

Amazon: competition or complement to OPACs

Maja Žumer

University of Ljubljana, Slovenia

Introduction

Research (e.g. Borgman 1996, Bates 2003 etc.) repeatedly confirms that end-users find OPACs difficult to use. And recently an interesting observation was often mentioned when discussing OPACs: some users use Amazon as front-end to OPACs. They find a book they want to borrow in Amazon and then use ISBN to find the call number in the catalogue (see e.g. <https://intranet.lib.byu.edu/wwg/?p=13>).

This may not be the prevalent current practice, but it certainly points to a problem: average OPAC interfaces seem to be less user-friendly than Amazon at least to some users. Amazon will of course not replace libraries, but librarians must have continuous interest in and understanding of what users need and want and how they search for information. We need therefore to analyse what other services do better and try to implement appropriate changes in our own services. Some library systems have already made the first steps in this direction.

Amazon and libraries

Libraries have more in common with Amazon than just the fact that they all deal with books. There are reports of Amazon being used as book and media vendor for libraries, even the economy of this is being investigated (Orkiszewski, 2005). Libraries are selling their unwanted books through Amazon Marketplace and buying used books for their patrons instead of interlibrary loan. Quite often there are links to Amazon directly from library OPACs. All these developments are interesting and their impact both on library operation and user satisfaction should be studied.

But this paper will focus on two additional issues of importance of Amazon for libraries:

- Amazon as model for better OPAC interfaces and functionality
- Amazon as front-end of OPACs

Criticism of library catalogues

Recently a lengthy analysis of University of California bibliographic services was published (University of California, 2005). The report is very critical of current state of affairs:

- »The current library catalogue is poorly designed for the tasks of finding, discovering, and selecting the growing set of resources available in our libraries. It is best at locating and obtaining a known item«

- »Our users expect simplicity and immediate reward and Amazon, Google, and iTunes are the standards against which we are judged. Our current systems pale beside them«

This criticism sounds very harsh, possibly too severe. Why? Is it because libraries are often perceived as old-fashioned and not ready for the challenges of the modern society? Is it because libraries are not (any more) seen as the primary information providers? And, finally, do these opinions come from non-users?

But we have proof that library users find library OPACs difficult to use and try to avoid them. A quote from a blog (<https://intranet.lib.byu.edu/wwg/?p=13>) shows an interesting searching behaviour:

“First, I search for the book on Amazon.com. Every book always comes up on Amazon.com, while for some reason only about one out of every ten books comes up on the library catalogue. Then, I get the ISBN number off of Amazon. Then, since there is no option to search for a book in the library catalogue by ISBN, I am forced to put the ISBN on the ILL request form and request the book. They are always very good about sending me a judgmental email, telling me to stop requesting books which are already in the BYU library. Happily, they always send me the reference. Where they get it from, I don’t know, but until the catalogue is significantly improved, I am forced to continue harassing them.”

It is interesting to observe that this user is willing to invest a substantial effort into finding materials in the library while strictly avoiding to use the OPAC. It is an important lesson for libraries.

Today’s library users and potential users are:

- WEB-savvy
- aware of alternatives in information provision
- behaving as consumers and taking an active role in choosing the information provider
- have high expectations

User studies prove that there are some common requirements of library users:

- Simplicity
Simple Web interfaces have become the standard against which all other interfaces are judged
- Direct access to resources
Users are not interested in references, they need the resources fast and easily
- Relevance ranking
The default order of current catalogues is by order of cataloguing and that does not make sense to the users. Although usually not enough context is available to enable real relevance ranking, the order can be improved, particularly combined with meaningful clustering (see below)
- Help with failed searches

User studies show that getting no hits is frustrating to the users. And they like search engines »because you always get something« as a result.

- Browsing/navigation
Users want to be able to explore the bibliographic universe by following the links to connected entities
- Meaningful clustering of large result sets
Current catalogues often present large result sets as long lists of seemingly identical records. This is not user friendly and could be avoided using FRBR.

Even critics point out that even traditional OPACs have some advantages. They usually offer more than one search interface, thus enabling at least some degree of personalization. They are efficient in supporting known-item searching. Subject access (i.e. searching on a topic) is well supported in OPACs (although could and should be improved), as well as name authority control. In addition, library metadata are of relatively good quality and consistent.

What is offered by Amazon?

When asked about what makes Amazon interface so attractive, users most often list:

- Simplicity of the interface
- Availability of book-cover images
- Recommendations to buyers
- Reviews (both writing and reading)
- Result ranking
- Full-text searching

1. Simple interface

Under the influence of Web tools such as search engines many if not most users expect a very simple interface, enabling searching on keywords. Libraries in fact already offer this option under keyword searching. Users also dislike very much getting no hits as result, which means they are not concerned with precision.

