

# The Use of Israeli Machine-Readable Cataloging by American Libraries: A Proposal

**Elhanan Adler**  
University of Haifa  
Haifa, Israel

## Introduction

Israeli manual cataloging practice has generally been based on American cataloging standards and procedures, with certain adjustments for Judaica and Hebraica, the most significant being the maintenance of a separate Hebrew catalog for Hebraica with all access points in Hebrew. With the advent of the computer, however, a major new gap has arisen between Israeli practice and American or international standards. While a conscious national decision has never been made on this question, Israeli bibliographic data is being encoded and stored using locally developed formats which solve specific local problems, but present a significant barrier to the "export" of bibliographic data. Some of these problems and their background were presented in my article "Hebrew cataloging and the computer—the view from Israel" (Adler, 1982).

Outside of Israel, there is considerable interest in the use of Israeli cataloging data. The development of a non-Roman capability by RLIN has led to solutions for many of the basic questions and format problems related to handling "vernacular" records within MARC (machine-readable cataloging). RLIN has already implemented Chinese-Japanese-Korean and Cyrillic vernacular cataloging, and Hebrew is now (Fall 1987) fully designed and close to implementation (Aliprand, 1987). Libraries with significant Hebraica collections expect to have a Hebrew capability available in the shared cataloging networks very soon and are beginning to be concerned about the source of Hebrew cataloging data. While libraries using OCLC or RLIN are willing to input some original cataloging, they generally expect to find the majority of their required cataloging data already in the system, ideally, at the time the library receives an item.

Since Israeli libraries are presumably purchasing virtually all local Hebraica close to publication and (presumably) cataloging these items on a current basis, and since most American libraries receive these items significantly later (because of notification delay, surface shipping, etc.), the ideal solution would be the availability of current Is-

raeli Hebrew cataloging in an American cataloging utility.

The proposal to make Israeli Hebrew cataloging available to American libraries raises many questions and problems with regard to cataloging practices, format, compatibility, etc. This article addresses these problems in some detail and poses several possible solutions which could lead to American use of Israeli cataloging data.

## Automating Hebrew Cataloging in the United States

Before dealing with the problems related to American use of Israeli cataloging data, several problems should be noted which relate to Hebrew cataloging in America in general.

1. *Hebrew Character Set*—Before any attempt at processing Hebrew data can be made, there must be agreement on the Hebrew character set and its coding. At a minimum, 27 characters must be accommodated (22 letters + 5 final forms). The RLIN character set is much more extensive, including vowel points and special versions of characters, particularly for Yiddish. (It should be mentioned that standard American and Israeli practice has been to use only the 27 consonantal characters—if only for the pragmatic reason that these are the only characters available on most Hebrew typewriters.) As a starting point, there should be agreement among LC, OCLC and RLIN as to the character set to be used and its coding.
2. *Romanization*—American manual cataloging practice for non-Roman languages has been to provide all access points in vocalized Roman script. The non-Roman transcription has been limited to the body of the entry, series statements, and some notes. With the expansion of the LC MARC cataloging service to non-Roman languages, the entire cataloging record is now entered in Romanized form. Since Hebrew is written virtually vowelless, and LC Romanization practice is to enter the text *as pronounced*, it is therefore necessary to reconstruct and insert all vowels (based upon grammar and context). This is an

extremely time-consuming and error-prone process. In a recent article on the Ohio State University online catalog, Amnon Zipin, Jewish Studies Bibliographer at Ohio State, notes:

*The area in which we encountered the greatest difficulty was in Romanizing our Hebrew and Yiddish records according to LC. . . . Even with a complete set of rules, LC Romanization is highly complicated. Vocalization of Hebrew according to LC's system demands a highly sophisticated knowledge of grammar and modern Hebrew usage. Such knowledge, combined with the additional time needed in the cataloging process to do complete Romanization is making the cost of producing the online Romanized record much higher than that of producing the traditional card. (Reversible transliteration which does not reconstruct missing Hebrew vowels, is much more cost-effective, and allows a future machine-conversion to original script display when the technology becomes available.) Only our dependency on LC cataloging and compliance with OCLC requirements convinced us to adhere to this system reluctantly. (Zipin, 1984, p. 56).*

A recent Association of Jewish Libraries Cataloging Workshop organized by Bella Hass Weinberg and Pearl Berger emphasized this problem: catalogers from some of the major Judaica libraries in the New York area were asked to Romanize certain Hebrew records, and the results were then compared with LC's Romanization of those records. The results showed considerable differences amongst the librarians participating in the experiment.

Already in 1977, both the Association of Jewish Libraries and the Council of Archives and Research Libraries in Jewish Studies called upon LC and the bibliographic utilities to use machine-reversible transliteration (letter-for-letter substitution) rather than phonetic transcription (Zipin, 1984, p. 53). This proposal was, however, not adopted.

