

Catalogue 2.0 and Bibliography 2.0: Collaboratively Created Structured Resource Lists and their Aggregation

Lobel Machala
National and University Library
Hrvatske bratske zajednice 4 p.p. 550, 10 000 Zagreb, Croatia
lmachala@nsk.hr

Krešimir Zauder
National and University Library
Hrvatske bratske zajednice 4 p.p. 550, 10 000 Zagreb, Croatia
kzauder@nsk.hr

Summary

Paper presents how web 2.0 and its tools affect traditional library services and products such as catalogue and bibliography. Using set of social tools and their integration into the library services the user may be given a new means for personalization of traditional library tools. A whole new range of possibilities and usage scenarios open up not only for citations but also for multimedia and hyper textual interactivity with bibliography lists. These tools for bibliographic record management became easy to use but, in the same time, powerful enough to be accepted by library professionals. The paper presents key issues and possibilities surrounding the catalogue 2.0 and bibliography 2.0, points out the differences and similarities between them and suggests the possibility of integration or a creation of a network of these tools.

Key words: Catalogue 2.0, Bibliography 2.0, Library 2.0

Introduction

“... looking is easy; finding what you want presents the difficulty!”
(Hildreth, 1995)

We are aware of definitive changes in our information environment in general and Web in particular as the latter grows from the network of essentially printed media in digital form to a digital medium that greatly differs from traditional ones and which is starting to realise its full potential. Some of the more obvious of these changes became transparent in roughly 2004/05. and were collectively dubbed and popularised as “Web 2.0”. At the core of the “2.0” concept, is a

move towards the social web, a move from a read-only medium to a read/write medium which brought a new paradigm into the architecture of a web page: a move from a page as a fixed document (a primarily printed construct) to a page as an application (a primarily digital construct). As we pair these changes with the growth of digitally available information, online identities, greater user participation, user generated content, ever increasing sizes of users' personal collections, new paradigms in web business models and design patterns and so on, it becomes obvious the new medium is having a great impact on information use and management as we understand it. In addition, one should not forget that we are looking at increased permanent connectivity via a plethora of devices that have increasing processing power and storage capabilities (Dempsey, 2009), some of which (i.e. e-paper, some smart phone uses) are specifically tailored to replace paper for reading texts but also in acquiring them.

In library world, the move of the library as an information space to embrace the new paradigm is frequently referred to as "library 2.0", a collage term denoting the libraries using the tools of the social web (blogs, wikis, social networking, tagging and personal collections, commenting and so on) and, more importantly, ideas, mechanisms and trends behind them. Some of the most important components of library 2.0 include increased user participation in the digital information space provided by the library, unique identification and tracking of users' actions and in communication with the library staff and other users online as well as increased librarian participation in online communities.

One of the most important mechanisms used to enable this is personalization, a creation of one's identity online in order to create one's own corner of the digital world. User might then use this online identity for activities such as communication in various contexts but also as a unique reference to a virtual information space unique for this user which might or might not be available to the general public. This space is then frequently used for the storage of digital resources. This way users are able to create own collections of resources in a certain context (e.g. books the user read or a personal bibliography of scientific articles). What should be noted is that users are frequently collecting not the full resources, but references to them. In other words they are structuring their own resource lists and dealing mostly with metadata. This makes sense in the web environment as resource lists also serve as direct access points to resources.

We will return to this in more detail in discussion of possibilities offered by catalogue/bibliography 2.0 as this idea is central to both. It is worth, however, to first explore what is "2.0" in this constructs.

Catalogue 2.0

"For at least the last 200 years, no service or image has been more closely associated with the library than its catalogue. Whether as a hand-written book, in card format, or as a digital entity, the catalogue both defined and represented the library. ... The whole point of a library was not just assembling the world's

knowledge, but assembling it in a manner which made it relatively easy to find, retrieve, and use.” (Kohl, 2007). This made the library catalogue an icon, the foundation of library identity. And, in the printed era, it has been so as the library functions and its position in the society have been devised and maintained by librarians. In today’s hybrid world, which shows rapid development towards a “mostly digital” one, library functions and its tools are changing rapidly and thoroughly.