2. Book-covers

Interestingly, the appearance of books was never in the librarians' focus. All features such as book cover colour and design have never been recorded in bibliographic records and they are not covered by any cataloguing rules. Yet these attributes are very important for known-item searching and (in FRBR terms) they support the 'identify' function. Realizing that, libraries now offer links to the cover image through, for

example, Amazon. We are not aware of any studies analysing the extent of use of this feature from library catalogues.

3. Recommendations

Amazon offers recommendations based on user behaviour. Typically, one is offered a list of books which were bought together with the book found or by people who bought the book in question. This is based on the assumptions that books bought by the same person are similar and that people buying the same books are similar. The concept is well known as similarity measures (using citations) in bibliometrics: bibliographic coupling and co-citation. Two documents with common citations form a bibliographic couple and two documents, cited together, are co-cited. In principle libraries could use the same concept using circulation data; the concern there is that there may not be the critical mass to enable such generalizations, unless used for large union catalogues.

4. Reviews

The possibility of writing and reading reviews has often been mentioned as an important feature of Amazon and other similar Web tools. Research would have to verify which aspect is perceived as the more important: to be able to write a review or to read others' reviews. Interestingly, Amazon users seem to write reviews only for bestseller, most other books have just one or two reviews.

Libraries are aware of the need of more active participation of users and many are already implementing or planning such features, often labelled Library 2.0.

5. Result ranking

Ranking of results is a serious weakness of traditional OPACs: the default arrangement is by date of cataloguing. While resembling the chronological order by publication date, this is by no means the order most useful or expected by the users. In most cases there is not enough context available to determine the most appropriate arrangement of results, but some experiments are starting. The area should be investigated further.

On the other hand sorting by popularity (according to circulation data) seems an appropriate alternative; again, the main concern is the lack of the critical mass.

6. Availability of content and full-text searching

Since users are looking for the content and not bibliographic records, direct availability and access to (digital) content (digitized or born-digital) is of the first priority. In addition, once the full-text is available, more searching tools can be offered. In addition to full-text searching (*Search Inside!*) two interesting tools are available from Amazon: *statistically*

improbable phrases (SIP) and *capitalized phrases* (CAP). The first are defined on the Amazon site as “the most distinctive phrases in the text of books in the Search Inside!™ program. To identify SIPs, our computers scan the text of all books in the Search Inside! program. If they find a phrase that occurs a large number of times in a particular book relative to all Search Inside! books, that phrase is a SIP in that book.” Capitalized Phrases are people, places, events, or important topics mentioned frequently in a book

Libraries are already enhancing their bibliographic records with full tables-of-content (Byrum, 2005) or other information available from the publisher.

Recommendations for future developments

Some libraries have already implemented some of the functions offered by Amazon and Google, but the vast majority are still waiting. There is no need for a revolution and total change. Libraries should keep and further develop existing services in which they already have an advantage over the competition, for example:

- Subject access. Librarians have developed tools such as classification systems and subject headings which enable access to materials when the topic is known. These tools and OPAC interfaces should be developed further to make searching easier for the end-users; some recommendations are listed in Bates (2003).
- Even current catalogue records contain information which enables navigation from one entity to another. Since exploration (browsing) is an important segment of searching behaviour, it should be facilitated by systematically offering links and relationships.
- Name authority control is gaining importance not any more as only the librarians’ tool for managing variant names and establishing authoritative forms, but as an important searching mechanism.

There are still many challenges ahead:

- The implementation of FRBR
 - For meaningful clustering of large result sets
 - For better exploration of the bibliographic universe and navigation between entities
- Multilingual and cross-lingual searching
- Avoiding failed searches with spelling control and presentation of alternatives
- Efficient and effective distributed/federated searching

There is probably no single simple and ideal solution for all, so there is no point in waiting for one. We need to start making smaller steps, but all leading in the same direction: providing a better service to our users. It is

better to make an occasional wrong step than stand still and disappear into disuse.

References:

Bates, MJ. 2003. Task force recommendation 2.3, research and design review: Improving access to library catalog and portal information. <http://www.loc.gov/catdir/bibcontrol/2.3BatesReport6-03.doc.pdf>.

Borgman, CL. Why are online catalogs still hard to use? Journal of the American Society for Information Science; 47 (7) Jul 96, p.493-503

Byrum, JD. Recommendations for urgently needed improvement of OPAC and the role of the National Bibliographic Agency in achieving it. 71th IFLA General Conference and Council, Oslo 14 – 18 August, 2005. (<http://www.ifla.org/IV/ifla71/papers/124e-Byrum.pdf>)

Orkiszewski, P. A Comparative Study of Amazon.com As a Library Book and Media Vendor. Library Resources & Technical Services, Jul2005, Vol. 49 Issue 3, p204-209, 6p

University of California Libraries. 2005. "How we provide bibliographic services for the University of California"
(libraries.universityofcalifornia.edu/sopag/BSTF/Final.pdf)