

Making it Easier to Connect with the Collection

It works as a kind of doppelganger search engine. All of us have had the experience meeting people and discovering that they appear to have the very same tastes we do -- that they really love the same obscure foreign films that we love, or that they are fans of the same little-known novelist whom we are obsessed with. If such a person recommended a book to you, you'd take that recommendation seriously, because cultural tastes seem to run in patterns. If you and your doppelganger love the same ten books, chances are you'll also like the eleventh book he likes. Collaborative filtering is simply a system that sifts through the opinions and preferences of thousands of people and systematically finds your doppelganger -- and then tells you what your doppelganger's eleventh favorite book is.

- Riedl and Konstan, *Word of Mouse*, 2002, xi

After the hard work of choosing a library collection, one measure of success is how well the collection is used by our patrons. To maximize the utilization of the collection and increase patron satisfaction, libraries need to find effective ways to connect people with the available resources. Should libraries supplement their traditional reader's advisory services by using collaborative filtering tools to connect patrons with their reading doppelgangers? Some people think so, including Joan Frye Williams, a library consultant and futurist, who writes that "It's time to think seriously about supplementing traditional reader's advisory service by layering user rating systems and collaborative filtering systems onto our online catalogs." (2002) Library patrons often struggle to find the information they need among the wide selection of resources available in the library's collections, and they are responsive to assistance and guidance. Williams explains that these collaborative filtering tools can "help us to collect and present the combined experiences and insights of users in ways that enhance both

retrieval and user satisfaction.” (2002) How do offering services like that fit into the library’s collection development and readers’ advisory services?

Collection Development & Reader’s Advisory

Librarians carefully develop and manage the library’s collection of books, periodicals, audio-visual materials and other resources with their audience and community in mind. Clearly then it follows that librarians should try to find ways to assist patrons in connecting to the materials they need or might enjoy. Libraries do this in a number of ways, from offering access to the library’s catalog, having librarians available at the reference desk to answer questions and find materials, setting up displays of books on various subjects, and publishing booklists and subject guides. Many patrons look to librarians to offer recommendations and opinions, and libraries should use all the tools at their disposal, including new technologies, to assist in connecting readers to the collections so carefully chosen to meet the needs of their community.

Readers’ Advisory services are offered to different degrees in different libraries, and its popularity has waxed and waned over the years. In her 1991 article, “Readers’ advisory service: new directions,” Catherine Sheldrick Ross explains that “Librarians must be more active as intermediaries in bridging the gap between the readers’ needs and the library’s store of books.” (1991) Libraries are also increasingly providing readers’ advisory online through their web sites, databases and even via email and online chats. Paula Wilson suggests that, “If libraries want to successfully provide online RA services they can start by studying the competition, surveying their patrons, and selecting the best technology to implement their services.” (2001)

According to a Readers Advisory Committee of the American Library Association, “Readers’ services in public libraries have been reinvigorated over the past decade.” They have found that, “the value of the human contact of a reader/librarian interaction has been recognized as a means for libraries to remain vital in their communities.” (RUSA, 2004) In readers’ advisory, the librarian tries to “discover why the reader finds certain types of literature interesting, then provide suggestions about books that might also be appealing to this particular individual.” It is recommended that the librarian try to “understand the user’s current tastes and needs as thoroughly and objectively as possible, because... suggestions that are based on the librarian’s tastes rather than the user’s will only be ignored.” (Bopp and Smith, 2001, p. 60)

Studies have shown that “users prefer personal recommendations to any other kind of information,” (Wiley, 1998) and in many cases recommendations from a readers’ advisory librarian are very well received. On the other hand, Williams feels that “users value peer recommendations (‘What did other people just like me think was useful?’) more highly than they value the advice of professional librarians.” (2002) The key here may be the desire to find those people just like you – like the doppelgangers described by Riedl and Konstan in their 2002 book, *Word of Mouse: The Marketing Power of Collaborative Filtering*.

Researchers Andreas Geyer-Schulz, Andreas Neumann, and Anke Thede have found that, “Due to – among other things – information overload and difficult quality assessment, information seekers are more and more incapable of compiling relevant literature from conventional database-oriented catalog systems in a time-efficient manner.” (2003) The authors point to a survey which found that students often rely

heavily on recommendations from their peers, and suggest that behavior-based expert advice services would free up valuable time for students, teachers and researchers.