We can differentiate three generations of online library catalogues. The first one resembled the card catalogue. The second, more sophisticated, one appeared in mid 80s and brought many new possibilities but with a high price in usability. In mid 90s an idea dawned for a third generation with increased functionality. Researchers (Hildreth, Borgman, etc.) pointed out their visions how new catalogues should look and feel like. Hildreth (1995) even called these catalogues E³OPAC, a name that represents: enhancing (functionality and usability), expanding (indexing, data records, collection coverage, “full-collection” access tool) and extending (through links, networks, and gateways to additional library collections, information systems and resources).

On the other hand, Borgman (1996) reports on poor functionality and difficult use of catalogues for over 20 years. Every change that was made on library catalogue interfaces was just a scratch on the surface and did not improve core functionality that would truly affect the users’ experience. There was increased need for the catalogues to become more user friendly by implementing natural-language searching, browsing, key ranked results, expanding coverage and scope, feedback methods (“more like this” or “not interested”), user-popularity tracking and all kind of different aids (spell corrections, synonyms, automatic term conversion etc.) and so on. However, except of some prototypes, the third generation catalogues did not appear till now.

In the meanwhile, web search engines developed what users missed in library catalogues. Web offered users easier and quicker (if not “better”) ways of finding information. Users got accustomed to natural-language searching and type multiple search terms without Boolean operators. They came to expect the same functionalities in library catalogues. Although users still see the catalogue as a trustworthy, well-organized and impressive tool (OCLC, 2005), the advanced features of catalogues are suitable more for the well-qualified librarian than for the average user. Today, a large and growing number of students and scholars routinely bypass library catalogues in favour of other discovery tools, and the catalogue represents a shrinking proportion of the universe of scholarly information (Calhoun, 2006).

While the features of the third generation OPACs are still being planned and sporadically implemented, ideas for some new features appeared out of web 2.0 pioneers such as del.icio.us and Flickr. The idea is that each user can create own collection of resources for future reference and that this collection is available online for the user to use from any device connected to the internet. Each user’s

collection is tagged via natural language keywords called tags to facilitate organizational scheme. The collections are interconnected so one user may access others' resources, add other users to friend lists, have his tags suggested from ones others used and so on. These tools came to be known as "social bookmarking" and are beyond this article. For discussion see Hammond et al. (2005); Macgregor, McCulloch (2006); Banek Zorica, Špiranec, Zauder (2007). This idea as an addition to library catalogue is the most central "2.0" component in "catalogue 2.0". Library may allow users to create profiles in order to personalise the catalogue and services. Using the profile a user may keep track of saved searches, save references to items to own collections, be referred to "recommended" resources, connect with other users and librarians, browse other users' (who may also be librarians) collections and so on. The combined actions of many users will then create a new layer on the catalogue: an access layer of user tags and collections may be used to both search and browse the library resources but also for analysis of users use and organization of resources (i.e. what do they read, what do they save for later reference and which terms do they use for organization). On this layer additional services may be built, such as automatic recommendation of tags, resources, and users with similar interests. Also, unlike other parties offering this kind of features, librarians are in a unique position of already possessing a large amount of quality metadata.

The most problematic part here is the software which needs to either be developed on top of already existent OPAC or implemented in existent commercial products. Breeding (2005) suggests that libraries will continue to use commercial library systems. The open source library systems usually cannot compete with the stability and support of the commercial ones which set high criteria. A possible development is in the direction of open source solutions which will provide additional possibilities to already present systems (Breeding, 2007). A live example is VuFind. Under the slogan "the library OPAC meets web 2.0", VuFind is a tool for creating library resource portals which support much of the features mentioned in this article. It can connect several commercial systems and using them create added value and pave the way for active users' participation in content and metacontent generation as well as for social networking and social bookmarking.

