# ELECTRONIC BOOKS IN THE ACADEMIC WORLD AND THE IMPLICATIONS FOR JEWISH STUDIES

*Heidi G. Lerner*

Imagine that you are teaching Jewish literature in translation. Very possibly you will incorporate as a text Robert Singerman's *Jewish Translation History: a Bibliography of Bibliographies and Studies* (2002). A search in your library's catalog directs you to the book on the shelf via a classification number. In addition, the catalog record contains a note, "Also available to [your university's] affiliates at eBrary." In other words, this book is available electronically via a commercial vendor to your institution's users on their computers.

Books have been available "electronically" for more than twenty years via CD-ROMs, diskettes, and other technologies. These works have not historically had much popular success, as the interfaces have not offered a pleasant environment for reading long texts when compared to looking at a printed page. Electronic versions of bibliographies, citation indexes, dictionaries, encyclopedias, abstracting and indexing services, and even journals have been much more successful.

The advent of the Internet has provided an avenue of distribution for electronic books (e-books) and texts that was unimaginable until a few years ago. The Internet not only offers a universal means of

access, it has also encouraged the development of new technologies to make electronic texts an effective medium for publishers and libraries to serve the scholarly community.



*Shaar Hamelech* by Rabbi Meyer Blumenfeld. Kew Garden Hills, NY, 1979. Courtesy of www.hebrewbooks.org

The advantages of electronic books are clear and abundant: the text is always there and accessible regardless of time or place, and editions will not go out of print. Features can include full-text searching, changeable fonts, hyperlinks that link to related resources within and beyond documents, note-taking, etc. The availability of this format in the academic environment is still limited, however, as libraries and institutions grapple with issues such as standards, proprietary and

incompatible formats, digital rights management, archiving, access, enhanced features, and content. More and more institutions, libraries, consortiums, publishers, vendors, and even individuals are beginning to create electronic book collections, but the number of books for use in Jewish studies is still small.

There has been much discussion on how to bring scholarly monographs into the digital environment. One of the first projects for creating a digital library was Project Gutenberg started by Michael Hart in the 1970s (www.promo.net/pg). Today it includes more than 10,000 e-books produced by hundreds of volunteers. These books are mostly older literary works in the public domain. All submissions must be accompanied by proof that Project Gutenberg may legally distribute the book, and with confirmation of its copyright status. Whenever possible, Project Gutenberg distributes a plain text version of an electronic book. Other formats are also accepted, but plain text is the "lowest common denominator." Text files are accessible on all computers and have demonstrated longevity which acts as insurance against future obsolescence. All works may be freely downloaded and read, and redistributed for non-commercial use. Titles are searchable by author, title, note, subject, language, file type, and Library of Congress classification number.

Project Gutenberg inspired Project Ben-Yehudah (benyehuda.org). This volunteer-supported project started by Asaf Bartov of Tel Aviv University, "aims to make accessible the classics of Hebrew literature (poetry and prose at first, and then essays, etc.) to the reader of Hebrew, amateurs and scholars." All that is required to view these works is an Internet browser that supports

Hebrew with diacritics such as Internet Explorer 5.0 (or higher). Texts are entered manually via a word processor and then converted into Hypertext Markup Language (HTML), and uploaded to the Web site. This method was selected because it enables full-text searching. This feature would not be possible if texts were scanned and loaded as image files. Optical Character Recognition (OCR) software is not being used since it does not adequately support vocalized Hebrew. So far, selections include works by Uri Gnessin, Rahel, Shelomoh ibn Gabirol, Yehudah Leyb Gordon, David Frishman, Abraham Ibn Ezra, Abraham Regelson, Ahad Ha'am, Eliezer Ben-Yehudah, and a host of others.

The number of other publicly accessible databases of electronic Hebrew texts is still small and these are mostly scriptural or religious in orientation. These include Mechon Mamre, which provides the Tanakh in Hebrew in four different editions, and an edition of Maimonides' Mishneh Torah (mechon-mamre.org). HebrewBooks.org is a project to preserve old Hebrew books published in the U.S. that are no longer in print or circulation (hebrewbooks.org). More than 1,200 works of American rabbinica have been scanned, and are currently available to the public in Portable Document Format (PDF) as well as for downloading and printing. Other freely accessible sources of online Jewish texts can be located via Jewish studies resource pages and reference guides. As in the above examples, texts are usually submitted to these databases either via scanning (sometimes with OCR), or by simply keying in the text. The advantages of these public-domain databases are clear. In tandem with their ease of access, however, there are concerns over

their reliability and accuracy.