“Moreover, in this scenario, they can also profit from the combined knowledge of all library users in contrast to the more restricted knowledge within their personal networks.”

(Geyer-Schulz, et al, 2003)

Librarians frequently create book lists recommending titles that may interest readers of particular authors or books (“If you liked Harry Potter, you might want to try..”) Already in the catalog you can often link to books on the same subject, why not import those reading lists so that Harry Potter in the catalog links you right to a listing of ten other books you might want to try? In low-tech ways, many librarians have also provided ways for “readers [to] share their enthusiasm for books.” In some cases, for example, libraries provide recommendation forms “with space for library patrons to write in the name of their favorite library book, what it was about, why they liked it, and how it helped them. Then [they] make these forms available for other borrowers to consult by putting them in protector sheets in a three-ring binder.” (Ross, 1991)

Because it can be overwhelming for many patrons to find what they need at the library, especially in large libraries with thousands of choices, some readers have developed their own systems to help find books they might like. Ross reports that, “Left to their own devices, some resourceful browsers develop makeshift ways of narrowing down choices and coping” with the overload of information. One reader reports, "At the library, there's a section of just-returned books and I always check those out first. I place faith in people having chosen popular novels and returning them." Another example Ross points to in her article is a patron who had discovered some other “reader of similar tastes

who initials the inside back cover of paperbacks she has read. Finding these initials to be reliable guides, she takes out any books bearing these initials and has rarely been disappointed.” Overall, the author has found that “the existence of these makeshift systems suggests a need for reading guidance that is not being filled.” (1991) An automated recommender system might be able to begin to fill some of that need.

Collaborative filtering software and recommender systems

The technology does now exist to help users find their doppelgangers in a more systematic way than flipping through three-ring binders, and “consumer acceptance and convenience of recommender systems are shown by the huge success of the broad variety of difference services offered at commercial bookstore sites.” (Geyer-Schulz, et al, 2003) When I go to Amazon.com, for example, I am presented with lists of recommendations of other books I might enjoy and I admit that I have often purchased titles from those recommendations. The web site explains that the recommendation services “are a way of making sure that you don't miss the perfect item.” The system examines items that I have previously purchased from them, items I've told them I already own, and items I've rated through their system. (Amazon.com) They then compare my activity on their web site with that of other customers, and try to find my doppelganger's suggestion for that eleventh book that I just will not be able to resist purchasing from them.

This same type of system could be used in libraries. Geyer-Schulz et al have found that “Library systems are a very promising application area for behavior-based recommender services.” Data mining can be used to glean data from circulation and searching log files from online public access catalogs. By creating user-centered portals

like Amazon offers, they believe that “Reductions in the search and evaluation costs of documents for readers, as well as an improvement in customer support and collection management for the librarians, are some of the possible benefits.” (2003)

Collaborative filtering is often defined as “a set of algorithms that uses the preferences of a community to recommend items to specific individuals.” (Riedl and Jonst, 2002)

Generally, the input into a recommender system, or "the ways customers demonstrate their preferences," takes the form of either explicit ratings assigned by the customer or "observed inputs generated from a customer's natural interaction with a site" such as purchase history and navigation (to identify current interests). This is combined with community based information such as ratings, purchase history, and navigation of others, as well as reviews others have written, and popularity measures. The output from the recommender system usually comes in the forms of suggestions, predictions of how much you might like a particular item, and ratings and reviews. (Riedl and Jonst, 2002)

While “most commercial applications use automated systems that gather customer preferences, identify customers with similar tastes, and use their experiences to recommend products for each individual,” it can also be done manually by hand-generating lists of suggestions. (Riedl and Jonst, 2002) Resnick and Varian prefer the term “recommender system” to collaborative filtering software for two reasons – first that it is not necessarily a collaboration because the recipients of the advice don’t necessarily know or work with those making the recommendations, and because it does not exclusively *filter out* information but rather suggests items that it thinks would be interesting to the recipient. (1997) There are various types of recommender systems,

including some very low-tech that librarians have been using for generations like the booklists and binders.

