Snunith Shoham and Moshe Yitzhaki
Department of Information Science & Librarianship, Bar-Ilan University

The Impact of Cultural and Technological Changes on Titles Content and their Use in the Process of Information Retrieval

Abstract: The use of the title as a source of information about books in library catalogs and as a device for retrieval has undergone many changes over time. The paper touches on its historical development, described the impact of technological changes, and analyzed the impact of the computers on the informativity of titles. Also, the differences between Western and Eastern cultures regarding use of title is described.

1. Historical Perspective
   The use of the title as a source of information about books in library catalogs and as a device for retrieval has undergone many changes over time, and titles were also treated differently by different cultures and societies.
   
   The title in ancient times, was actually the book's opening words (Avrin, 1991). In many Babylonian libraries there was title information on each tablet together with other information such as the name of the owner (e.g., Amarna tablets of 1400 B.C.), name of the scribe (e.g., tablets excavated in the capital city of the Hittites, dated about 1300 B.C), the tablet number and the first words of the following tablets (e.g., in tablets excavated from the library/archive of Ashurbanipal, in the city of Nineveh, dated about 630 B.C.) (Strout, 1956).
   
   In large ancient libraries, such as that of Ashurbanipal, the clay tablets were kept in earthen jars and the jars were kept in orderly rows on shelves. Each tablet bore an identification tag, indicating the jar, shelf, and room in which it was to be found. On the walls of each room, beside the door was a list of works to be found in it; and something like a subject catalog or descriptive bibliography has been found on tablets, kept near the door of each room. These tablets include title of each work, number of tablet for each work, number of lines, opening words, important subdivision, and locational symbol (Johnson and Harris, 1966).
   
   Some lists of titles were written on walls of temples or pyramids, though the Egyptian books themselves have not survived. The Egyptians would sometimes write the title of the book, a summary of its contents, and, the name of the author on the reverse side or the cover of the scroll.
   
   In the literature found in Mesopotamian libraries, a colophon is often used in the last column. The word colophon derives from the Ionian city of that name (Glaister, 1979), but the colophon already flourished in ancient Mesopotamia and Egypt. The colophon gave information details about the books such as title (which consisted of the opening words), name of scribe, and his patron; seldom was the author named (Avrin, 1991).
   
   Throughout Greek literature there are references to book collections. However, most of our knowledge about the Greek's views of books is derived from the Pinakes compiled by Callimachus in the Alexandria library in about 250 B.C. The Pinakes, which is some sort of catalog or possibly a bibliographical work, is arranged according to incipits—the first word or words of the text. In it the number of lines in each work, the opening words, and some biographical data on the authors were recorded.
   
   The main purpose of the Greek colophon in the Hellenistic and Roman periods seems to have
been to establish the authenticity of the text copied. The counting of the lines originated as an assurance to the reader or purchaser of the scroll that the text was copied properly in its entirety, and it also determined the scribe's fee (Avrin, 1991).

The cataloging and bibliographical procedures of the period were began with a few general "subject" categories. The scholar, after checking the general subject, next looked for the author (title was secondary). Indeed, the concept of the author entry originated with the Greeks; it never appears in any works that have survived from earlier civilizations of the East. It is probably rooted in the democratic belief in the importance of the individual (Strout, 1956).

The evolution of books from papyrus surface to parchment and from roll to codex began in the first centuries and was not entirely established until the fourth century. In the European Middle Ages books took the form of the parchment codex, and were made by monasteries. The scribe frequently began the text with the "incipit" ("here began") and ended it with "explicit" ("explicitus," a tradition from the scroll period) or "finit". In manuscript books, a concluding statement, a colophon, indicated the title of the work, the name of the copyist, date and place of copying, a blessing for the patron or client, and threats of excommunication to unauthorized copiers.

Throughout the Middle Ages, cataloging was limited to an unorganized inventory list, probably representing shelf arrangement, which may have been done according to size or chronology. In fact these were merely lists of books, some arranged by author, some by title, and others by catchword from the title or first line, or sometimes a combination of all three. Some of these lists were originally kept on strips of parchment tacked to the side of the book chest; others were kept in codex form (Johnson and Harris, 1976).

Sometimes, among several works that were bound together only the title of the first work was mentioned. Most works were essentially anonymous, and the book bore the name of the scribe who copied it and not its author. Indeed, throughout the Middle Ages the ancient world's indifference to the proper naming of books and to their authors was sustained (Steinbrg, 1974). Many colophons did not contain the book's titles. Sometimes the name of the copied book appeared on the book's cover (De Vinne, 1972); and the copyist saw himself as exempt from repeating the title because the parchment and paper were expensive. Sometimes the opening words were written in red ink, or, beginning in the sixth century, capital letters rather than the book's regular letters were used to highlight the first words of the book (Riegler, 1995).

