Desirable search features of web-based scholarly e-book systems

Shiao-Feng Su

Graduate Institute of Library and Information Science, National Chung-Hsing University, Taiwan, Republic of China

Abstract

Purpose – To examine what search capabilities are desirable in web-based scholarly e-book systems.

Design/methodology/approach – To present the author's opinion and interpretation about what search capabilities are desirable in web-based scholarly e-book systems through investigating studies of information seeking behavior and exploring current multimedia and hypertext techniques.

Findings – The findings are several desirable browsing and known-item search features of web-based scholarly e-book systems.

Originality/value – The author's suggestion on the implementation of a dedicated, user-oriented and user-friendly interface, which includes the search features recommended in this paper to unleash the power of multimedia and hyperlinks.

Keywords Electronic books, Internet, User interfaces, Information searches

Paper type Viewpoint

Introduction

E-books are deemed "the obvious next step . . . to ring a full line-up of Web-based basic library resources" and trusted "going to be increasingly important" in scholarly communication (Dillon, 2001, p. 113 and 121). E-books refer to "digital texts that are issued as individual works and designed to be accessed by using special software for text navigation and ease of reading" (Hughes, 2003, p. 984). In this paper, web-based scholarly e-books are PC and laptop-based, can be read on the computer monitor via connection to internet, and are produced for searching, studying, and researching with academic intent. Their target audience is usually researchers, professors, and students. netLibrary is a typical example of web-based scholarly e-book vendor.

A representative of a major e-book vendor recently proudly assured a roomful of professors that e-books would soon be as convenient in use as paper-based books. E-book vendors' effort in mimicking the presentation of paper-based books is indeed plausible. However, e-books being as good as their paper counterparts may not be enough to attract already satisfied paper-based book readers, they have to be better in order to be viable in academic community. Snowhill (2001) asserted that "What will really make e-books viable for academic use is added functionality over printed version". Making e-books easier to find and to use than their paper counterparts can be a good beginning to win some people over.

Currently, not many libraries attempt a dedicated search system for their e-book collection. Many libraries provide links for users to search e-books through the vendors' systems. Users then have to visit multiple links and navigate on several vendors' interfaces of different idiosyncrasies for an e-book, if the library works with

multiple vendors. Some libraries integrate e-book records with their paper-based counterparts in the OPAC. The single-record approach makes serious e-book users scroll endlessly down search result screens to dig out e-books. To justify the creation of a separate record for electronic resources, Crowley (2000, p. 32) argued: "not...OK to lump all of these different formats onto one record. What is the difference between that [combing print and the electronic records] and taking a video, a sound cassette, and several differing paper editions of a title and lumping them all onto one record?!"

A single-record approach supporter claimed that users "simply won't look" at the separate records for internet resources (Grenci, 2000, p. 35). However, a whole new library clientele becomes dependent on full-text electronic resources. The majority of university faculty members used electronic resources in their research several times a week (Hafner, 2004). Students of rapidly increasing distance learning programs consider e-textbooks and e-reserves a major solution to their learning and information needs (Hyatt and Connaway, 2002; Okerson, 2001). Users of real time live digital reference services, such as chat references, are also in favor of e-books (Snowhill, 2001). These pro-e-book users would most likely prefer a dedicated place for better focusing their searching activities.

The California Digital Library (CDL) Ebook Task Force (2001) outlines eight elements that are considered important to the study regarding academic use of e-books; one of them is about enhancement and ideal e-book features. It is worthwhile to examine what search features users might like to have in the e-book environment. Since remote users usually conduct searches without much instruction, a dedicated, user-oriented e-book search system is obviously preferable. The system should facilitate the search strategies frequently used by information seekers, not simply copying those adopted by the traditional OPAC. Since the ultimate type of e-books is hypermedia, which integrates textual material with video, sounds and pictures, and provides alternative reading paths (Feather and Sturges, 2003, p. 170), the system features should unleash the power of multimedia and hyperlinks vested in e-books.