Bibliography 2.0

"Bibliography is about books", claims R. B. Stokes before the age of the dot.com. As Encyclopaedia Britannica puts it, the tasks for compiling bibliography consist of finding materials on particular subject, describing them item by item and assembling results entries into useful arrangements for reference and study. Today, when "Shift happens" (Fisch, 2008), the context of bibliographing changed rapidly. Changes in ICT made the realm of information access and control impossible to separate from the new technologies. Major reference works appear only online, most electronic publications are searchable in full-

text, publishing world is shrinking through consolidation and mergers, self-publishers are growing in numbers, and the list goes on (Parent, 2007). However, the core of the bibliographic work, gathering, assigning metadata and presenting, remains unchanged.

On the web, such activities may be noticed since the early days as various structured and semi-structured resource lists have been active and range from efforts such as Yahoo! directory or simple lists of “related links” to subject gateways such as Intute. These lists serve bibliographic functions and are often published by people without specialized skills in information organization who follow non-rigorous selection and organization procedures. Still, this publishing activity is fundamentally important since it structures information locally, creating a patchy network of secondary access points for human users but which are also frequently used by large-scale web search engines such as Google as input for their indexing and ranking algorithms.

Producing this kind of resource list is, in its essence, bibliographic work, albeit very often amateur in nature. The new possibilities of collaboration in the digital environment, whether via automatic connection of aforementioned constructs or via user contribution to the “bibliography” has led some researchers to call this type of work “collaborative bibliography” (Hendry, Jenkins, McCarthy, 2006). A large number of these constructs (the “related links” pages of many sites) have been made by individuals for use of others, and differ greatly in style and fullness of description while the structure is most frequently a flat list of resources. It is in this respect that the “2.0” paradigms make a difference from these types of “collaborative bibliographies” to “bibliography 2.0”. The “2.0” features are mostly the same as those displayed by catalogue 2.0, and they include a single system in which individual’s reference lists are tied to his or hers profile making the list an individual’s own personal collection (structured by tags) of references to resources. In this type of approach, the single system may provide the basic description elements, ensuring that it is uniform across users’ collections, and automatic metadata harvesting possibilities, ensuring metadata of some quality. The collaborative part in this respect is mostly gained through network of these collections manifesting itself through interconnectedness of user profiles, tags and resources. The main difference from the catalogue 2.0 is that user’s collections are not an additional layer on the catalogue of a library but that the service allows references to resources of certain types (e.g. scientific texts) without regard to ownership of those resources, which we think to be, as explained later in the text, to be more characteristic of the digital medium and expect it to be more prominent in future systems.

There are a few products developed and freely usable which would fit the description of a bibliography 2.0 tool. Examples include: LibraryThing (a service to help people have structured lists of books online) and CiteULike or Bibsonomy (services applying the idea of social bookmarking to scholarly texts). Some other products are primarily designed for personal desktop use, but fea-

ture an online component which facilitates remote synchronisation and collaboration between users. They share some similarities with bibliography 2.0 tools in that they support networked personal structured resource lists. An example of this kind of software is Zotero, a popular personal reference management tool.

Catalogue 2.0 and Bibliography 2.0: Unification possibilities

As more resources are accessible online in their entirety, the users' distinction between catalogue and bibliography begins to blur. The physical location begins to matter less in the digital world as users wish to satisfy their information needs as efficiently (or at least as quickly and effortlessly) as possible. In other words, in a world of ubiquitous openly accessible information sources, "information gateways" play an increasingly more important role to the user than "information warehouses", as the actual ownership of a resource begins to matter less and less to the user providing it's in open access. Although we currently live in a hybrid world, and users need to ask themselves "Can I find this in digital form?" and "Where can I access this?/Who will buy it for me?", given the amount of freely available information on the web and the "googlization" of the information world, it is no wonder a web search engine is frequently a typical user's first stop for solving his or hers information need, whether that engine is a viable choice as a tool for satisfying it or not. Also, it is to be expected that the amount of digitally available resources will only grow as various digitalisation projects bear fruit and as digital-only resources become the norm for some areas of human knowledge and activity and as the reading technology changes (e.g. e-paper) the current problem some have with reading long texts or screen might very well disappear bringing even more convenience to usage of digital resources.