Librarians have been recommending titles in these ways since the early days of librarianship. Searchable databases, such as the online library catalog, are in a way also recommender systems already. Riedl and Jonst write that, “When the database is indexed in meaningful ways... customers may be able to narrow their search significantly just by following the categories. Indeed, the category descriptions are a form of recommendation - they recommend sets of items the marketer thinks are useful to view together.” (2002, p. 12) However many online catalogs do a poor job of grouping titles in helpful, meaningful ways. For example, searching for “The Sisterhood of the Traveling Pants” in the online catalog of one library gives you the subject heading “*Pants -- Fiction.*” while the sequel, “The second summer of the sisterhood” is labeled “*Jeans (Clothing) – Fiction.*” (Santa Clara County) A link to a hand-selected or automated list of books that you might like if you liked that title (including the sequel) might be more helpful to a library patron trying to find something to read next.

In their research, Geyer-Schulz et al, found three major reasons why more scientific libraries had not adopted these tools – privacy, budget restrictions and data size. (Geyer-Schulz, et al, 2003, p. 2) Rather than an individual library system developing the software to implement a recommender system, perhaps the big online catalog vendors will eventually add in the option and allow the pooling of data to provide access to a much greater data set. The issue of privacy still needs to be resolved, but if patron identifiers are not attached to the data in any way, the anonymous information should not pose any real privacy concerns.

Some libraries and catalog vendors are experimenting with different ways of bringing in additional information. Sirsi's "Ibistro e-library" (www.infobistro.com) -- includes hotlinks "to what others are reading" (Wilson, 2001) Some libraries are linking right to Amazon's entries from their online catalog to provide this information. "Most catalogs don't tell you what the book is about, what other people think of it," said Tacoma PL Community Relations Officer David Domkoski. While patrons at any library with Internet access can search Amazon.com's web site, the link from the catalog, he said, "makes an already easy-to-use program at Amazon easier to use." (Oder, 1998) Carol Tenopir agrees, writing, "Now the best solution, for public librarians at least, may be to link from their catalogs to Amazon to take advantage of all of Amazon's good implementation." (2003) But this could be construed as a commercial endorsement, and many libraries have policies explicitly forbidding that type of linkage.

In addition, many libraries subscribe to online services that meet some of these needs. Wilson explains that "Libraries with more money than staff time can start by offering remote access to online databases like Ebsco Publishing's Novelist (<http://novelist.epnet.com>) or What Do I Read Next? from the Gale Group (www.galegroup.com). Novelist provides extensive subject and keyword access to more than 100,000 fiction titles, full-text book reviews, book discussion guides, and related links to author home pages and other fiction-related Web sites. What do I Read Next? includes access to both fiction and nonfiction titles by subject, genre, time period, geographic location, character, and reading level." (2001) Integrating services such as these further into the online catalog may be a future option. Perhaps libraries will even be able to offer patrons a choice of whether they want global data (since one's

doppelgangers could be anywhere!) or specific localized information (since one's community may have distinct reading patterns and tastes that you share.)

Effects on collections and patrons

Geyer-Schulz, et al, suggest that future research should be conducted on the impact of the recommender systems on collection management. They find that “repeat-buying models aggregated at the level of document categories could be used for implementing a more customer-centric collection policy.” (Geyer-Schulz, et al, 2003, p. 12) By really understanding user patterns and preferences, collection management decisions can be more closely aligned to the real needs of the patrons.

Geyer-Schulz et al have found that commercial entities using recommender systems generate improved revenues by facilitating cross- and up-selling, and that the systems support marketing research by continuous consumer panel analysis. For their customers, these systems “reduce search cost and lead to a better overview of available products, support the discovery of related products and product groups, [and] reveal market leaders and standard products.” (Geyer-Schulz and Hahsler, 2002) And while libraries are not interested in up-selling other titles to increase revenue, circulation statistics are an oft-cited measure of library success and being able to show that more of the library's materials are circulating more often is helpful to securing public support of the library.