Desirable e-book search features

Search capability development has been usually governed by system designers and librarians, but not necessarily according to user needs. The CDL Ebook Task Force (2001) recommends several "advanced" search capabilities, including truncation, Boolean, fielded, and proximity. Unfortunately, the uses of Boolean logic and proximity operators are not usually fully understood by scholars (Bates, 1996), let alone less experienced users. Korfhage (1997, pp. 56-7) pointed out: "Most users of information systems are not well trained in Boolean operators and their rules of precedence, combined with the fact that many users do not access an information retrieval system on a regular basis, presents a major barrier to the effective use of Boolean retrieval systems".

Ideally, a search system should facilitate all information seeking techniques. Bates (1989, p. 414)argued: "If we want to meet users' needs, we should enable them to search in familiar ways that are effective for them. ... From the standpoint of general effectiveness in searching ... other things being equal, the search with the widest range of search strategies available is the searcher with the greatest retrieval power".

In addition to traditional author-title-and-subject searching, preferable e-book search features should facilitate other frequently used search strategies identified in

research on users' information seeking. Desirable e-book search features are discussed in this writing in two sections for convenience: features of known-item search and features of browsing. Discussion of the desirable browsing features is further partitioned into two sections: shelf browsing across the e-book collection and content browsing within an e-book. In actual system, the user should be able to interchangeably use all features.

Known-item search

Traditional wisdom seems to suggest the superiority of author-title-and-subject searching to others. Some frequently used search strategies identified in research are therefore overlooked. Examples of such are footnote tracing and citation searching, which are popular search strategies among scholars and graduate students (Ackerson, 1996). Footnote tracing is to follow up footnotes or endnotes found in books or articles for further studying. Using this technique, one can find prior publications relevant to the subject of the book or article at hand. Citation searching is to find publications that cite the book or article at hand as a reference note. The former technique is for backward searching, and the latter forward searching, in terms of publication date. The CDL Ebook Task Force (2001) mentioned cited sources from footnote as one of the "additional desirable" search tools.

Traditional library OPAC paid little attention to automating footnote tracing and citation searching. The employment of such strategies is quite tedious and very time-consuming for both librarians and users in manual environments. The implementation of such capability with paper-based books requires indexers manually generate lists of citations and copious linking notes. Despite indexers' effort, users still have to conduct known-item searches one-by-one for each citation of interest.

Fortunately, it is relatively straightforward to effect and use the footnote tracing and citation searching strategies in electronic environments. Librarians need simply make necessary hyperlinks between e-books and their cited and citing references. Users can leap around among full-text books and articles through simple point-and-click. Not enhancing the e-book search system with these features would be a complete waste of the hereditary hyperlinking capacity in e-books.

Browsing

Information professionals must rid the mythical perception that browsing is casual searching. Browsing is actually an important work engagement among knowledge workers (Choo *et al.*, 2000). A well-designed interface with enhanced browsing capabilities would be valuable to web-based e-book users. The discussion of desirable browsing features is partitioned into two sections: shelf browsing and content browsing.

Shelf browsing. Shelf browsing is considered "the quintessential form of browsing in manual environments" (Bates, 1989, p. 417). It enables the user to browse books that are physically collocated with the book found earlier. Since paper-based books are usually arranged on shelves by library classification number, their physical collocation implies the possible closeness in subject contents. Many users scan over the bookshelves in a subject area of interest.

Although there are no physical bookshelves for the e-book collection, the system can provide an artificial shelf list for browsing arranged by a certain classification scheme. In fact, there can be several shelf lists employing different classification schemes, such as DDC, LC, and others. The shelf list browsing in electronic environments can be quite flexible. For instance, once the DDC number of an interested book is identified, the user can "hop" around the number for further browsing – to the root of the number or to a certain meaningful level by increasing or decreasing the number. The IFLA Task Force on Guidelines for OPAC Displays made several recommendations for good practice in OPAC displays for users who want to begin searching immediately without much instruction. The principle 22 recommends "Display the hierarchical relationship between a classification number and the entire classification" (Yee, 1999).