One of the earliest listings of a medieval library is dated in the eighth century. It consists of a list of brief titles, with authors added to a few of them (Strout, 1956). One of the first lists that can be designated as a catalog is from St. Martin's Priory at Dover (1389). It is divided into three sections, two of which are arranged by call numbers and give information on books including a short title. The third section is a catalog of analytical entries (which is a genuine innovation) alphabetically listed, some under author, some under title followed by author, some beginning with such words as "book," "part," or "codex".

The first printers followed the form of the manuscript book, and continued to use the colophon. The earliest known printed colophon is from 1457 in a book titled "Psalmorum codex" and published by Johannes Fust and Peter Schoffer of Mainz. (Clair, 1976). Peter Schoffer of Mainz was also the first to use a label title in Pope Pius's Bull against the Turks, which he printed in 1463. It consisted of two short lines:

"Bulla cruciata sanctissimi domini nostri Papa cotra turchos"
(cited by Clair, 1976, p. 115).

The first label titles took the form of a brief mention of title and author's name, placed at the
top of an otherwise blank protecting leaf at the beginning of a book. These early label titles probably served to prevent the first printed leaf from becoming soiled while lying about in the print shop before being given to the binder (Steinberg, 1974). However, printers soon realized the inherent potential of a full title. Title-page was first used by Erhard Ratdolt, of Augsburg and Venice, in his edition of Johannes Regiomontanus's astronomical and astrological calendar of 1476. He included a page that proceeded the calendar and gave some details about the book. It was set in a woodcut frame that clearly distinguished it from the text.

In 1500 Wolfgang Stockel, in Leipzig, was the first to issue a book with complete title page, giving subject-title, name of publisher, name of printer, date and place of printing (Clair, 1976). For producers, the title page offered more than the technical advantage of a protective cover; it was also a cheap and effective means of advertising the book. Indeed, the first books to use a title page were all new publications that needed some introduction to the public (Steinberg, 1974).

Title pages became increasingly common, by 1500 the title page had established itself; thus facilitating the production of book lists and catalogs while acting as advertisements in themselves (Eisenstein, 1983), and they were decorated and used extensive wording.

During the Renaissance, works again came to be identified with their authors (a concept which has originated with the Greeks). Titles, therefore, were neglected in the single-entry catalogs. The catalog or index was arranged according to author, often using the Christian name rather than the surname. The title was recorded in the catalog, but not as an access point.

In a bibliography arranged in chronological order by the German bibliographer and librarian, Johann Tritheim in the late fifteenth century, an alphabetical author index was included (Strout, 1956). A classified catalog of Syon Monastery, Isleworth, England, from the early 16th century also included an alphabetical author index.

Beginning in the mid-16th century, bibliographers rather than librarians took the lead in improving the catalog. Florian Trefler published at Augsburg a treatise on the keeping of a library (1560). He advocated a five-part catalog consisting of alphabetical author catalog, a shelf list, a classified index to analytics, an alphabetical index to the classified index, and a list of books that were not kept with the main collection (Strout, 1956), although he did not include a title catalog.

In 1595 Andrew Maunsell, an English bookseller, wrote rules for entry in the preface to his catalog of English printed books. He used surname rather than Christian name. He placed anonymous works either by title or else by the subject matter, and he was actually the first to set up the principle of uniform title - he established uniform entry for the Bible. This marked the first recognition that the title can also be used for retrieval (Strout, 1956).

In the 17th and 18th centuries, the concept of the catalog changed. The catalog became a means of finding books rather than inventory. The accepted retrieval entry was that for the author, and subject catalogs provided an author index. This was the case with the catalog of Oxford University, which was prepared under the supervision of the library director, Sir Thomas Bodley early in the 17th century.

The title was usually not regarded as a retrieval device, but it was recorded in the catalog records. The Frenchman Frederic Rostgard asserted in regard to catalogs that the word order of the titles as found on the title page should be preserved.

In the French government's code of 1791, libraries were directed to use card catalogs (because of the wartime shortage). The title page had to be transcribed on the card, and the author's surname underlined as a filing word. If there was no author, the key word in the title had to be underlined (Strout, 1956)

It is interesting that if during the 17th and 18th centuries, an anonymous work did not have a title, a word or words from the document's subject were used (Domanovsky, 1974). That is - there was no clear distinction between an author-title catalog and a subject catalog.
The 19th century was a period of much argument over the relative merits of classified and dictionary catalogs. The codes of that century emphasize alphabetical arrangement of the catalog, and usually the author is the main (if not the only) entry. Panizzi, in his 91 rules (British Museum, 1841), preferred, when there was no author, to use the name or the event or person that the book was about (Rule XXXIII), or the name of any society or body, or place (Rule XXXIV), or of the editor or even the translator (Rule XXXV), only in cases where none of these rules could be applied, he suggested the use of the first substantive word in the title (Rule XXXVIII). He also recommended the use of a uniform title for the Bible (Rule LXXXIX), acts, memoirs, journals (Rule LXXX), almanacs, and calendars (Rule LXXXIII).

