Predecessors of Scientific Indexing Structures in the Domain of Religion

Bella Hass Weinberg

Abstract

Many scientific indexing structures thought to have been developed in the computer era were invented about a millennium earlier, in the domain of religion. Hans Wellisch traced the origins of alphabetical indexing to a fourth-century index to a compilation of sayings of the Greek Church fathers.

The Masoretes standardized the text of the Hebrew Bible and in the course of their work created alphabetical lists of words and phrases. These tenth-century lists may be viewed as the predecessors of thirteenth-century Latin biblical concordances. Structurally, Masoretic lists and biblical concordances are analogous to KWOC (keyword out of context) indexes. Masoretic lists anticipated search features, including truncation and adjacency, developed in the latter half of the twentieth century. The Masoretes also created permuted indexes and produced frequency counts of biblical words. The first complete Hebrew biblical concordance, compiled in the fifteenth century, contains a "stoplist," or list of words not indexed, that is similar to contemporary English stoplists. The *Manipulus Florum*, a Latin collection of quotations compiled in 1306, contains a sophisticated network of cross-references.

Hebrew citation indexes covering religious texts date back to the end of the twelfth century. The first Latin biblical citation index, *Tabula Septem Custodiarum*, covers the commentaries of the church fathers on the Bible. The appended bibliographies, arranged alphabetically by author and title, resemble the source index of *Science Citation Index*.

The call for papers for this conference stated that emphasis should be on the period from World War II on, the era when scientific information systems are believed to have been developed. Herbert Ohlman says that "ideas leading to mechanical and automatic indexing go back 150 years. However, not until World War II, when appropriate tools became available, did these ideas become reality" (1999, p. 187).

In this paper I demonstrate that many scientific in-

dexing structures thought to have originated in the computer era were invented as much as a millennium earlier, in the domain of religion. A. Taylor (1966, p. 31) observed that "theology was . . . the queen of sciences" in the Middle Ages; this phrase recurs in the introduction to a book of medieval Jewish philosophy (Saddiq, 2003, p. 29); the editor traces the concept back to Philo (Wolfson, 1947, vol. 1, pp. 156-157; J. Haberman, personal communication, 9 December 2002). I have demonstrated that a higher percentage of early theological works than medical works had indexes (Weinberg, 1999, p. 115). This paper surveys alphabetical index structures and citation indexes developed in the field of religion, with an emphasis on ancient texts indexed in the medieval period, and relates them to contemporary indexes and electronic search capabilities in the field of science.

The Origins of Alphabetical Subject Indexing

Alphabetical order is found in the Bible (Weinberg, 1999, p. 114). Hans Wellisch (1994b) traced the origins of alphabetical indexing to a fourth-century compilation of sayings of the Greek Church fathers titled *Apothegmata*. Not having seen this ancient work, I cannot say whether it was a word index or a true subject index; the latter would include words for concepts not expressed in the sayings.

Roberto Busa (1971) surveyed the history of *concordances*—another term for word indexes. He cites an earlier encyclopedia article by Henri Leclerq (1948) that describes a rudimentary Greek concordance of geographic names, not in alphabetical order, that was produced in the sixth or seventh century. Busa says that "Biblical concordances were very probably in existence in the seventh and eighth centuries" (p. 595).

My recent research (Weinberg, 2001) has focused on the Masoretes, who began working in the eighth century to standardize the text of the Hebrew Bible (Even-Shoshan, 1982, vol. 1, p. [xvi]). The Hebrew word Mesorah means "tradition," and the Hebrew term for Masoretes is Ba'ale Mesorah (masters of the tradition). In the course of their work, which spanned several centuries, the Masoretes created many alphabetical lists of biblical words and phrases. These may be viewed as selective concordances. Some of the lists, which date from the tenth century, include parts of the verses in which these terms are found; in other words the keyword was placed next to its context. Using traditional indexing terminology, we may say that the keyword was the "heading," and the biblical verse the "locator" (Borko & Bernier, 1978, p. 11). At the time the Masoretes worked, there were no chapter or verse numbers in the Bible, but learned Jews knew the entire text by heart (Wellisch, 1985, p. 70), and part of a verse served as the locator for a single word.

