Personalization of Information Services

Yaşar Tonta
Department of Information Management
Hacettepe University
06532 Beytepe, Ankara
Turkey
tonta@hacettepe.edu.tr

Introduction

The Internet removes the temporal and spatial barriers and enables information professionals to offer information services to remote users 24 hours a day and seven days a week. Users experience "instant satisfaction" when they get "instant access" to networked information sources such as bibliographic databases, full-text electronic journals and electronic document delivery services. The access paradox of "single source-single user" for printed information sources can be overcome by providing remote access by multiple users to the networked copy of the same single source. Libraries are increasingly opting for collection management practices that are based on "just-in-time" approach rather than the "just-in-case" approach that has been prevalent over the years. Members of the Association of Research Libraries (ARL) in the United States spend as much as 25% of their total material acquisition budgets for networked information sources and services in recent years. This paper reviews some of the current trends in information management (ownership vs. access, disintermediation, and personalization) and emphasizes the importance of providing personalized information services along with a discussion of some of the issues involved.

Ownership vs. access

Currently, collection management practices of library and information centers are increasingly slanted towards access rather than ownership. Ownership usually dictates the use of centralized information management models whereas instant access to networked information sources and services requires decentralized/distributed ones. Information managers are no longer responsible solely from their locally held physical collections. They have to assume further responsibilities in order to provide access to a wide variety of networked information sources that they do not necessarily own. This necessitates the development of cooperative or consortial collection management practices involving other information centers, publishers and/or aggregators.

Economic models based on centralization are replaced with those based on distributed and personalized information services. For instance, in a centralized model an information manager has to decide once a year as to which printed journals to subscribe to, thinking that those journals would satisfy information needs of most users, make a one-time payment and have them used as often as possible. For unsatisfied information needs, traditional interlibrary loan or document delivery services are available. Yet the use of such services is usually not encouraged nor are they terribly convenient to use. Compare this with licensing full-text electronic journals where the economic model is based on such criteria as the frequency of use, number and types of users and sites (e.g., students, researchers, "walk-in" users, simultaneous users, distance learners), and perpetual access to back issues. Moreover, if users wish to get access to full-texts of journals that are not licensed, different payment schemes (e.g., pay-per-view) are usually in place for electronic document delivery services.

Information managers have to sign separate licences for a large number of electronic resources and develop separate use, processing, maintenance and storage policies. No longer do they have the luxury of determining their own information management policies. Instead, they are "interdependent" on their colleagues as well as on information producers/providers and library consortia to develop and coordinate their own policies.

Disintermediation

While distributed, networked access to information sources and services gives users "instant satisfaction", it also eliminates face-to-face communication between users and intermediaries. For example, reference services are usually provided through a single reference desk in a traditional library. This does not serve the needs of remote users at all. Similarly, remote users would prefer electronic document delivery services rather than paying a visit to the library to get photocopies of journal articles. It is of no use advising them to come to the library in person to use reference services or the printed journal collection. Once users get accustomed to instant access to networked sources, they would look elsewhere for instant satisfaction. Intermediated services can mostly be provided by means of centralized information management models whereas disintermediated services can be offered through distributed models. Remote users can easily get instant access to both sources and services spread around the network without intermediation.

The impact of networked access to information services and disintermediation is also reflected in use statistics of ARL libraries. For example, the total number of interlibrary loan (including document delivery) transactions of 125 ARL libraries has almost doubled between 1991 and 2000 whereas the total number of reference queries decreased 12% and the total number of materials borrowed decreased 6% during the same period (Kyrillidou & Young, 2001). Such statistics indicate that libraries are getting more dependent on each other while, at the same time, users tend to go to libraries less often for in-house services. They seem to prefer instant access to networked information sources and electronic print (e-print) archives.

The trend towards remote access and disintermediation is observed elsewhere, too. For example, the number of intermediated searches performed at the Turkish Academic Network and Information Center (TANIC) between 2000 and 2002 has been halved. In 2002, the total number of on-site searches performed by users themselves was equal to one-third of what it was in the year 2000. Yet the Center has witnessed, during the same period, a 13-fold increase in the use of its Web-accessible medical database (www.ulakbim.gov.tr).