Take the following case as an example. A user finds out an e-book of interest whose DDC number is 351.4, Governments in Europe. If taking one level down, s/he can browse e-books under 351.41, Government of United Kingdom. If wishing to expand the geographic area, s/he may increase 351.4 by 0.1 or 0.2 for titles under 351.5, Governments in Asia, or those under 351.6, Governments in Africa, respectively. Or, the user may move several levels up to 350, Public Administration, or all the way up to 300, Social Science. In addition, since linear physical arrangement does not exist in electronic environments, each e-book title can be assigned with multiple classification numbers as necessary. Users can make more effective jumps that enhance the probability to discover more along the search route, some unexpectedly.

Some shelf browsing capabilities have been established on library OPAC and interfaces of commercial online bookstores. For instance, after finding an item in pursuit on bn.com (the online store of Barnes and Noble) search interface, the user can click on the "Find Related Books with Book Browser" button for links to more titles on related subjects. The user can also click on the "Other books by" button to browse works by any one of the authors. Another smart feature on some commercial online bookstore systems is the statistics look-alike listing of books that people usually purchased at the same time as the book in pursuit (e.g. People who bought this book also bought. . .). A study of a user circulation log demonstrated the possibility of non-hierarchical association of DDC classification numbers. The results indicated that, for instance, users who borrowed a 284 (Western Mythology) book usually also did an 870 (Western Literature) one (Pu and Yang, 2003). The e-book interface may incorporate such a feature for the user's reference. The interface can display books that are usually checked out or reserved at the same time, or associated classification numbers for further browsing.

Visualization can be an excellent aide for browsing. Bates (1989) suggested the use of spatial metaphors to help users recall where a certain book was placed in online environments. The e-book search interface can create a virtual "book room" that imitates paper-based book arrangement on shelves, visualizing one in a physical library. Upon entering the "room", the interface presents multiple library bookshelves containing e-book images. Each shelf is labeled with its beginning and ending classification numbers. The user may move the mouse pointer to "walk" around the bookshelves in the virtual book room. S/he may touch a certain bookshelf icon for further expansion to browse books on that particular shelf. Several projects on user study have implemented the bookshelf visualization feature to successfully assist searching (Borgman *et al.*, 1990).

The visualization feature can be personalized, with a password option. The feature allows the user to label an e-book by placing a computer-generated sticky note on a book spine image, or "laying down" and/or "pulling outward" a book image. Of course, the above-mentioned features request e-book publishers' goodwill. Since true color and graphics can help users in locating a previously seen book, it is preferable that e-book publishers digitize the book cover and the spine to be true to the print original, including graphics and color. In fact, the final report of the CDL Ebook Task Force (2001) has already made such demand on e-book contents.

The e-book system could also create a "new arrivals" area for users to browse the most recent additions of e-book titles in the collection. Such feature can be particularly helpful to the regular user. In physical library settings, many librarians shelved the books in the new book area to gain information about good reads (Hoffert, 2002). On the one hand, through seeing what book icon moves and what does not in e-book system, librarians can acquire an equally intimate understanding of their patrons' reading tastes without leaving their chair. On the one hand, users can see how popular an e-book is. The information could serve as a judging clue for borrowing decision. Alternatively, a user can request a virtual robot librarian to "tidy up" the shelves before his or her browsing activity begins.

With netLibrary alone digitizing 20,000 titles a year, the "new arrival" area should be able to attract some traffic. The feature helps encourage users to periodically come back for e-books.

Circulation and other information can be found by click the image of the e-book of interest. For example, on clicking the book spine image, information such as the frequency of being checked out (e.g. 123 times since last Monday), estimated return date, previous borrowers' comments, related book reviews, and other relevant information (e.g. the top x picks in y subject area) may pop out. Such information can serve as a reader's advisory.

Content browsing. People seek out familiarity of print in the web environment (Wathen and Burkekk, 2002). They tend to locate criteria used in judging paper-based materials to evaluate the web-based resources. Summarizing the results of interviews to evaluate the scholars' reactions to experimental e-books (termed as "online book" then), Summerfield (1999) reported that scholars feel that browsing in an online book could have great utility but that the online system must mimic the way in which users browse print books, e.g. to page smoothly and quickly through a book or to focus on a graphical element and read the material on the pages around it. Equally the user should be able to move quickly, preferably with a clicking feature, from the Table of Contents to a chapter or from an index to a point referenced there.