Chapter numbers were assigned to the Latin Bible by Christians, an action that paved the way for the compilation of complete concordances in the Latin language in the thirteenth century. By 1247 an alphabetical concordance was developed under the direction of Hugo of St. Cher (Rouse & Rouse, 1990, p. 219). There were two elements to an index entry in this concordance (Weinberg, 2000, p. 5): the heading (the biblical word) and the locator (the chapter number plus a letter indicating relative location within the chapter). Richard and Mary Rouse (1979, p. 10) note a disadvantage of this concordance: "It did not cite words in context." In subsequent editions a third element was added to each entry: a subheading (the phrase surrounding the word).

In terms of format thirteenth-century Latin biblical concordances with subheadings may be compared with modern KWOC (keyword out of context) indexes, in which entries consist of a word extracted from a title, the full title, or the part of the title surrounding the extracted word and a number leading to the bibliographic data.

The Christians compiled the first full biblical concordance, but at least three centuries earlier the Masoretes produced an amazing variety of lists that anticipated many of the index structures and search capabilities commonly thought to have been developed in the latter half of the twentieth century. Ohlman (1999, p. 189) states that "the earliest reference to permutation techniques was The Art of Making Catalogs of Libraries," written by A. Crestadoro in 1856. Nearly a millennium earlier, however, the Masoretes created lists of words in permuted order: for example, tsedek u-mishpat (righteousness and justice) / mishpat tsedek ([a] judgment [of] righteousness); shoftim ve-shotrim (judges and officers) / ve-shotrim ve-shoftim (and officers and judges). Thus the Masoretes anticipated the Permuterm subject index invented by Eugene Garfield (1976) and published with Science Citation Index in the 1970s. The difference between the old and new index structures is that the Masoretes recorded permutations actually found in the Bible, while the computer-produced Permuterm index shows all possible combinations of two words in titles.

Masoretic lists anticipated search capabilities developed by Lockheed and SDC (System Development Corporation) in the latter half of the twentieth century. The Masoretes noted the occurrences of significant twoword expressions, such as *tel 'olam* (everlasting ruin), long before "adjacency" searches were possible on-line.

The Masoretes produced lists of words beginning and ending in certain letters. While their purpose was orthographic accuracy, these lists anticipate the search capability of "truncation." A common example of left truncation in the sciences is a search on all words ending in *mycin*, to retrieve a class of antibiotics. Because Hebrew is written from right to left, the Masoretes were actually doing right truncation.

Trudi Bellardo Hahn (1998) has shown that many of the capabilities of Internet search engines were developed several decades prior to the creation of this network of networks, in the context of on-line searching. Yet in this paper I have shown that the roots of these search capabilities are evident in Masoretic lists compiled a thousand years earlier.

Masoretic notes on the Bible often include frequency counts, which one might think could have been produced only with the aid of a computer. Keep in mind the scope of the Masoretic frequency counts: they cover all occurrences of a word or phrase in the twenty-four books of the Hebrew Bible. Scholars I have consulted are divided as to whether the Masoretes had the entire Bible memorized or used an indexing system (Weinberg, 2001, pp. 183–185).

Stoplists

One scholar, Jordan Penkower, believes that the Masoretes must have had an indexing system because their lists include function words, such as prepositions, which are often placed on "stoplists" (lists of words not to be



Figure 1. The stoplist of the first Hebrew biblical concordance, Me'ir Nativ, compiled in the 1400s and first printed in Venice in 1523. Courtesy Library of the YIVO Institute for Jewish Research, Rare Book Collection.

indexed). How could anyone remember all the occurrences of the word *on*, for example, in twenty-four books?

The Masoretes did not exclude function words because every word in the Bible counted, and the standardizers of the text wanted to be sure that even non-content words were copied correctly. They prepared lists of the occurrences of similar-sounding function words, such as *el* (to) and *'al* (on). The invention of the stoplist is credited to Hans Peter Luhn (1959), in conjunction with the development of the KWIC (keyword in context) index. Eugene Garfield (personal communication, 24 December 2001) informed me that Ohlman was the codiscoverer of KWIC indexing, but Ohlman's paper (1999) on this topic does not discuss stoplists.

