : #b shilṭon u-veḥirot, he-'ašor ha-aḥaron / #c Shevah Ṷals, Dayid Eshkol.
260		תל אביב : תל אביב : #b לוגוס #c, [2005]
260		Tel Aviv : #b Logos, #c [2005]
300		277 p. : #b ill. ; #c 21 cm.
500		Title page partially vocalized.
650	0	Elections #z Israel .
651	0	Israel #x Politics and government .
700	1	אשכול, דוד.
700	1	Eshkol, David.

Before going to the trouble of learning the Hebrew keyboard layout, I was encouraged by my supervisors to explore the functionality of the much touted language macro. There are transliteration macros for both Hebrew and Yiddish that, when invoked, produce the vernacular field parallel to a romanized one, or vice versa, will produce the romanization from the vernacular. Since the issue was more with the time needed to input vernacular, I was most interested in testing the roman-to-Hebrew conversion macro.

Unfortunately, thus far, I've found the Macro to be lacking. I chose this record as an example, "בחרתיך ומעסתך : שלטון ובחירות, העשור האחרון."

● ● ● | OCLC Record with Macro Conversion, cont.

שבבוס, דוד ישכול א' / שלטון בחירות, העשור האחרון ב' : בחירת ומערכת
 Behartikha u-me'astikha : #b shilton u-vehirot, he-'asor ha-aharon / #c Shevah Vais, David Eshkol.

Looking at the title more closely, we can see several anomalies. Because I didn't want the slide to get too crowded, I only illustrate one example of each error type: The כף סופית of בחרת is given as a כ; the ו of ומעסתיך is missing, and the system doesn't know what to do with the hyphen used in romanization to indicate the presence a "little word"; מעסתיך also illustrates a case of כתיב חסר, where transcription requires a י; אחרון is missing its א. These are more of the same; but in וייס, the double-י goes undetected. And for some bizarre reason, the א of אשכול is given as a י. These are all matters that affect access.

● ● ● | Macro conversion, OCLC to Aleph

OCLC display:

בחרתיך ומאסתיך : #b שלטון ובחירות- העשור האחרון / #c שבח וייס, דוד אשכול.

Aleph display:

בחרתיך ומאסתיך : שלטון ובחירות-- העשור האחרון / שבח וייס, דוד אשכול.

OCLC display:

בחרתיך ומאסתיך : #b שלטון ובחירות- העשור האחרון / #c שבח וייס, דוד אשכול.

Aleph display:

בחרתיך ומאסתיך : שלטון ובחירות-- העשור האחרון / שבח וייס, דוד אשכול.

Even if you chose to make all the corrections manually and only keep the דג points generated with the macro, in this case, those for the ב, ש's, and the כ, what happens upon export to Aleph is that the marks are shifted: the דג of the bet to the n; the ש's to the ל, ב, and כ, respectively, and that of the כ to the ו – rendering it a שורוק. In other words, you get a distorted OPAC display.

So, for now, we are left to input the vernacular manually, with no bells and whistles. But at least in so doing we get a clean – and accurate – display.

Another issue I had when inputting Hebrew characters arose when I tried to be efficient and copy the letters from the RLIN record for the book I was cataloging. I'm not sure if it's because I can only access the RLIN catalog via a browser and not the client, but OCLC doesn't understand the Unicode values of the characters pasted in, so the record fails validation. Granted, with the merging of the two utilities, this will soon be a moot point.



Bi-directional text

245	1	0	י-ברזילי #c / שבע-עשרה שיחות עם אסא כשר
245	1	0	Sheva'-ešreh śiḥot 'im Asa Kasher / #c Ve
246	1		#i Title on verso of t.p.: #a Seventeen conv
246	1	4	שיחות עם אסא כשר 17
246	1	4	17 śiḥot 'im Asa' Kasher

- Undo
- Cut
- Copy
- Paste
- Delete Field
- Copy Field
- Copy Control Number
- Control Single Heading
- Uncontrol Single
- Transliterate
- Right-to-Left Reading Order

Our next topic is the input and display of bi-directional text. Reading order can be complex in mixed fields, such as those with both Hebrew letters – read Right-to-Left – and numerals – read Left-to-Right. For this title, “שבע-עשרה שיחות עם אסא כשר,” the number 17 is spelled out and we have an added title with the numerals, which we can see have been moved to the end of the vernacular 246 field. We’d like for the number to be transposed, so it comes at the beginning of the vernacular title, just as it does for the romanized title. Using the right mouse button, we can select the optional “Right-to-Left Reading Order” feature.

● ● ● | Bi-directional text, cont.

245	1	0	ורד לוי-ברזילאי #c / שבוע-עשרה שיחות עם אסא כשר	
245	1	0	Sheva'ešreh śiḥot 'im Asa Kasher / #c Ṿered Leyi-Barzilai.	
246	1		#i Title on verso of t.p.: #a Seventeen conversations with Asa Kasher	
246	1	4		17 שיחות עם אסא כשר
246	1	4	17 śiḥot 'im Asa' Kasher	

Uncontrol Single
 Transliterate
 Right-to-Left Reading Order
 Insert from Cited Record

 Link Fields
 Uplink Fields

 MARC Field Help

Aleph display:

17 śiḥot 'im Asa' Kasher

17 שיחות עם אסא כשר

Having done so, the field is flipped and our numeral does appear at its beginning; we can verify that we're using the Right-to-Left reading order feature by right clicking again. And though the default OCLC view is the one we saw above, with the numerals and words reversed, the OPAC displays the title correctly.

● ● ● | Bi-directional text, cont.