Basic content browsing and linking activities, such as jumping from the Table of Contents to a chapter and from any back-of-the-book indexes to a point referenced there are fundamental features. Observing that library users often flip through pages and randomly sample paragraphs to decide whether or not to check out the book at hand, Bates (1989, p. 420)argued that: "It would be easy to program a command that would produce a series of randomly selected passages, or 'snapshots', each two or three paragraphs in length. Such passages should be truly randomly selected – just as happens when we flip through an article or book – because it is precisely what is not indexed that we want to sample".

In addition to Bates' "lucky draw" feature, users can opt to read exactly keyword-matched paragraphs to check the author's perspectives or the depth of the content. With full-text searching capability and probable term weighting techniques, the search interface can easily locate paragraphs with the closest relevance to the search term(s). If the selected paragraphs located in multiple pages, the interface could provide options of reading either paragraph to paragraph in continuity or with page-flipping animation.

Conclusions

Dillon (2001, p. 121) posed a vital question regarding the spirit of the e-book: "The speed of e-book development has scattered a cloud of dust over the key question 'what is an e-book?' Is it merely a digital version of the long familiar printed codex, or is it something more, or possibly something else?"

E-books that are embedded with multimedia and hyperlinks should be considered as a completely different kind of medium from the paper-based book. The search interface must unleash the power vested in multimedia- and hyperlink-embedded electronic resources in order to be more effective and user-friendly, especially for remote users. The design focus of the search system ought to be placed upon features that are actually desirable by the information seekers.

As users seek out familiarity of print in electronic environments, the ideal e-book search interface should exploit multimedia and animation techniques to make users feel that they are almost present in the physical library and conduct searches in familiar ways with the print environments. Prior research results also suggest that the e-book system mimic the way in which users browse paper-based books.

Through embedded hyperlinks in e-books, many well-established search strategies that were not implemented in library OPACs, such as footnote tracing and citation searching, can be easily accomplished in e-book interfaces. Most of the features suggested in this paper can be easily implemented in electronic environments with readily digitized multimedia- and hyperlink-embedded e-books. Although currently being exceedingly "flat", e-books are destined to be dynamic. Snowhill (2001) asserted that the current mindset is to replicate the print version of a book, but further development needs to recognize the potential scholarly significance of increased integration of unconventional media with text in e-books.

Multimedia is changing the nature of reading itself, even that of education (Hofstetter, 2001, p. 3 and p. 9). Multimedia-embedded e-books could play an important role in the possible change of reading habits. The author suggests that librarians and system designers attempt to work toward the implementation of the recommended features in this writing when developing new versions of software in the future. With better searching capabilities and multimedia presentation, there will probably soon be a day when users start their search with an e-book system. Loyal paper-based book readers can later go back to the paper-based book catalog to perform a known-item search for availability. Eventually, people may check out paper-based books only when e-books are not available. As Lancaster (1982, p. 108) prophesied 20 years ago: "...print-on-paper publications give way and are eventually completely replaced by electronic publications... A research 'library' of the future, then, need not contain any printed materials at all. It could be a room containing only terminals. Apart from archival repositories of printed records of the past and institutions designed primarily

to lend inspirational/recreational reading materials, the libraries of today could well disappear".

Perhaps in our wildest dream, the library in the future could simply be a *halo deck*, like one in the *Star Trek* Movies. Upon entering, the user sees a virtual library full of virtual bookshelves stacked with e-books. S/he can touch an e-book of interest and start virtually living the world in the book. Then, searching, browsing, and reading a "book" will have a whole new meaning.

References

Ackerson, L.G. (1996), "Basing reference service on scientific communication: toward a more effective model for science graduate students", *RQ*, Vol. 36 No. 2, pp. 248-60.

Bates, M.J. (1989), "The design of browsing and berrypicking techniques for the online search interface", *Online Review*, Vol. 13 No. 5, pp. 407-24.