Regardless of display of vernacular Hebrew, accurate Romanization of Hebrew bibliographic data will still remain crucial because:

- a) The LC/RLIN standard for processing

Chinese-Japanese-Korean (CJK) records (which seems to be the standard approach to non-Roman cataloging) mandates a *full* Romanized record to which are optionally appended parallel vernacular fields.

b) Access to these records by libraries without Hebrew display capability (e.g., for interlibrary loan) will always be via the Romanized form. It is particularly in these libraries that the expertise required for accurate Romanization will be most lacking.

c) Any attempt at machine identification of identical records (a subject becoming of greater and greater importance in bibliographic utilities and union catalogs) will be severely hampered by variations in Romanization.

The fact that RLIN does not mandate a uniform Romanization scheme and that NYPL's version of the ANSI standard for reversible Romanization and ALA/LC Hebrew Romanization are both being used serves to further compound these problems. [Subsequent to the writing of this paper, RLIN's BibTech committee voted against endorsing ANSI as an alternative standard.—Ed. (B.H.W.)]

3. *Vernacular Access Points and Hebrew Orthography*—As mentioned above, American manual cataloging practice has been to record all access points in either Roman or Romanized form, regardless of the script of the title, series, etc. There was therefore no need for vernacular access points (other than title/series added-entry cards in some libraries), or to supply the authority control which vernacular access points would require. LC/RLIN practice with regard to CJK cataloging indicates that access points will also be entered in the Hebrew vernacular. Besides the additional authority control required, a special problem exists with regard to Hebrew.

Nonvocalized Hebrew is written in either of two officially recognized forms: *ktiv haser* (defective or minimal form) or *ktiv male* (plene form) (in the latter form, two "consonantal-vowels" are added to aid in pronunciation). Phonetically, there is no difference between the two. Current Hebrew practice is inconsistent in the use of these forms, although the trend seems to be towards the fuller form. The Israeli library tradition, however, favors the minimal form.

Since a reader cannot be expected to know whether a given title, person, corporate entry, or series appeared in the full or short form, the Israeli solution has been to record all Hebrew *access points* in a standard form, regardless of the spelling in the work.

The title, edition, and imprint are recorded, however, exactly as written. To further complicate matters, Israeli libraries have been divided as to which system to use (for details see Adler, 1982, pp. 241–242). Only recently has a decision been taken mandating the minimal form in the inter-university cataloging network.

The extended use of Hebrew access points (other than titles and series) by American research libraries makes imperative a decision by LC and the major Hebraica libraries on the forms to use and the means for authority control.

4. *Input and Display of Data and Printing of Catalog Cards*—Hebrew presents a problem not encountered in CJK cataloging in that Hebrew is written from right to left, while the natural orientation of a MARC record is from left to right. While it is theoretically possible to enter and edit Hebrew data from left to right (and it has been done on occasion), this is an extremely tedious and error-prone method. Preferably, the cursor should move from right to left (as is done on Israeli terminals) but this would clash with the MARC apparatus (tags, indicators, subfield codes, etc.). One approach is to "freeze" the cursor after the escape-to-Hebrew sequence while each character keyed would appear at the cursor position, pushing the remainder of the Hebrew text to the right (the data, of course, would be stored in the order keyed). Because of problems in maintaining the logical sequence of bilingual text, RLIN has opted for a more complicated solution: embedding directional "flags" in the data field (for details, see Aliprand 1987, p. 8). Special care must be taken also in the display and printing of Hebrew data, particularly in systems showing a "user-friendly" form of the record.

Since many large Judaica libraries do not yet contemplate closing their catalogs, there will be a demand for Hebrew catalog cards. Here also, aside from the basic hardware capability, there are complicated programming changes required to maintain the logical flow of what is often bidirectional data. In some cases, it may be necessary to split lines in mid-paragraph to maintain coherency, as was done with the New York Public Library's Dictionary Catalog (Malinconico, 1977). LC (which has more experience in formatting Hebrew cards than any other library) and the major Judaica libraries should be consulted in setting up the exact parameters for producing Hebrew cards, including the question of vernacular access points which may appear in the record. RLIN has avoided this problem to date, and Hebrew catalog cards are not part of its current Hebrew program (Aliprand 1987, p. 13).

## Israeli Machine-Readable Cataloging

As mentioned above, and as detailed in my previous article (Adler, 1982), Israeli libraries have de facto opted for a non-MARC solution to Hebrew cataloging problems rather than creating a local MARC format to accommodate them. Similarly, they have taken simple local solutions to presenting the Hebrew character set (either substitution for lower-case Roman or, recently on the IBM PC, substitution for graphics characters at ASCII decimal nos. 128–154—which clashes with the MARC extended character set). In retrospect, this approach was probably justified: the library automation scene in Israel is too small to justify custom-made terminals and printers, and the enhanced graphic capabilities of microcomputer-based terminals are just now coming into wide use. With regard to cataloging format, trying to solve both the general problems of non-Roman cataloging and the specific problems of Hebrew within the MARC format would probably have set development of automated Israeli library systems back several years and would have required extensive (and expensive) participation in international organizations and committees.