Some librarians point out the limitations of relying on automated recommender systems. Barbara Hoffert, for example, writes that “the delicate act of connecting reader and book... cannot easily be mechanized, and librarians make a big mistake when they

abandon it to Amazon.” (2003) She writes that “librarians might stop and think that they are better situated than the folks at Amazon or the New York Times to suggest what their patrons will like.” She advocates for the personal interaction where the readers’ advisory librarians make suggestions which “address community needs because they are done by the people who know the community really well.” She also points out that “members of the community also know their librarians, which can make their hot picks more trustworthy than a reading list furnished by an anonymous critic.” (Hoffert, 2003)

Others, like Leigh S. Estabrook find that, “Exploiting user information from automated systems can allow us to develop targeted services. We can create ‘communities of interest’ by grouping our users. Then we can use the network to expand our services across space and time.” (1996) She imagines more personalized profiles of readers and that “As the library acquires titles that fit these user profiles (based on information they have provided or information culled from circulation patterns), an automated message could be generated to notify those readers. This same information about user interests could become the basis for organizing reader/user special interest groups (electronic or face-to-face).” (1996) Again, privacy concerns would need to be addressed here, but the idea of using technology including recommended systems to target readers with information specifically selected for them could be a great service to library patrons.

Allowing subjective reviews from library patrons into the catalog, of course, may lead to other problems. Even Amazon.com has seen abuses of its comment system when authors inflate their own works or try to downgrade their competition. In the library, would you filter the comments left by other readers to protect against abuse,

inappropriate material and bad advice? Or would that open the library up to charges of restricting free speech? Allowing comments and reviews would open up a very complicated can of worms. Even behavior-based input could be misused and manipulated, particularly if the system is using a very small data set to generate the recommendations. A parent might be very disturbed to see a controversial adult book listed among the recommendations for their child if safeguards are not set. Again, this may cause more trouble than is worth for the benefit it could bring to other users.

Moving Forward

In a recent panel of the library-based comic strip *Unshelved* (<http://www.overduemedia.com>), Tamara the children's librarian is asked by a young patron to tell him about the book he is holding up. She replies simply that "It's wonderful!" and is asked, "Can you be more specific?" She doesn't want to ruin the plot for him, and explains that she does not "want to be like those trailers that tell you everything about a movie." Tossing the book aside, his response is, "There's a movie? Great! I'll take that instead!" (Barnes and Ambaum, 10/30/04) While this is poking fun at what may in fact be a typical exchange, the fact is that many librarians are asked their opinions on books in the collection and asked to make recommendations on what people should read next. Librarians face questions of whether to offer only neutral, non-judgmental information about books, or provide more subjective reviews and connect readers with their peers through reviews, recommender systems or other methods.

Riedl and Jonst write that, "What's important from the customer's vantage point is what preference and product information they need to supply, and how recommendations

are presented to them." (2002, p. 14) Eastbrook writes, "To compete effectively, they must improve their ability to target service to users that will help them sort information wheat from chaff, find what is relevant and interesting, and meet their specific information needs." She adds that, "Our automated systems and networks give us the tools to provide such services. Can our policies and values accommodate them?" (1996) Libraries should be careful to respect patron privacy and to not eliminate the personal contact and connection that librarians can provide.

It is clear that finding effective ways to connect readers with the books in the library's collection is a worthy goal that will lead to better patron satisfaction and better utilization of the collection and resources purchased by the library. Automated recommendation systems alone are not enough to offer, but implemented carefully they can be useful tools, and they are already being successfully used by competing providers. As Wilson writes, "Librarians have never been at a more pivotal juncture to use the power of technology to promote their collections and a love of reading." (2001)

By making behavior-based suggestions, subjective reviews, and recommendations from other readers available to library patrons, libraries can augment the readers' advisory services, allow patrons to benefit from information gathered from outside of their immediate circle of contacts and help make the library's collection more accessible and responsive. If connecting people to readers just like them – people who have checked out or highly rated the same books – will make it easier for people to connect to the materials in the library's collection, then we should look seriously at the options available to help us facilitate these connections. Whether this should be done by purchasing "collaborative filtering software and allow our user communities to create

more locally relevant review databases” (Williams, 2002) or through some other means, it is clear there is much room for improvement in our catalog and other tools, and the competition is raising the expectations placed upon us.