More and more academic libraries and large national repositories are converting their printed book collections into digital formats. These collections are in various stages of development. Of crucial importance is the need for these institutions to have a centralized mechanism so that they can record information such as which items have been or are about to be digitized, where and how they can be accessed, and what technical standards and specifications have been followed. The Digital Library Federation (DLF, diglib.org) has taken the initiative to create a registry of digital formats that all participating institutions may contribute to and use. National libraries and archives, bibliographical networks, and major universities in the U.S., Canada, and the United Kingdom have shown some early interest in this project.

In 2000, the Bibliothèque nationale de France (BNF) issued the largest single library available online. The BNF made available to the public more than 35,000 volumes from the Middle Ages through the early twentieth century on its Gallica Web site (gallica.bnf.fr). Multiple search criteria are available in Gallica: author, title words, date, subject, and the full-text of the catalog entries. Researchers visiting the Web site can access a wide range of retrospective texts in the areas of Jewish history, language, religion, and society including works by Julius Furst, Jacob Guttman, Elijah Levita, Johannes Buxtorf, Josephus, Maurice Vernes, Bartolocci, and many other texts from the sixteenth to the nineteenth centuries.

The Bibliotheca Rosenthalia is digitizing its holdings of works by Menasseh ben Israel (cf.uba.uva.nl/en/collections/rosenthaliana/menasseh). The electronic version of the Menasseh ben Israel

collection consists of more than 11,000 digital images. Users can navigate easily through the collection and the individual editions. Brief descriptions of the books are provided, and each entry has a link to the full text of the description with annotations and to the pages of the book itself.

The Jewish National and University Library (JNUL) in Jerusalem is planning a major project to scan out-of-print books to enhance preservation and security in house, as well as to provide outside access. The project will begin with 5,000 rare titles.

Consortia and collaborations have also been formed to promote the production of electronic books. As a response to the relatively small number of titles available electronically in the humanities and social sciences, the American Council of Learned Societies (ACLS, www.historyebook.org) initiated the History E-book Project. This project aims to offer both new titles in electronic format (eighty-five history titles through 2004), and also digitally convert retrospective history titles. These titles are searchable by author, title, and subject, as well as full-text. The number of Jewish studies titles available at this Web site is very limited.

The California Digital Library (CDL, www.cdlib.org) is a collaborative effort of the ten campuses of the University of California. Among its offerings is the eScholarship Editions (texts.cdlib.org/escholarship), which provides a growing collection of electronic editions of academic monographs, a significant number of which are freely available to the public. Current offerings in Jewish studies number about fifty to sixty titles. eScholarship uses Extensible Markup Language (XML) and Java-based technologies to support its

electronic editions. What this means is that the structure and meaning of a document can be specified by tags that enable sophisticated searching and displays. CDL uses tags defined by the Text Encoding Initiative (TEI, www.tei-c.org), an international project set up "to develop guidelines for the preparation and interchange of electronic texts."

Increasingly over the past three or four years, printed texts have begun to be packaged and offered commercially as electronic books. Commercial vendors such as eBrary, netLibrary, and Questia accompany their texts with an array of research tools including dictionaries, encyclopedias, and other features. There is not, as of yet, a single accepted standard for commercially produced electronic monographs that serve both publishers' needs

and those of the end-users. These texts are usually "published" in a proprietary format that may require reformatting or scanning of the print version to conform to the requirements of a particular vendor's system. Texts found in eBrary are offered in a proprietary PDF format, allowing for the original look and layout of the book to be maintained. The texts are acquired directly from publishers and sold to libraries rather than to individual users. Its list of publishers includes many of the leading academic presses. A simple subject search in eBrary reveals 56 titles under Judaism, 65 under Jews, 69 under Jewish, 3 under Hebrew, and 5 under Yiddish.

More and more institutions, commercial publishers, consortia, and individuals are taking the initiative to move printed books

into the digital environment. The number of titles of importance to Jewish studies is limited and narrow in scope. It is imperative that Jewish studies scholars including publishers, academics, and librarians join together to coordinate the planning and development of standards for selecting, accessing, describing, and archiving their important books in an electronic environment.

*Heidi G. Lerner is the Hebraica/Judaica Cataloger at Stanford University Libraries.*