Barbara Flood (1999) described the stoplist used at Biological Abstracts, developed shortly after Luhn's articulation of the concept of not indexing function words. Flood noted that Luhn did not use the term *stoplist* in his paper on KWIC indexing (Luhn, 1959). He simply wrote that "the following 16 words are disregarded" (p. 231). Flood believes that Biological Abstracts may have been the first organization to use the term *stoplist*.

Learned Jews knew the Hebrew Bible by heart. Therefore, according to Hans Wellisch (1985, p. 70), Jews did not need concordances. But a Hebrew concordance was developed in the 1400s for the purpose of theological debates with the Christians, and the author, Mordecai Nathan, used Christian chapter numbers as locators, even though a different division of the text had been established by Jewish scholars a millennium earlier.

This first Hebrew concordance, called *Me'ir Nativ* (Lights the path), was not printed until 1523. It includes a one-page stoplist (Figure 1). The types of words on the list—prepositions and conjunctions—are very similar to those found in contemporary stoplists. Amazingly, the author uses the Hebrew equivalent of the term *function words* to characterize the stopwords. The author admits to having modeled his Hebrew concordance on a Latin one, but from the wording of the passage on the stoplist in the (unpaginated) introduction, I infer that this stoplist was an original compilation. The author wanted to complete the concordance in a reasonable time, and so he omitted frequent nonsubstantive words.

Function words were not headings in the first Latin concordances, but lists of the omitted words were not provided. In the fifteenth century the meanings of some of these words became central to Christian theological debates. To help locate all their occurrences, a supplementary concordance of them was created (Walker, 1894, p. 9). The words were characterized as "indeclinables," that is, words other than nouns—which are declined in Latin—and verbs, which are conjugated. Among the words on a printed list of indeclinables, titled "Index partium indeclinabilium" (*Concordantiae*, 1531, [folio A1 verso]), is *amen*.

Concept Indexes and Vocabulary Control

Up to this point we have discussed word indexes, or indexes in which the headings are extracted from the text, and no links between synonyms and related terms are provided, except for grouping morphological variants of a word. The first sophisticated subject indexes were, as with word indexes, produced in the domain of religion. Rouse and Rouse (1979) devoted an entire book to a description of the *Manipulus Florum*, an alphabetically arranged reference tool compiled in 1306. This work falls into the genre of "florilegia," or collections of quotations from various authors. The *Manipulus Florum* draws on the writings of the church fathers. The excerpts are arranged alphabetically under topical headings. Each entry is assigned a locator—one or two letters, written in the margins.

In a volume of *Chemical Abstracts* produced in 1947, I noted a similar system of locators, with the letters running down the space between the two columns (Crane, 1947). Like the *a*-*f* locators in the first Latin concordance, which represent relative position in a biblical chapter, the letters *a*-*i* in *Chemical Abstracts* mark relative position on a printed page. In contrast, the locators in the *Manipulus Florum* are linked to entries, not to their position on a manuscript page. *Chemical Abstracts* started numbering entries in later volumes but retained letters as suffixes (for check digits). The editors could have borrowed the idea of sequential numbering of entries from a religious manuscript produced in the fourteenth century.

Because the entries in the *Manipulus Florum* are arranged alphabetically, it is a self-indexing work. Since there is no auxiliary index, why were such precise locators necessary? The Rouses (1979, p. 120) believe that the locators were provided solely for the purpose of cross-referencing: if a quotation was germane to two topical headings, it was given in full under one, with a cross-reference from the other heading to the precise locator of the entry under the first heading.

The *Manipulus Florum* had a better system of crossreferencing between entries than was provided in *Engineering Index* in the mid-1970s: in *Engineering Index* if an abstract was relevant to two subject headings, it was placed under one, with a cross-reference from the second (e.g., "GRAPHIC METHODS See Also AIR POLLUTION—Philadelphia, Pa.," Engineering Index, 1975, p. 232). There is no semantic relationship between these headings. In earlier work I pointed out the unusual function of "see also" references in *Engineering Index* (Weinberg, 1982, p. 25).