Bates, M.J. (1996), "The getty end-users online searching project in the humanities: Report No 6: overview and conclusions", *College & Research Libraries*, Vol. 57 No. 6, pp. 514-23.

Borgman, C.L., Gallagher, A.L., Krieger, D. and Bower, J. (1990), "Children's use of an interactive catalog of science materials", *Proceedings of the ASIS Annual Meeting*, Vol. 27, pp. 55-67.

California Digital Library Joint Steering Committee for Shared Collections Ebook Task Force (2001), "eBook Task Force Final Report", available at: www.cdlib.org/about/publications/ ebookappendixf.pdf

Choo, C.W., Detlor, B. and Turnbull, D. (2000), "Working the web: an empirical model of web use", *Proceedings of the Hawaii International Conference on System Sciences*, 4-7 January, Maui.

Crowley, J.M. (2000), "Classification of internet resources: an AUTOCAT discussion", *Cataloging & Classification Quarterly*, Vol. 29 No. 4, pp. 19-38.

Dillon, D. (2001), "E-books: the University of Texas Experience, part 1", *Library Hi Tech*, Vol. 19 No. 2, pp. 113-24.

Feather, J. and Sturges, P. (Eds) (2003), *International Encyclopedia of Information and Library Science*, 2nd ed, Routledge, New York, NY, pp. 168-71.

Grenci, M. (2000), "Classification of internet resources: an AUTOCAT discussion", *Cataloging & Classification Quarterly*, Vol. 29 No. 4, pp. 19-38.

Hafner, K. (2004), "Old search engine, the library tries to fit into a google world", *The New York Times*, available at: www.nytimes.com/2004/06/21/technology/21Libr.html?ex = 1088993258&ei = l&en = 477360266097e820

Hoffert, B. (2002), "Book report 2002: the amazon effect", *Library Journal*, Vol. 127 No. 3, pp. 38-40.

Hofstetter, F.T. (2001), *Multimedia Literacy*, McGraw-Hill, Boston, MA.

Hughes, C.A. (2003), "E-books", in Drake, M.A (Ed.), *Encyclopedia of Library and Information Science*, 2nd ed., Marcel Dekker, New York, NY, pp. 984-9.

Hyatt, S. and Connaway, L.S. (2002), "Utilizing e-books to enhance digital library offerings", available at: www.ariadne.ac.uk/issue33/netlibrary

Korfhage, R.R. (1997), *Information Storage and Retrieval*, Wiley, New York, NY.

Lancaster, F.W. (1982), *Libraries and Librarians in an Age of Electronics*, Information Resources Press, Arlington, VA.

Okerson, A. (2001), "Wanted: a model for e-reserves", *Library Journal*, Vol. 126 No. 14, pp. 56-8.

Pu, H. and Yang, C. (2003), "Enriching user-oriented class associations for library classification schemes", *Electronic Library*, Vol. 21 No. 2, pp. 130-41.

Snowhill, L. (2001), "E-books and their future in academic libraries: an overview", *D-lib*, Vol. 7 No. 7-8, available at: www.dlib.org/dlib/july01/snowhill/07snowhill.html

Summerfield, M. (1999), "Online books: what roles will they fill for users of the academic library?", available at: www.columbia.edu/cu/libraries/digital/texts/paper/

Wathen, C.N. and Burkekk, J. (2002), "Believe it or not: factors influencing credibility on the web", *Journal of the American Society of Information Science and Technology*, Vol. 52 No. 2, pp. 134-44.

Yee, M. (1999), "Guidelines for OPAC displays", available at: www.ifla.org/IV/ifla65/papers/098-131e.htm.

(Shiao-Feng Su is currently an Assistant Professor at the Graduate Institute of Library and Information Science, National Chung-Hsing University, Taiwan, Republic of China. Her research interests include interface design, digital reference, and information seeking behavior. She can be reached by e-mail: sfsu@dragon.nchu.edu.tw or sfadasu@seed.net.tw, by cell phone: 886-963063801, or by mail at Graduate Institute of Library and Information Science, National Chung-Hsing University, 250 Kuo Kuang Road, Taichung, Taiwan, Republic of China.)