Concerning the user, he/she is more and more frequently walking in the shoes of the information expert as user collections of various information constructs (textual documents, photos, multimedia files, various reference lists such as favourites, reading and listening lists, references to scholarly work and so on) grow to the size of that previously owned only by institutions and select individuals and as new types appear. In the digital world, which is bombarding the user with the amount of accessible information, such collections and listings might be necessary as users need to model their own information environment to suit their needs. User collections of references to items are prominent especially due to the nature of the medium where having a unique reference to a resource is frequently like having a resource itself (at least where time needed to access the resource is concerned; legal, security and preservation issues aside) and where resources' content is changing fluently, not in discrete editions (e.g. a link to a wiki page vs. a locally stored wiki page).

What we are dealing with in both catalogue 2.0 and bibliography 2.0 is a system of networked tools that provides additional value as it creates a layer of access and automatic reasoning (e.g. "recommended resources" feature, gained through

a network of users, tags and resources) which, unlike traditional web search tools, has user decision making at its core and supports additional discovery mechanisms such as serendipity in browsing (not unlike traditionally browsing the shelves but with different placement), identification of users with similar interests (which then function as a recommendation mechanism), detection of popular items and so on. One should bear in mind that while most users are amateurs when it comes to information organization, and may need assistance in the field, they may very well be excellent subject experts in the field from which they are collecting resources.

Personalised structured resource lists as online library services?

Both catalogue and bibliography may be viewed as structured resource lists mainly differing in body of written knowledge from which the references are derived, in fullness of description and in organizational approach. A significant difference in the printed world is also that the catalogue, unlike bibliography, may be used to facilitate access to some possessed instance of the resource. However, as these tools move online some of the distinctions begin to blur as both may be used for a direct access to resources. The differences organizational and access possibilities are also diminished as new ways of structuring the lists may be done automatically and the structure of a list may change on demand for a single user. One of the most significant abilities of the digital medium in this respect is to tie these tools together and provide additional layers of possibilities for the users of the system (both library users and librarians) without necessarily modifying the original data.

From the users' perspective, personal information management tools are needed more as the need for modelling one's own information space raises. No single system may currently serve all the user's needs, mainly due to the fact that new types of information constructs users need to manage are still rapidly appearing and it is questionable which should be included in the library systems. However, literature lists in the broad sense are surely one of the needs of the new generation and libraries are in a unique position of already possessing quality metadata and employed information experts which could provide an enviable basis for quality implementation of the "user collections" or "personalised structured resource lists" idea. By mashing up these tools and approaches (catalogue, bibliography, personalised resources lists, automatic metadata harvesting, user networks and so on) a library as an online information space may be a lively place which can attract users, a tool to be used for various user analyses and, in general, a construct which suits the new media.

Benefits and problems of personalised structured resource lists

Using these tools, the library can offer the user a service for personalising his library information space or information space in general, depending on the service, the community with an additional access layer to library or other informa-

tion resources and the librarian with the data about users' use and organization of resources. In addition these tools may make a digital library a lively place with a (hopefully) active community, make the library valued beyond "literature warehouse" and serve to keep the librarian in the position of an information professional both in printed and digital worlds.