Personalization

In order not to shy away remote users, information managers try to offer some services through the Web nowadays. Although it is important to have remote and disintermediated access to networked information services, users simply need more than that, however. They wish to customize their information environments and be able to get personalized information services. "Personalization is the process whereby a program 'follows' what a user does on a website and tries to 'match' the user's behavior by providing information related to what the user has previously done" (Kotwica, 1999). Active and passive methods of personalization are both used. Information display environments, content and information services can be personalized. Several technologies are used to gather data about users, ranging from the fill-in profile forms to cookies, to click-stream analysis/web usage mining systems and collaborative filtering (Bonett, 2001).

Users can take an active role and collaborate with information managers by explicitly describing their interests to make use of personalized "alert" or table of contents (TOC)

services. Or, the interests of users can be identified by analyzing their previous use patterns and behaviors that are stored in cookies. For instance, profiles of users of an electronic commerce web site can be created unobtrusively by simply analyzing the items that they retrieve and display on the screen during their visits. Such a system can keep track of five types of actions. If the user buys the item displayed on the screen, this is interpreted as strong positive feedback. If she simply browses without buying, this is interpreted as weak positive feedback. On the other hand, if the user explicitly skips an item when presented or if she removes the item from the list, these actions are interpreted as negative feedback and strong negative feedback, respectively. If the user simply does not get to an item on the list, this is interpreted as "no change" in her "interest representation." These feedback values are used to update the user's profile unobtrusively (Mostafa, 2002, p. 10). Online bookstores such as Amazon.com use similar techniques to update users' preferences and recommend further titles.

Users can personalize their information display environments by setting some parameters. They can make their favorite databases, e-journals or bookmarks always visible on the screen and change the display layout as they wish. News portals, banks and libraries offer such personalized information display environments (e.g., MyYahoo, MyCNN, Wells Fargo, and MyLibrary).

Currently, information retrieval systems display standard content to all users regardless of their access methods (e.g., remote vs. on-site) or privileges (e.g., student vs. faculty). Yet the content can be personalized if individual users are recognized by the system when they log on. For example, authorized users can enjoy more privileges (i.e., access to full-text e-journals, electronic reserve collections or electronic document delivery services). On the other hand, unauthorized users getting access to the system from off-campus machines may not even be informed of the availability of, say, web access to full-text e-journal or e-reserve collection.

Providing personalized information services is perhaps the most advanced application of personalization as such. Such services can be offered using both "pull" and "push" technologies once the mechanism to create and update user profiles is in place and the individual user is recognized every time she logs on to the system. Some banks already offer personalized services such as electronic fund transfers. Many phone companies already feed personalized information on share prices, weather forecast and match scores to GSM numbers of their customers. Similar services can also be provided by information centers. Users can be informed of newly acquired books or current additions to the full-text article database in their areas of interest by analyzing their previous transactions (e.g., books checked out or articles downloaded). Unless recalled by another user, they can renew their library books through the Web. Users can be informed of the availability of recalled books. The PDF copies of articles obtained through electronic document delivery services can be sent to users' electronic mail (e-mail) addresses.

Personalization of collections and services on the basis of individual preferences and privileges requires the use of more sophisticated techniques. Information about users should be kept and updated as long as they continue to use the system. Although libraries can provide personalized information services similar to those of Amazon.com, they are reluctant to do so for security and privacy reasons. Lynch (2001) points out that "circulation systems typically break the link between a patron and a book that has been borrowed when that book is returned" and thus libraries lose the opportunity of providing more personalized services. Libraries usually evaluate user data in aggregates and later destroy it. Lynch also emphasizes the fact that it is quite difficult to implement personalization in a distributed information environment as personalization "occurs separately within each system that one interacts with" and "[i]nvestment in personalizing one system (either through explicit action or just long use) are not transferable to another system."