Besides sophisticated cross-references between entries the *Manipulus Florum* has a good syndetic (linking) structure for topical headings: "see" references link synonyms, and "see also" references link related terms. This 1306 manuscript does not mark the invention of cross-references, however. At the Vatican Library I found "see" references in a Hebrew dictionary of philosophical terms written in the twelfth century, and I believe that cross-references in narrative texts occurred earlier, for the purpose of saving space.

Wellisch (1994a), in describing the first printed indexes, noted that a graphic display of term relationships was appended to a work of St. Augustine. The thesaurus is believed to have been invented in the middle of the twentieth century, and the first graphic display is considered to be that of the *Euratom Thesaurus;* a page from the second edition, published in 1967, is reproduced in the NISO thesaurus standard (NISO, 1994, p. 47). Those who create visualization interfaces of semantic relationships probably do not know that there was a print predecessor thirty-five years earlier, and they are surely unaware that there were graphic displays of terms in "incunabula" (books printed before the year 1500).

Citation Indexes

The history of alphabetical indexes may be unfamiliar to many of the attendees at this conference, but some participants may have seen my paper on "The Earliest Hebrew Citation Indexes" (Weinberg, 1997). Citation indexes follow subject indexes in this paper because the latter are older. I have been able to trace citation indexes back to the twelfth century, while word indexing began eight centuries earlier. In the following paragraphs I summarize my earlier work on citation indexes and then present subsequent unpublished research.

Science Citation Index, a tool that shows which older documents are cited in newer journal articles, was invented by Eugene Garfield. (I recently explained the concept of citation indexing to my sixteen-year-old daughter, Kira, who was tracing a theme for a Jewish Studies paper. After thirty seconds of explanation she said, "I see; it's a backwards index!"). Garfield credits the idea to Shepard's Citations, a legal reference work begun in 1873 that indicates whether older cited cases have been overruled or upheld by later courts. In 1992 Fred Shapiro described a legal citation index titled Raymond's Reports that was published in England in 1743. Knowing that major Hebrew citation indexes had been published in the 1500s, I decided to do a systematic study of the earliest printed and manuscript Hebrew citation indexes.

The first manuscript Hebrew citation index documented in the literature, *Mafteah ha-Derashot* (Index of homiletics), is arranged alphabetically by biblical phrase; it shows where the phrase is cited in a wide variety of Rabbinic sources. The manuscript is ascribed by Yehuda Leib Maimon to Maimonides, a physician who lived from 1135 to 1204. Maimonides is said to have compiled this citation index in his youth. Menahem Kasher challenges the authorship of the index for two reasons: Maimonides was a genius and did not need indexes, and had Maimonides compiled a biblical citation index, it would not have had such an illogical arrangement (the logical method of arranging such a citation index is canonically, i.e., in the order of the text). And indeed all other citation indexes to the Bible and Talmud are in canonical order, despite the absence of standard locators for either work prior to the 1500s, as well as a lack of standardization of the order of the books of the Bible and the chapters of Talmudic tractates. If *Mafteah ha-Derashot* was the first citation index, however, its compiler would not have had any models to follow, and Kasher's criticism of its arrangement is inappropriate.

Biblical citation indexes served both for the preparation of sermons and for questions of law. Talmudic citation indexes were primarily legal in nature, showing where a Talmudic passage was cited in codes of Jewish law. In an edition of the Talmudic citation index 'En Mishpat (Fount of justice) that was published in 1714, a symbol was included to indicate that the decision in the Talmud was overruled by a subsequent code of Jewish law (Weinberg, 1997, p. 230). This notation preceded the equivalent code in Shepard's Citations.

My 1997 paper on Hebrew citation indexes was based on research done in libraries in New York City. In subsequent research on Hebrew citation indexes that I conducted in libraries abroad, Maimonides figured prominently. At the Vatican Library I found a biblical citation index appended to the Hebrew translation (from Arabic) of his philosophical work *Guide for the Perplexed*, copied in 1205 (catalog no. Vat. Ebr. 263). At the Institute of Microfilmed Hebrew Manuscripts of the Jewish National and University Library in Jerusalem, I identified fragments of citation indexes that lead from the Bible to the Talmud (compiled between the third and fifth centuries of the common era); to a legal code by Rabbi Yitzhak Alfasi (1013–1103), called the Rif; and then to the code of Maimonides.