There are a few problems in implementation of described services in libraries. These systems require an active development which requires funds and teams able to pull it off both conceptually and technically. Also, collective intelligence can be achieved only when a critical mass of participation is reached. There has to be sufficient number of frequent users using the service to enable the service reach its potential and become valuable (Anderson, 2007). Some researchers even claim that library communities are too small to achieve that critical mass (Wenzler, 2007). In addition, while libraries are always late in use of new technologies, other hand users move very quickly toward another source if they are not instantly gratified. Another problem is motivation of users to participate actively in library catalogue. Do we have users that are willing to help altruistically or will they participate only when they can also fulfill their private incentive. This problem is somewhat alleviated by the fact that this systems may be built in a way that implements collaboration on the level of aggregation, so that the service gains value if even it has a community where each user works for herself only.

To sum up, these services are currently out of the scope for many libraries due to lack of funds, active users and/or teams who can pull it off both conceptually and technically. One of the solutions to the problem is in products which may be implemented as components on top of current library solutions such as VuFind, which are developed by a certain community and then released for use and customisation by others. This may not solve the problem of attaining a critical mass of users, but it will at least lessen the amount of resources needed to implement the service and serve to attract the users with the new possibilities offered and with being in trend.

Conclusion

A modern day library functioning as storage is but an information island outside the network. This doesn't mean everything should be online, but it does mean that, in the internet era, library websites (or hubs to libraries) should be lively places as they are the face of the library that more and more users will first see and use. Besides, a social library website presents the librarian with many new possibilities.

A modern day user quite frequently does not need just the hard copy of written knowledge but a reference to a recommended piece to retrieve for which he or she may or may not need the library. Given the amount of written information around, a social approach may be an interesting bibliography tool for dealing with online resources: access, evaluation, personal information tool.

Given the number of resources and current search problems, two points are to be made. First, a help in the selection of quality resources plays an important role as current web search tools are great for known-item retrieval but subject based searches are much more problematic. It is here that the user will need most help, whether via direct tutoring or advice or via resources (subject gateways, bibliographies) either specially prepared by professionals, gained through “collective intelligence” or compiled with the combination of these approaches. Second, as users’ collections continue to grow both in number and in size, they offer an important device for the single user, who has a device for keeping his collected information resources in one place, for the community, which can benefit from another type off access layer, and for the librarian, who can benefit from having data about users’ use and organization of resources.

In both points, information literacy is of paramount importance and presents one of the possible challenges for libraries in the future: if these types of activities are done in the (digital) library this presents an opportunity for the library to educate the community. Also, given the quality of library metadata and staff already employed on “quality control” in various guises, librarians are uniquely poised to provide metadata of greater quality than most other institutions and to add value to users’ collections working as behind-the-scenes information professionals.

The success of catalogue/bibliography 2.0 depends on both parties involved: libraries need to design social tools that are attractive, intuitive and useful, and users need to contribute and use the services provided by the catalogue. In order to realize their true potential in the digital world and, libraries need to bring convenience, trends and quality close together. To put it figuratively: Libraries should be more “tree-focused” rather than “forest-oriented” in developing software tools for users. In the end, well tended trees will produce fine forests.

References

- Anderson, Paul. What is Web 2.0? : ideas, technologies and implications for education. // *JISC Technology and Standards Watch*. February, 2007. URL: <http://www.scribd.com/doc/300024/What-is-web-20-Ideas-technologies-and-implications-Paul-Anderson?autodown=pdf> (20.06.2009.)
- Banek Zorica, M., Spiranec, S., Zauder, K. Collaborative Tagging: Providing User Created Organizational Structure for Web 2.0 // *The Future of Information Sciences : INFUTURE2007 - Digital Information and Heritage*. Sanja Seljan and Hrvoje Stancic (ed). Zagreb : Department of information sciences, Faculty of humanities and social sciences, 2007, pp 193-202
- Bates, Marcia J. An exploratory paradigm for online information retrieval. // *Intelligent information systems for the information society : proceedings of the Sixth International Research Forum in Information Science (IRFIS 6), Frascati, Italy, September 16-18, 1985*. / Brooks, B. C. (ed.). New York: North-Holland (Elsevier), 1985, pp 91-99.
- Borgman, Christine L. Why are online catalogs hard to use? : lessons learned from information retrieval studies. // *Journal of the American Society for the Information Science*. vol. 37 (1996) no. 6 ; pp 387-400.