Jewett (1853) referred to title entry in Rule XXIII of his code, according to which translations of works of unknown authorship were to be entered, like other anonymous works, under the first word of the original title.

Cutter (1891) used the title entry in anonymous works (Rule 68) and in all works of prose fiction (Rule 75). In cases where the author's name was known, he suggested a cross-reference from the title to the author (Rule 70). For the first time, a cataloging code stipulated that the title always be used as a retrieval access.

In the German and Anglo-American codes of the early 20th century, the use of uniform titles (together with the use of uniform names) became compulsory (Domanovsky, 1974). This constituted an important step in the use of titles as a tool for retrieval.

In the mid-20th century, the informativity of documents' titles began to increase (Buxton and Meadows, 1977; Yitzhaki, 1994). This was manifested in two ways: (a) the replacement of trivial or literary words by substantive and more meaningful ones, and (b) an increase in the length of titles, thus including a larger number of substantive words. Many studies have dealt with this issue; most, however, focused on journal articles' titles and only a few on book titles. This may be partly explained by the fact that whereas from ancient times up to the late 19th century the book was the main source of knowledge, since then the journal article has become the most important channel for scholarly communication.

2. An Increase in Title Informativity

In the mid-20th century, the informativity of documents' titles began to increase (Buxton and Meadows, 1977; Yitzhaki, 1994). This was manifested in two ways: (a) the replacement of trivial or literary words by substantive and more meaningful ones, and (b) an increase in the length of titles, thus including a larger number of substantive words. Many studies have dealt with this issue; most, however, focused on journal articles' titles and only a few on book titles. This may be partly explained by the fact that whereas from ancient times up to the late 19th century the book was the main source of knowledge, since then the journal article has become the most important channel for scholarly communication.

Some studies left it to persons to judge the informativity of titles (e.g., Peritz, 1984). This subjective approach is based on classifying the titles checked into two distinctive categories: informative and noninformative. A title is considered informative if it conveys at least some general idea of the paper's context. Other scholars relied on counting the number of substantive words in the title (e.g., Tocatlian, 1970; Buxton and Meadows, 1977; Yitzhaki, 1992, 1994, 1995).
Diodato (1982) examined the occurrence of title words in the abstracts, first and last paragraphs, and cited titles of research papers in chemistry, economy, history, mathematics, and philosophy in the 1960s and 1970s. The best reflection of title-word occurrence was in the abstracts, followed by first paragraphs, last paragraphs, and cited titles, respectively. Longer-than-average titles demonstrated a higher frequency of title word occurrence in first and last paragraphs than did titles in general. These findings strengthen the assumption that longer titles have a larger number of useful words for document retrieval than do shorter titles.

Montgomery and Swanson (1962), Bottle (1970), Miller (1971), and Frost (1989), on the other hand, analyzed the matching of keywords or descriptors from the index entry against the title keywords of the corresponding documents. Ghosh (1974) employed a different approach, examining each title in a specific subject bibliography to see whether the word prostaglandin(s) or its synonymous terms existed in the title.

The increase in title-informativity findings probably indicates a growing awareness among editors and authors of the importance of titles in the process of information retrieval. This trend has progressed at different rates in various fields, generally speaking, the highest increases in title informativity were found in the sciences, with the social sciences and humanities showing lower rates (Yitzhaki, 1992).

The increasing number of substantive words per title over a wide range of subjects before 1960s may be have resulted from the increase in the number of articles published, thus easing the scanning of lists of titles in journals for selecting papers to read (Bird and Night, 1975). The key-word-in-context (KWIC) permuted-title index was introduced by Luhn in 1958 as a relatively inexpensive means of building a dissemination index whose task is to prompt notification to new material. Luhn pondered a temporary bridge between the content of the current literature and readers, awaiting the completion of the more slowly prepared conventional indexes (Luhn, 1960).

Although the trend toward more informative titles has preceded the introduction of KWIC and KWOC indexes (e.g., Buxton and Meadows, 1977; Yitzhaki, 1992), these indexes have undoubtedly contributed much to the growing awareness of the importance of informativity of titles. Indeed, one may associate the "wave" of increase in title informativity as observed for example by Yitzhaki (1992) between 1960 and 1970 to the KWIC and KWOC indexes, but also to the "current contents" that were initiated by Garfield in 1961 for chemical and physical sciences.