There is no question that ultimately, the Israeli library community will wish to align itself with international bibliographic standards and will demand additional character sets (Arabic at the very least). In view of the difficult financial situation in Israel in general (and particularly in Israeli universities), it is highly doubtful whether this worthy goal will be achieved in the near future.

Since there is interest in the use of Israeli cataloging data abroad, the question remains: given that there is no Israeli MARC format, is it still possible to manipulate, "massage," and otherwise upgrade an Israeli cataloging record to make it useful to the American bibliographic community? The answer to this question is a qualified yes.

## Adapting Israeli Cataloging Data to US MARC

All Israeli university libraries (with the possible exception of Bar-Ilan which, as of October 1987, was still undecided) have opted to use the Hebrew University's ALEPH system for their cataloging (for details on ALEPH see Levi, 1984). While ALEPH in theory allows great flexibility in defining fields, in actual practice Israeli libraries have agreed upon a standard set of field codes to facilitate the interchange of cataloging data. The cataloging record of an Israeli university library using LC classification and subject headings would contain most of the following fields—some in exact (title page)

Hebrew, some in normalized (minimal form) Hebrew, and yet others in Roman characters:

Main entry (Normalized Hebrew)

Uniform title (Normalized Hebrew)

Title, parallel title, subtitle, author statement (Exact Hebrew, separate fields)

Edition statement (Exact Hebrew)

Place, publisher, date (Exact Hebrew or Roman—depending on data, separate fields)

Collation (Normalized Hebrew, divided with ISBD punctuation)

Series traced (Exact Hebrew, divided with ISBD punctuation)

Series untraced (Exact Hebrew, divided with ISBD punctuation)

Note (Hebrew, exact or normalized)

Added entry (Normalized Hebrew)

Added title (Normalized Hebrew)

Added series (Normalized Hebrew)

Personal subject (Normalized Hebrew or English)

Topical subject: English (LCSH, with minor changes for Judaica)

Classification (LC, with minor changes for Judaica)

LC card number (virtually nonexistent in Israeli publications)

ISBN (becoming more frequent in Israeli publications)

Local data

An attempt to “map” such a record into USMARC format would encounter the following difficulties:

1. *Character set*—Translating the character set from one coding scheme to another should present no problems.

2. *MARC Tags*—While for some fields (LC classification, ISBN, title, imprint, etc.), there is one-to-one equivalence, for others, the level of detail is much lower in ALEPH. MARC distinguishes among several types of main entries, notes, etc. There will have to be some human, professional intervention, perhaps prompted by a computer-aided suggestion (e.g., key words in fields could trigger an “intelligent guess” as to the exact field type).

3. *MARC Indicators*—Many of the field indicators are unchangeable or determinable (e.g., title traced). Some will have to be manually set; once again, it should be pos-

sible to provide computer-aided selection in most cases.

4. *MARC Subfield codes*—The subfields of MARC tags 245 (title) and 260 (imprint) are separate fields in ALEPH and should be simple to append. Also, subfield codes of fields 300 (imprint) and 4xx (series) should be reconstructible from the ISBD punctuation. Some subfield codes will have to be manually set (e.g., subdivisions under subject headings), again with computer assistance in most cases.

5. *MARC fixed-length data fields* (particularly field 008)—Some of the data required can be derived from the record (e.g., language, place of publication), and some can be deduced (manually) from the record (intellectual level, government publication, fiction, etc.).

6. *Miscellaneous MARC 0xx fields: 039, 040, etc.*—The data required in these fields should usually be automatically obtainable.

In addition to these specific format questions, the following changes will also need to be made:

1. *Romanization*—All vernacular fields must become secondary fields (field 880), to be replaced by either systematic Romanization or established English-language forms. ALA/LC Romanization, which requires addition of vowels, cannot by any stretch of the imagination be performed automatically by computer.

2. *Authority control*—In order for a Romanized Israeli record to be used freely by American libraries, it must go through some kind of authority validation, preferably by checking against LC name authorities. Given time, it should be possible to create a Hebrew-English authority file which could also handle much of the Romanization of personal and corporate access points, solving part of the preceding problem.

3. *Within the vernacular fields* themselves there are differences between American and Israeli practice, e.g., in the method of recording Jewish-era dates. These will generally have to be handled manually. In addition, the decision by American libraries as to what form of orthography they prefer: full, short, compromise or straight title-page form, may create the need for changes within the vernacular access-point fields (straight title-page form in fields where Israeli libraries are using normalized Hebrew may often require reference to the work itself).