Ezra Chwat, a scholar at the institute, called to my attention manuscripts that might be citation indexes, including the one just described. Some of the manuscripts are just lists of references, however; they do not go forward in time. Maimonides's legal code was controversial because it lacked references, just as a scientific paper without references to the prior literature would generally be unacceptable. "Vertical legitimation" is an important concept in Jewish law (Weinreich, 1980, p. 207): earlier authoritative sources must be cited to justify a ruling.

The Institute of Microfilmed Hebrew Manuscripts

holds photostats of fragments of source notes for Maimonides's legal code, written by his son's father-in-law, Rabbi Hananel ben Shmuel. This documentation activity may be compared with that of a scientist writing a paper who sends his or her graduate assistant to do a literature search so that references can be appended to the paper. But the analogy is not precise: Rabbi Hananel no doubt provided the source notes voluntarily, to lessen the controversiality of his son-in-law's *magnum opus*, which remains authoritative today.

In the summer of 2001 I traveled to England to examine the first Latin biblical citation index, called Tabula Septem Custodiarum (Index of the seven custodies). This work was described by the historians Richard and Mary Rouse (1991b), who note that all copies of the manuscript are in England. They suggest (1979, pp. 18-19) that compilation of the Tabula may have begun in the thirteenth century under the direction of Adam Marsh, who died in 1259, but all the manuscripts of this work described in one of their later books (1991b, pp. xcviii-xcix) date from the fourteenth and fifteenth centuries. The earliest dated manuscript of the Tabula is "Oxford, Bodleian Library, Bodley 685 (AD 1339)" (Rouse & Rouse, 1991b, p. xcix). That date is more than a hundred years after the death of Maimonides, and more than a century and a half after Maimonides's youth, the period in which he is said to have compiled a citation index. The Christians who compiled the Tabula may have gotten the idea for a biblical citation index from Hebrew manuscripts, just as the idea for the first Latin biblical concordance probably emanated from the study of Hebrew Masoretic Bibles. The fact that both types of Latin reference work were produced in the midthirteenth century lends support to this theory.

During my brief stay in England I could not examine all the extant copies of the *Tabula*, but among the ones I saw were several manuscripts of extraordinary beauty: one written in fancy calligraphy, held by the Bodleian Library (Oxford, New College, no. 315), and another decorated in color, in the library of Balliol College, Oxford University (nos. 216–217). This recalls a point made by the world-renowned authority on Hebrew manuscripts, Malachi Beit-Arie (1993, p. 81): Latin manuscripts were almost always produced under institutional auspices, while Hebrew ones were written by individuals.

The focus here is not on the aesthetics of indexes but rather on their structure. The earliest Latin biblical citation index was structurally similar to its Hebrew counterparts that were arranged in canonical order. As Rouse and Rouse put it, "For theologians, the order of topics was defined by the order of the Scriptures" (1991a, p. 201). The main difference between the first citation indexes in the two languages lies in the nature of the works covered. Hebrew citation indexes analyzed the biblical references in anonymous classics, such as the Talmud, Midrash (legends), and Zohar (a Kabbalistic work), while the *Tabula* covered the works of the church fathers, that is, personally authored works (see Figure 2).

Because only a few works were covered by most Hebrew biblical citation indexes produced in the Middle Ages, the works indexed were often listed on the title page of a manuscript or printed book. But the list of works covered by the *Tabula* was more extensive, and several of the manuscripts include a bibliography that is analogous to the source index of *Science Citation Index*. The same goes for many copies of the *Manipulus Florum*, discussed above: a bibliography of the works analyzed for the florilegium is appended to the text, although the bibliography is not arranged alphabetically by author.

The manuscript of the *Tabula* that is illustrated in Figure 2 is bound together with a biblical concordance. This conjoining demonstrates that the two reference works were considered companions: one could look up a word in the concordance and find all its occurrences; then one could go to the citation index and locate commentaries on the relevant passages.