Works Cited

- Barnes, B. & Ambaum, G. (2004) *Unshelved*. Accessed October 31, 2004, from <http://www.overduemedia.com/archive.aspx?strip=20041030> .
- Belcastro, P. (1995). A Matchmaking Tool For Readers. *Wilson Library Bulletin* , 69, p34-7. Retrieved November 19, 2004, from Library Lit & Info Full Text database.
- Bopp, R., & Smith, L. (2001). *Reference and information services*. 3rd ed. Engelwood, CO: Libraries Unlimited.
- Chelton, M.K. (1999). What we know and don't know about reading, readers, and readers advisory services. *Public Libraries*, 38(1), 42-47. Retrieved November 5, 2004, from HWW Library Lit & Info Full Text database.
- Conhaim, W. (2003). Let the Internet Choose your Reading. *Information Today*, 4(29), 32-33. Retrieved November 5, 2004, from HWW Library Lit & Info Full Text database.
- Estabrook, L. S. (1996). Sacred Trust or Competitive Opportunity: Using Patron Records. *Library Journal*, 121(2), 48-49. Retrieved November 10, 2004, from Business Source Premier database.
- Geyer-Schulz, A., Neumann, A., & Thede, A. (2003, Dec). An architecture for behavior-based library recommender systems. *Information Technology and Libraries*, 22(4). Retrieved November 10, 2004, from <http://www.ala.org/ala/lita/litapublications/ital/2204geyer.htm>.
- Geyer-Schulz, A., Neumann, A., & Thede, A. (2002). Behavior-Based Recommender Systems as Value-Added Services for Scientific Libraries. Retrieved November 10, 2004, from <http://www.wipd.ira.uka.de/~mart/presentations/nf02.pdf>
- Hoffert, B. (2003). Taking Back Reader's Advisory. *Library Journal*, 128(14), 44016. Retrieved November 5, 2004, from HWW Library Lit & Info Full Text database.
- Oder, N. (1998). Tacoma PL Links Catalog to Amazon. *Library Journal*, 123(7), 16. Retrieved November 5, 2004, from HWW Library Lit & Info Full Text database.
- Reference and User Services Association Collection Development and Evaluation Section Readers Advisory Committee (2004). Recommended readers' advisory tools: Collection Development and Evaluation Section (CODES) Readers' Advisory Committee. *Reference & User Services Quarterly*, 43(4), 294(12). Retrieved November 12, 2004, from Expanded Academic ASAP database.

- Resnick, P. & Varian, H. R. (1997, Mar). Recommender systems. *Communications of the ACM*, 40(3). Retrieved Nov 10, 2004, from <http://www.acm.org/pubs/cacm/MAR97/resnick.html>.
- Riedl, J. and Jonst, J. (2002). *Word of mouse [electronic resource] : the marketing power of collaborative filtering*. New York : Warner Books, 2002. Retrieved November 9, 2004, from San Jose Digital Library, <http://ebooks.sjlibrary.org/>.
- Rippel, C. (2003). What public libraries can learn from superbookstores. *Australasian Public Libraries and Information Services*, 16(4), 147(9). Retrieved September 5, 2004, from Expanded Academic ASAP database.
- Ross, C. S. (1991). Readers' advisory service: new directions. *RQ*, 30(4), 503(16). Retrieved November 12, 2004, from Expanded Academic ASAP database.
- Santa Clara County Library, Online Catalog. Accessed November 2004, from <http://146.74.92.11/ipac20/ipac.jsp?profile=scco>
- Tenopir, C. (2003). Running with the Amazons. *Library Journal* , 128(20), 34- . Retrieved November 19, 2004, from HWW Library Lit & Info Full Text database.
- Williams, J. F. (2002). 6 Hot Technologies With A Purpose. *Library Journal*, 127(2), 51. Retrieved November 1, 2004, from Business Source Premier database.
- Wiley, D.L. (1998) Beyond Information Retrieval: Ways to Provide Content in Context. *Database*, 21(4), 18-22. Retrieved November 15, 2004,
- Wilson, P. (2001). Reader's Advisory Services: Taking It All Online. *Public Libraries*, 40(6), 344-5. Retrieved November 5, 2004, from HWW Library Lit & Info Full Text database.