A standard Israeli Hebrew record would thus require extensive upgrading and editing in order to make it usable by the American li-

brary community. The questions remaining are: who should do this upgrading, how, and at what cost?

## Possible Solutions

A possible solution, mentioned only in passing because I do not consider it practical, would be for raw Israeli cataloging data to be passed on to an American library for upgrading and conversion. The logical candidate for such an arrangement would be the Library of Congress, which might subsequently distribute the records in a manner similar to that employed for its new CJK service (the CJK service consists at this point of LC's cataloging only). While this would result in high-quality cataloging data, particularly with regard to descriptive and subject authority work, successful implementation would require a significant addition to the LC Hebrew cataloging staff, which already faces a considerable cataloging backlog. Furthermore, resolution of many of the above-mentioned problems may require reference to the work itself, which may not have yet reached the Library of Congress. The amount of additional work to be done would seem to be so great that any benefit from having an Israeli “draft” record would be marginal.

A more practical, speedy, and cost-effective approach would be for an Israeli university library (either the Hebrew University's Jewish National and University Library or the central library of one of the non-technical universities) to do the necessary upgrading and conversion as part of its processing system. The resultant record would be in full USMARC format, usable by American libraries via one or more of the American cataloging utilities, in the same way as an original cataloging record input by an American member of the network. The effort involved (most of which is unnecessary from an internal Israeli standpoint) would have to be justified by external funding (at least at the initial stage); ultimately, it should be seen as the price of Israel's participation in the international bibliographic community.

An Israeli university library might undertake a pilot project together with a suitable partner (LC or one of the American bibliographic utilities) to attempt to convert its current Hebrew cataloging into USMARC format and “deposit” the resulting data in a shared cataloging network. This could be done in either of two ways:

1. Cataloging data would be input *directly* into the network data base via a terminal, using international and national communications networks. This would allow utilization of authority files within the network as

well as ensure availability of the data with maximum speed. Under this option, *all* data would have to be keyed into the system. Factors to be taken into account include system availability (because of the time gap), maintenance of equipment (in particular, any unique hardware adaptations), and communications costs.

2. Current cataloging data would be copied and manipulated within the University's own computer to produce a first draft of the MARC record, which would subsequently be edited to correct and add all data requiring human intervention. The corrected record would be sent to the network either by magnetic means or via electronic mail. Authority checking would be manual, using LC name authorities (microfiche edition) and other sources as appropriate. This option would require extensive software development; however, communications costs, availability, and maintenance problems would be minimal. It would, in any event, require the inputting Israeli library to use an American library standard terminal/micro-computer with multiple character sets.

Other in-between options are also possible (e.g., use of the network online only for authority checking), but these options seem to be the two basic viable possibilities.

### Conclusions

Current Israeli Hebrew cataloging could and should be available within American shared cataloging networks. The problems of format differences, while formidable, are not insurmountable; however, the extra costs involved in such an exchange of bibliographic data would have to be shared by the American library community.

### References

- Adler, Elhanan. "Hebrew Cataloging and the Computer—the View from Israel," *Information Technology and Libraries*, vol. 1, no. 3 (Sept. 1982), pp. 238–245.
- Aliprand, Joan M. "Hebrew on RLIN," *Judaica Librarianship*, vol. 3, no. 1–2 (1986–1987), pp. 5–16.
- Levi, Judith. "ALEPH: an Online Real-Time Integrated Library System," *Judaica Librarianship*, vol. 1, no. 2 (Spring 1984), pp. 58–63.
- Malinconico, S. Michael; Grutchfield, Walter R. "Vernacular Scripts in the NYPL Automated Bibliographic Control System," *Journal of Library Automation*, vol. 10 (Sept. 1977), pp. 205–225.

Zipin, Amnon. "Romanized Hebrew Script in the Online Catalog at the Ohio State University Libraries," *Judaica Librarianship*, vol. 1, no. 2 (Spring 1984), pp. 53–57.

*Elhanan Adler is Assistant Director of the University of Haifa Library, and has been involved in all its automation activities. He has also co-authored Hebrew cataloging texts (with Aviva Shichor), based on Anglo-American practice.*

Garland is pleased to announce the completion of:

# THE RISE OF ISRAEL

A DOCUMENTARY RECORD FROM THE  
NINETEENTH CENTURY TO 1948

A facsimile series reproducing over  
2,000 documents in 39 volumes

General Editor

HOWARD M. SACHAR

Editors

MICHAEL J. COHEN

ISAIAH FRIEDMAN

AARON S. KLIEMAN

For free descriptive brochure contact Henry Dalby



GARLAND PUBLISHING, 136 Madison Avenue  
New York, NY 10016, (212) 686-7492