A citation index is maintained on cards at the Western Manuscripts Division of the Bodleian Library of Oxford University. Arranged by collection name and manuscript number, the index leads to modern works in which the manuscript is cited or described. This card index may be compared with the section of *Science Citation Index* that is arranged by patent number. (The patent index of *Chemical Abstracts* is not a citation index.) In using the print or electronic version of *Science Citation Index*, it is worth recalling that this index structure is more than eight hundred years old.

Citation Analysis and the Study of the Earliest Indexes

Most of the early indexes discussed in this paper are described in books and journals in the field of religious studies, with no links to the information science literature. Eugene Garfield (personal communication, 23 May 2002) called to my attention a recent reference that reinforces this point: a book review of a Latin citation index, published in *Zeitschrift für Kirchengeschichte* (Journal

Solor non om Butan quepmit à sab fareme at the a fa for ingusta to the court of to gue our personaly pun to prese que our pressant services net mp 20 Jopao gennegethoman + fice eno alenter later Gor cypet augo alt 10 For Thet wit wer or Dechincana thigh Section ato white BB St Strangito AmonArata ymagine 104 fortingen Sno in the policymet in most + of 10 28 10 moz 28 c Gung Papilo of type ony 18 708:28 antono Sennt lagião augulaje como The mon a germant water ad fo bern for two lig - e w for 10 an fort fely abe into moner to? milgap + c? Juob genne Toppe X8 Gung orpor ango affinnt 400 Fange log is penpe it do in Boc el Doktmer anyomman ampet. solie Jopph Eng mano 28 Gungarpo ango Se muse + gentra per 7 e pres gracon Te pentes & connexione + of perange fange 6027 filmo to 30 apellate + mp asse more for Onos meanothigh war tong Som of na Water Mayam annage t. Gor ango 6 8 On voer spr (ve Samafteg c. No 1 to t cypina Seponfanono filo te qt. Omes gentarones is + 8 6/12 ad spine Figure 2. A page from the Tabula Septem Custodiarum, the first Latin biblical citation index. Manuscript number Harley 3858 f.61. By permission of the British Library. (Rouse and Rouse [1991b, p. xcviii] give the date of this manuscript as "before 1395.") © 2004 by the American Society for Information Science and Technology

tor libraria Dunothy Aber soneraanno in spi 28 Brig cyper, bale of Sur Augo of Faft Bor in and sinon Ac Man te of Gango So inte Dar ho 15 c 23 ibi pegy in fingt matter + cope as grig

mo agarhene apit al alaga & comanit 19 m. 2 drather capit de alcaya + ennant as popet 1 e of ange se go goth of gone ffile so file abasha bor sope ses amb to en a ors sates al abasham of as espoe le best le for the o e c 6. S25" notand graphoring clain 30 pfert as bet e off to one mare le 1 e v a is t apre o clain of some boto a c yount 12109 80 Se mangiom by piliozipt unprea ile

These goment place 27 a Setta - Port and as for lies of lace of a Om fint mos Talamon gomme Goog Se part Gor amb Regito Por of Pue of Sof Ilud On Emt salamon

Sund we gennet onlamone excague funt Vie Fambio mott allego Rog Polure & fis Mas Om Fuess Pat Guus export ang Se maris sume Proper to po o in Cin non assite 10000 mano to

12 for and me construction of the one pre Calmoh Low another mint gonacones + Sit al alige comajo com anaturecom to? bein/8

Tweet Sepontata mate ign reath Bugo R Scoat Bis pro ma for the off Bugo R Scoat Bis pro ma Gerbi Of & Sepontacous not to ge BA Or muto of ous 20 a the panio: E Sca stanta unifus eft and 10 ge Buge Se meterieta Thigh menycut uncuta cp m St Gais 200, 20 Bung & cope ango & mps Sago ftrute liof c & Juqua & quano ounday toma concepciones conforma concepciones conformation to +0. Timouta of this se and sta har and que or Summen in que Suplem Summe be consume star as seen low point fi-

Procono programmite? porceps a memo on cer mais als Bu onempro so carmer on i 2" f 16. Bu on car morno colme a of מושק שב שלים שיי והדיום לי ב איר שמוי Jun cet By mftuo trolut a Smit