Recently, personalized electronic books (e-books) emerged. Users can add their own annotations or hyperlinks to the existing text and would like to see them there whenever they use the same e-book (Ohene-Djan & Fernandes, 2003). As the use of e-books in library and information centers increases, it remains to be seen if libraries can incorporate those personalized additions to e-books and keep multiple individual copies of e-books in their collections. This would mean that the number of objects that the library has to deal with would multiply several orders of magnitude as each copy of an e-book should be stored along with a number of personalized annotations and hyperlinks. More sophisticated database management systems would be needed to handle different "versions" of the same e-book and to recognize several "owners" of each e-book as they log on to the system.

Although personalized e-books came into being within the last decade or so, personal libraries containing "associative links" among different information objects in the database is not a novel idea. Bush (1945) had predicted that such personal libraries complete with personal annotations and links could be built. In his seminal article entitled "As we may think," he called this personal library "memex" (memory expander) and gave the detailed description of its components (building, indexing, creating links among information items, underlining certain parts of text, etc.). Memex is considered to be the predecessor of the current World Wide Web that is based on hypertextual links as envisioned by Bush some 60 years ago. We are able to build personal libraries of e-books similar to Bush's memex (Ohene-Djan & Fernandes, 2003). The challenge today is to build personalized libraries in a distributed environment involving several digital collections. This would certainly involve much more than just sheer computational power and large bandwidths.

Issues

We can safely predict that the number of personalized information services offered will increase tremendously in the near future. Users managing their money from afar using web sites of banks would certainly appreciate it if they could renew their books using the library's web site.   It is of no use telling them that library's web site and its circulation system are not "interoperable." In the near future, users will not be content with simply renewing their books through the Web. They will ask if they could download the contents of every book (e-book or otherwise) that they are interested in that the library owns or provides access to. They will demand electronic document delivery requests to be delivered to their desktops or personal digital assistants (PDAs). Institute for Scientific Information (ISI) has already advertised that it will incorporate personalized alert services to its Web of Knowledge (Personalization, 2002). The British Library (BL) recently signed an agreement with Elsevier Science and Adobe to provide print-quality copies of articles delivered to users' desktops (British Library, 2002). The agreement allows BL to supply PDF copies of journal articles from over 1700 key Elsevier titles. Users will pay a fixed annual subscription fee (500GBP) plus downloading (4.5 GBP) and copyright charge (varies) for each article that they download (Kraan, 2002).

Users will soon demand similar services from all library and information centers. Libraries providing traditional services up to now should therefore strive to transfer those services to the Web. In addition, they should develop new, personalized services that will enable remote users to get instant satisfaction.

Libraries contemplating to offer personalized information services should have a sound network infrastructure as well as access to personal, local, regional and wide area networks. Security and privacy mechanisms should be thought of as part of the network infrastructure and standards should be developed in order to function in a distributed environment. Personalized information services should be provided on the basis of predefined user rights and privileges. Personalized services offered by libraries should be integrated with those,

which are offered by other institutions. For example, personalized information services should be an integral component of an electronic university. For this, systems used by the library should be interoperable with other on-campus systems (registrar's office, financial systems, health systems, to name but a few). They should also be integrated (and interoperable) with off-campus systems such as other electronic learning (e-learning) institutions and electronic banks. For instance, users should be able to transfer money from their bank accounts to the library's account to pay for, say, electronic document delivery services.

More sophisticated budgeting, pricing, use and training models that are required to deliver direct and personalized services to individual users should be developed. The "one-size-fits-all" approach will not satisfy information needs of ever-demanding users. There will always be novice users who are at the bottom of the learning curve and need some "hand-holding." Although not inexpensive, some remote training sessions can be delivered through the Web using what is called "customer relationship management" software (also used for e-reference services).