- Breeding, Marshall. The new landscape of the automation business. // *Computers in libraries*. vol. 25 (2005) ; pp 40. URL: <http://www.librarytechnology.org/lgtg-displaytext.pl?RC=11590> (20.06.2009.)
- Breeding, Marshall. The sun sets on Horizon. // *Computers in libraries*. vol. 27 (2007) ; pp 38. URL: <http://www.librarytechnology.org/lgtg-displaytext.pl?RC=12736> (20.06.2009.)
- Calhoun, Karen. The Changing Nature of the Catalog and its Integration with Other Discovery Tools : final report for Library of Congress. March 17, 2006. URL: <http://www.loc.gov/catdir/calhoun-report-final.pdf> (20.07.2009.)
- Fisch, Carl. Shift Happens : did you knew 3.0. 2008. URL: <http://www.youtube.com/watch?v=jpEnFwiqdx8> (29.08.2009.)
- Gatenby, Janifer. Today's information consumer tapping into international library services: making it a reality. November 2006. URL: <http://www.oclc.org/content/1400/pdf/article-informationconsumer-internationallibraryservices.pdf> (20.08.2009.)
- Hammond, T. et al. Social Bookmarking Tools (I) : a general review. // *D-Lib Magazine*. vol 11 (2005), no. 4. <http://www.dlib.org/dlib/april05/hammond/04hammond.html> (07.08.2009.)
- Hendry, David G.; Jenkins, J.R.; McCarthy, Joseph F. Collaborative bibliography. // *Information Processing and Management: an International Journal*, vol 42 (2006), no 3.
- Hildreth, Charles R. Online catalog design models: are we moving in the right direction? : a report submitted to the Council on library resources in August 1995. New York : Palmer School of Library and Information Science, Long Island University, 1995, URL: <http://myweb.cwpost.liu.edu/childret/clr-opac.html> (10.08.2009.)
- Keen, E. Michael. Designing and testing an interactive ranked retrieval system for professional searchers. // *Journal of Information Science*. vol. 20 (1994) no. 6 ; pp 389-398.
- Kohl, David F. Alas, poor catalog, I knew thee well---. // *Journal of academic librarianship*. vol. 33 (2007), no. 3; pp 315-316.
- Macgregor, G., McCulloch, E. Collaborative tagging as a knowledge organisation and resource discovery tool. // *Library Review*. vol. 55 (2006), no. 5; pp 291-300.
- Merčun, Tanja ; Žumer, Maja. New generation of catalogues for the new generation of users : a comparison of six library catalogues. // *Program: electronic library and information systems*. vol. 42 (2008), no. 3; pp 243-261.
- OCLC (2005), Perceptions of libraries and information resources : a report to the OCLC membership. Dublin, OH : OCLC, 2005, URL: <http://www.oclc.org/reports/2005perceptions.htm> (01.07.2009.)
- Parent, Ingrid. The Importance of National Bibliographies in the Digital Age. // 73rd IFLA General Conference and Council, 19-23 August 2007, Durban, South Africa, URL: <http://www.ifla.org/iv/ifla73/index.htm> (21.06.2008.)
- Stokes, Roy B. Bibliography. // *Encyclopedia of Library and Information Science*. 2nd ed. New York : Marcel Dekker, 2003.
- Wenzler, John. LibraryThing and the Library Catalog: Adding Collective Intelligence to the OPAC. // A Workshop on Next Generation libraries held in San Francisco on September 7, 2007, URL: <http://online.sfsu.edu/~jwenzler/research/LTFL.pdf> (05.08.2009.)