The Put cam traduce Gor ango At IA 3 the the to Te Ste Dece Bit Eng 10 Bi Diato Joppf memo wepet infino + nollat at take no ar Di sam gua con Do De allogin Sam non so fra Bat 1am velue confegment exitemule a Bulton the Ret ann traduct 1. Sr

Face angelno 86 28 frug express anges Scopen & and c. 18 1B1 Sum & Bone open

hor and as the one for to 2. 5 1. IT Beilin gingern than greang The G TT 18 Alienationat protos a pot of vor 181

under tornono pr 20 Commy out

of church history). Besides the fact that scholars tend to read only the literature of their own discipline, the use of different terms for index structures in the fields of religious studies and information science impedes interdisciplinarity.

The majority of the journals in the category of information science and library science were transferred from *Science Citation Index* to *Social Sciences Citation Index* in the 1980s. Perhaps a shift of this category to *Arts and Humanities Citation Index* should be considered because the roots of scholarly information systems are in the domain of religion.

Conclusion

I have explored several facets of the history of indexes in languages with which I am familiar. It may well be that earlier exemplars are in languages I have not mastered, such as Arabic and Chinese; in any case I hope I have convinced readers that the indexing structures of scientific information systems were not invented after World War II but rather many centuries earlier.

Acknowledgments

The research described in this paper was supported by the Eugene Garfield Foundation, the Jesselson Foundation, the H. W. Wilson Foundation, and St. John's University. My husband, Gerard Weinberg, M.D., assisted me on research trips to the Vatican Library and various manuscript repositories in France, England, and Israel. The paper was presented and edited for publication during a research leave from St. John's University. Pennie Bjork word processed several versions of the paper, and Liliane Blattner assisted with the production of the transparencies for the conference.

References

- Beit-Arie, M. (1993). Hebrew manuscripts of East and West: Towards a comparative codicology (The Panizzi Lectures, 1992). London: British Library.
- Borko, H., & Bernier, C. L. (1978). *Indexing concepts and methods*. New York: Academic Press.
- Busa, R. (1971). Concordances. In *Encyclopedia of library and information science* (vol. 5, pp. 592–604). New York: Marcel Dekker.
- Concordantiae maiores Sacrae Bibliae svmmis vigiliis (1531). Basileae: Froben.
- Crane, E. J. (Ed.). (1947). *Chemical Abstracts* (Vol. 41). Columbus: Ohio State University.
- Engineering Index. (1975). *The Engineering Index monthly and author index* (Vol. 13[11]). New York: Engineering Index.

- Even-Shoshan, A. (1982). A new concordance of the Bible (3rd ed., Vols. 1–3). Jerusalem: Kiryat Sefer. (In Hebrew.)
- Flood, B. (1999). Historical note: The start of a stop list at Biological Abstracts. *Journal of the American Society for Information Science*, 50(12), 1066.
- Garfield, E. (1976). The Permuterm subject index: An autobiographical review. *Journal of the American Society for Information Science*, 27(5/6), 288–291.
- Hahn, T. B. (1998). Text retrieval online: Historical perspective on Web search engines. *Bulletin of the American Society for Information Science*, 24(4), 7–10. Also available: http://www.asis.org/Bulletin/Apr-98/hahn.html (accessed 21 June 2002).
- Leclerq, H. (1948). Concordance. In F. Cabrol & H. Leclerq (Eds.), *Dictionnaire d'archéologie chrétienne et de liturgie* [Dictionary of Christian archaeology and of liturgy] (Vol. 3, part 2, cols. 2,493–2,494). Paris: Librairie Letouzey et Ané. (Busa [1971] gives the date as 1914.)
- Luhn, H. P. (1959). Keyword-in-context index for technical literature (KWIC index). (Technical report RC-127, IBM Corporation, Advanced Systems Development Division). Reprinted in C. K. Schultz (Ed.). (1968). *H. P. Luhn: Pioneer of information science: Selected works* (pp. 227– 235). New York: Spartan Books.
- Nathan, M. (1523). Me'ir nativ: hanikra Konkordanses [Lights the path, which is called Concordance]. Venice: printed by Daniel Bomberg.
- NISO (National Information Standards Organization). (1994). Guidelines for the construction, format, and management of monolingual thesauri: An American National Standard (ANSI/NISO Z39.19-1993 [R1998]). Bethesda, MD: NISO Press.
- Ohlman, H. (1999). Mechanical indexing: A personal remembrance. In M. E. Bowden, T. B. Hahn, & R. V. Williams (Eds.), Proceedings of the 1998 Conference on the History and Heritage of Science Information Systems (pp. 187–192). Medford, NJ: Information Today.
- Rouse, R. H., & Rouse, M. A. (1979). Preachers, florilegia, and sermons: Studies on the "Manipulus florum" of Thomas of Ireland. Toronto: Pontifical Institute of Medieval Studies.
- Rouse, R. H., & Rouse, M. A. (1990). Concordances et index. In H. J. Martin & J. Vezin (Eds.), *Mise en page et mise en texte du livre manuscrit* [Page layout and arrangement of text in the manuscript book] (pp. 218–228). Paris: Editions du Cercle de la Librairie-Promodis.
- Rouse, M. A., & Rouse, R. H. (1991a). Statim invenire: Schools, preachers and new attitudes to the page. In Rouse & Rouse (Eds.), Authentic witnesses: Approaches to medieval texts and manuscripts (pp. 191–219). Notre Dame, IN: University of Notre Dame Press.
- Rouse, R. H., & Rouse, M. A. (1991b). The *Tabula Septem Custodiarum* and the *Registrum*. In Rouse & Rouse (Eds.), *Registrum Anglie de libris doctorum et auctorum veterum*