If I can borrow the analogy that was used in the report of the Committee on Information Strategy for the Library of Congress to describe the current LC cataloging practices, developing personalized information services requires "relationship-centric" approach to information management rather than "resource-centric" approach that has been prevalent in libraries (Committee, 2000; cited in Lagoze, 2000). As Lagoze (2000) noted in a somewhat different context, "… the resource-centric descriptive model upon which current cataloging practices are built, whereby discrete descriptive records are associated with fixed information artifacts, is incompatible with networked digital information. This new context has radically different information entities, decentralized information production and management, and troublesome questions about authenticity and trust. It requires a model that can flexibly express the relationships between resources, abstract concepts, and multiple descriptions of those resources and concepts" (Lagoze, 2000). One should also add "relationships between resources and users" to the list given above, if we are to offer personalized services in a distributed environment.

Conclusion

It should be kept in mind that instant satisfaction is only possible with the availability of instant access to networked resources and services. If information professionals get caught unprepared for this new era of personalized information services and our services are not instantly accessible through the Web, libraries may easily be bypassed by remote users. Some other institutions –likely to be more prepared than libraries- may emerge to supply those services. More demanding users will then "take their business elsewhere" to get instant satisfaction. Moreover, one should be concerned that potential users who might think that libraries cannot provide such services anyway may ignore libraries altogether. Therefore libraries should strive to transform their services and make them visible and accessible through the Web as expeditiously as possible.

References

Bonett, M. (2001, June 22). Personalization of web services: opportunities and challenges. *Ariadne,* No. 28, [Online]. Available: http://www.ariadne.ac.uk/issue28/personalization/intro.html [02 April 2003].

The British Library (2002, December 3). Encryption technology heralds new era in electronic document delivery - The British Library, Adobe and Elsevier Science set standard for

secure EDD (press release) [Online]. Available: http://www.bl.uk/cgi-bin/press.cgi?story=1310 [29 March 2003].

Bush, V. (1945 July). As we may think. *Atlantic Monthly*, 176(1): 101-108. [Online]. Available: http://www.theatlantic.com/unbound/flashbks/computer/bushf.htm [02 April 2003].

Committee on Information Strategy for the Library of Congress. (2000). *LC21: A digital strategy for the Library of Congress (2000)*. Washington, DC: National Academy Press.

Kotwica, K. (1999 July-November). Survey: Website personalization. [Online]. Available: http://www.cio.com/behavior/edit/survey7.html [28 March 2003].

Kraan, W. (2002, December 11). More journals to your desktop, courtesy of the British Library, Adobe and Elsevier. [Online]. Available: http://www.cetis.ac.uk/content/20021211203949/printarticle [29 March 2003].

Kyrillidou, M. & Young, M. (2001, August 21). ARL statistics trends: an introduction. [Online]. Available: http://www.arl.org/stats/arlstat/00pub/intro.html [02 April 2003].

Lagoze, L. (2000). Business unusual: how "event-awareness" may breathe life into the catalog? [Online]. Paper prepared for *Bicentennial Conference on Bibliographic Control for the New Millennium, Library of Congress, November 15-17, 2000.* Available: http://lcweb.loc.gov/catdir/bibcontrol/lagoze_paper.html [02 April 2003].

Lynch, C.A. (2001). Personalization and recommender systems in the larger context: New directions and research questions. [Online]. Keynote paper presented at *Second DELOS Network of Excellence Workshop on Personalisation and Recommender Systems in Digital Libraries, Dublin, Ireland, June 18-20, 2001*. Available: http://www.ercim.org/publication/ws-proceedings/DelNoe02/CliffordLynchAbstract.pdf [29 March 2003].

Mostafa, J. (2002, November-December). Information customization. *IEEE Intelligent Systems*, 17(6): 8-11.

Ohene-Djan, J. & Fernandes, A.A.A. (2003). Personalising electronic books, *Journal of Digital Information*, 3(4), [Online]. Available: http://jodi.ecs.soton.ac.uk/Articles/v03/i04/Ohene-Djan/ [02 April 2003].

Personalization and alerting to be added to ISI Web of Knowledge. (2002, November 25). [Online]. Available: http://www.isinet.com/isi/news/2002news/isi/8146251/ [02 April 2003].