The further increase in title informativity that occurred between the 1970s and 1980s and then between the 1980s and the 1990s was probably influenced by the construction of online bibliographic databases and the introduction of OPAC (Online Public Access Catalog) into the libraries. With the advent of computers, the search for words from titles has become one of the main modes of subject search.

It is interesting to note that the increase in title informativity also contributed to the improvement of subject indexes. Studies of indexers have shown that they rely a great deal on titles for their subject analysis (Bottle and Feibish, 1970; Feinberg, 1973). The indexing in PRE-MED database, which provides access to key journals in clinical medicine, nursing, and hospital administration, is based on the title of the article as it appears in the table of contents of the journal (Piternick, 1985). Thus Piternick suggests using terms taken from titles in addition to other subject terms. He especially stresses the usefulness of terms taken from titles in new subject areas, for which generally accepted terms have not yet been established.

3. Cultures

The identification of a work with its author or with its title is to a large extent related to the cultural ambience. The Greek's identification of a work with its author reflected their democratic
orientation and their esteem for the individual. The resumption of that identification since the end of the Middle Ages stemmed from a similar orientation in the new era. In oriental cultures, which are less individualist, the traditional entry for a book was its title.

Until the publication of the first edition of the Nippon Cataloging Rules (1943), a title main-entry system had been used in almost all of the cataloging codes in Japan (Takawashi et al., 1989). According to Takawashi's explanation, this can be attributed to the Japanese attitude toward books; the Japanese tend to consider a work as existing independently apart from its author. Older Chinese and Japanese books continue to be entered under title.

A similar situation prevailed with Hebrew books. Most Hebrew books published up to the 20th century were in the field of Talmudic research, in its widest sense, that includes also those dealing with the Hebrew language. An analysis of their titles has revealed that the vast majority consisted of two-words expressions, usually taken literally from one of the books of the Bible (Haberman, 1968). Even a short biblical book such as Canticles (The Song of Songs) has been a source for no less than 50 titles over the past 1500 years (from the completion of the Talmud to our day): for example, "Locked garden" or "Nuts garden" (Riegler, 1993).

A more in depth examination of Hebrew book titles has shown that only in some cases the title is informative, which means that there is any connection between the title and the content of the book. It seems that in most cases when authors had to decide what title they would give a book that they had worked on possibly for many years, having an informative title that would enable information retrieval about the content was not major consideration. It seems that usually the author chose certain biblical phrase as a title mainly because this phrase included his name, whether explicitly or implicitly -- for example "Jacob's Tent" or "Yoshua's faces." Authors regarded this as a suitable way of perpetuating their names, and it is indeed a well-known historical fact that the actual names of authors of Talmudical books tended to be forgotten among the Jewish people, whereas the book titles became widespread surrogates, for these names.

4. Summary

The concept of the title has undergone many changes over the course of human, social and cultural history. In ancient civilizations books did not have a distinctive title; instead the opening words were used to designate the document. In some cultures, such as the Hebrew one, the book's title was chosen without any relation to its content, but instead constituted of a biblical phrase, sometimes with the author's first name implicit in it.

The advent of printing led to the appearance of the title page and the use of titles that represented a book's content in order to advertise and sell the book. As self-serving publicists, early printers issued book lists. They would put their firm's name, emblem and shop address on the front page of their books. Their use of title pages involved a reversal of scribal procedures: they would put the name of their firm first, whereas the scribal colophon had come last. The early printers also extended their new promotional techniques to the authors and artists whose work they published.

Since then the title has constituted a source of information about the book (or article). This process has accelerated during the 20th century, to a large extent because of computerization, with titles becoming more and more informative.

Throughout much of history, there was no clear distinction between a document's subject and its title. There were periods when terms related to the subject appeared in the author/title catalog; in other periods a catchword from the title was used as a subject. However, beginning with the appearance of theorists of knowledge organization in the late 19th century, a clear distinction was made between author catalog, title catalog and subject catalog, with each designated to fulfill a different role in the retrieval of information.
More recently, new retrieval methods that were developed for the computerized systems again caused an overlap between titles and subjects, and we are currently witnessing blurring of the distinction between title and subject approaches.

Today the title is a very important element of any scientific or scholarly book or article. Its primary function is to draw readers' attention to a book or article and indicate its content in a short glimpse, thus influencing its initial selection or rejection. Titles are also an important source for subject searches in most of the automated catalogs and databases.

Authors should be aware of this situation and should provide their writings with informative and indicative titles that will contain as many substantive keywords as possible as opposed to trivial words.

References
Frost, Carolyn (1989). Titles words as entry vocabulary to LCSH: correlation between assigned LCSH terms and derived termed from titles in bibliographic records with implications for subject access in online catalogs. *Cataloging and Classification Quarterly, 10*, 165-179.