[Corpus of British medieval library catalogues, (Vol. 2)] (pp. xcvii–cxxvi). London: British Library, in association with the British Academy.

- Saddiq, J. ibn. (2003). *The microcosm of Joseph ibn Saddiq*. English translation, introduction, and notes by Jacob Haberman. Madison, NJ: Fairleigh Dickinson University Press; London: Associated University Presses.
- Shapiro, F. R. (1992). Origins of bibliometrics, citation indexing, and citation analysis: The neglected legal literature. *Journal of the American Society for Information Science*, 43(5), 337–339.
- Taylor, A. (1966). *General subject-indexes since 1548*. Philadelphia: University of Pennsylvania Press.
- Walker, J. B. R. (1894). The comprehensive concordance to the Holy Scriptures... Based on the authorized version, with an introduction by M. C. Hazard. Boston: Congregational Sunday-School and Publishing Society.
- Weinberg, B. H. (1982). Multiple sets of human indexing for civil engineering documents: Comparison of structure and occurrence rates in full text. Science and Technology Libraries, 2(3), 13–33.
- Weinberg, B. H. (1997). The earliest Hebrew citation indexes. Journal of the American Society for Information Science, 48(4), 318–330. Reprinted with a postscript in T. B.

Hahn & M. Buckland (Eds.). (1998). *Historical studies in information science* (pp. 51–64). Medford, NJ: published for the American Society for Information Science by Information Today.

- Weinberg, B. H. (1999). Indexes and religion: Reflections on research in the history of indexes. *The Indexer, 21*(3), 111–118.
- Weinberg, B. H. (2000). Book indexes in France: Medieval specimens and modern practices. *The Indexer*, 22(1), 2–13.
- Weinberg, B. H. (2001). Index structures in early Hebrew biblical word lists: Preludes to the first Latin concordances. *The Indexer*, 22(4), 178–186.
- Weinreich, M. (1980). *History of the Yiddish language* (Shlomo Noble, Trans.). Chicago: University of Chicago Press.
- Wellisch, H. H. (1985). Hebrew Bible concordances, with a biographical study of Solomon Mandelkern. *Jewish Book Annual*, 43, 56–91.
- Wellisch, H. (1994a). Incunabula indexes. *The Indexer, 19*(1), 3–12.
- Wellisch, H. (1994b). Indexing. In *Encyclopedia of library his*tory (pp. 268–270). New York: Garland.
- Wolfson, H. A. (1947). Plato: Foundations of religious philosophy in Judaism, Christianity, and Islam (Vols. 1–2). Cambridge, MA: Harvard University Press.