Entered Right-to-Left as:

245	1	0	עליתו ונפילתו של ה-#Y / #c יהונתן גפן ; עורך, רפי קינן
245	1	0	'Aliyato u-nefilato shel ha-Y / #c Yehonatan Gefen ; 'orekh, Rafi Kenan.

Default view:

245	1	0	יהונתן גפן ; עורך, רפי קינן #c / -Y עליתו ונפילתו של ה
245	1	0	'Aliyato u-nefilato shel ha-Y / #c Yehonatan Gefen ; 'orekh, Rafi Kenan.

Right-to-Left view:

245	1	0	עליתו ונפילתו של ה-#Y / #c יהונתן גפן ; עורך, רפי קינן
245	1	0	'Aliyato u-nefilato shel ha-Y / #c Yehonatan Gefen ; 'orekh, Rafi Kenan.

Here's another example, this time with a Roman/non-Roman character mix. Whenever I'm faced with a case such as this, I prefer to input the characters in the Right-to-Left reading order, just for the sake of clarity. Again, though, the default view will shift the Roman character to the beginning of the non-Roman word string. Even more confusing, when I return to the Right-to-Left reading order, the ISBD punctuation is interfiled with the Roman letter that belongs to the non-Roman title proper. Thankfully, the OPAC displays the characters in proper order.



Bi-directional text in OPAC

catalogusmai

[sign in](#) | [end session](#) | [my account](#) | [help](#)

[basic search](#) | [advanced search](#) | [multi-campus search](#) | [recent searches](#) | [basket](#) | [results](#) | [preferences](#)

[CATALOG](#) | [JOURNALS](#) | [COURSE RESERVES](#) | [choose campus](#)

▶ You are here: [Catalog - University of Maryland, College Park](#) | [library home](#)

'Aliyato u-nefilato shel ha-Y / Yehonatan Gefen ; 'orekh, Rafi Kenan.

עלית ונפילתו של ה-Y / יהונתן גפן ; עורך, רפי קינן.

● ● ● | Bi-directional text, cont.

(העורכים, מנחם מור ... [אחרים] / #c מחקרים בתולדות ישראל בעת העתיקה : מוגשים לאוריאל רפפורט #b : לאוריאל
 Le-Uri'el : #b meḥqarim be-toldot Yisra'el ba-'et ha-'atikah : mugashim le-Uri'el Rapaport / #c ha-'orkhim, Menahem Mor ... [et al.]

Le-Uri'el : meḥqarim be-toldot Yisra'el ba-'et ha-'atikah : mugashim le-Uri'el Rapaport / ha-'orkhim, Menahem Mor ... [et al.]
 ([אחרים]) : לאוריאל : מחקרים בתולדות ישראל בעת העתיקה : מוגשים לאוריאל רפפורט / העורכים, מנחם מור ... [אחרים]

700	1	מנחם, מור.	■ <u>Mor, Menahem.</u>
700	1	<u>Mor, Menahem.</u>	■ מור, מנחם.

Punctuation is also transposed to the opposite end of the Hebrew subfield; here we see square brackets that, though input correctly, are reversed and closed at the beginning of the Hebrew field; and here a period is moved. Again, in both these cases, the OPAC display is correct.



Bi-directional text, cont.

Default view:

הופק עבור ערוץ 8, נגה תקשורת; חי כהן, צילום, הקלטה, עריכה, הפקה, ובימוי; הקרן החדשה לקולנוע - #c [videorecording] / #h מודל מוקנדה וטלויזיה = Mukanda Model / Produced for Channel 8, Noga Communication photographed, recorded, edited, produced, and directed by Hai Cohen; New Israel Foundation for Cinema and Television.

Right-to-Left view:

מודל מוקנדה #c [videorecording] / #h הופק עבור ערוץ 8, נגה תקשורת; חי כהן, צילום, הקלטה, עריכה, הפקה, ובימוי; הקרן החדשה לקולנוע וטלויזיה = Mukanda Model / Produced for Channel 8, Noga Communication photographed, recorded, edited, produced, and directed by Hai Cohen; New Israel Foundation for Cinema and Television.

Romanization:

Model Mukandah #h [videorecording] / #c hufak 'avur 'Aruts 8, Nogah Tikshoret; Hai Kohen, tsilum, ha-klatah 'arikhah, hafakah, u-vimui; ha-Keren ha-hadashah le-kolno'a ye-televizyah = Mukanda Model / Produced for Channel 8, Noga Communication; photographed, recorded, edited, produced, and directed by Hai Cohen; New Israel Foundation for Cinema and Television.

Long titles are particularly challenging to read in OCLC. Where the text spans more than one line, the cataloger needs to train his eyes to follow the language; someone reading from right to left would need to figure out that, after הקרן החדשה לקולנוע, he needs to read the next line from after the = sign, וטלויזיה.

In the right-to-left display, reading the Hebrew is straight-forward, but reading the English requires visual gymnastics: edited, produced, and directed ---, by Hai Cohen. And you'll note the period that is placed at the end what would be a Hebrew line.

So while having the parallel field with the vernacular is nice, in cases of long text it's probably easiest on the eyes to work with the transliterated field, which reads in only one direction.



Conclusions for the future

- Excited with new vernacular feature
- Room for improvement
 - Macro
 - Directionality of text

Thanks!

In summary, during this last year, I've enjoyed working with the new Hebrew character feature. Patrons appreciate the added information and access, and this far outweighs any frustrations I may have experienced. Of course, I am also hopeful that improvements in coming releases will make Hebrew cataloging in OCLC even better.

Thank you for your attention.

For more information, please contact Jasmin Nof, Judaica and Hebraica Cataloger;
2200 McKeldin Library, University of Maryland, College Park, MD 20742;
301-405-9330; jnof@umd.edu