IFLA
Gothenburg
2010

Open access to knowledge
– promoting sustainable progress

From Dominance to Decline? The Future of Bibliographic Discovery, Access and Delivery

Martin Flynn
Head of Information Services
Victoria and Albert Museum
London, United Kingdom
m.flynn@vam.ac.uk

Meeting: **71. Art Libraries**

Executive Summary:

Library catalogues have been moving from a position of dominance to one of decline in bibliographic discovery, access and delivery in recent years. Their contents are largely invisible to search engines. The process of cataloguing is complex and difficult to master and users struggle to understand catalogues and therefore take full advantage of their added value elements. There is an extraordinary amount of duplication of effort in catalogue creation across the sector and considerable doubt about whether the current model of localised catalogues is sustainable. In general catalogues tend to be automated versions of their printed card predecessors so have not incorporated the enhanced features users have come to expect in a digital environment. They also tend to concentrate mainly on printed materials to the exclusion of digital formats. Endemic problems of catalogue, such as confusing search methods, poorly organised results and search dead-ends, have still not been addressed on any perceptible scale yet these failings have been tackled by search engines. Delivery of the content of search results either through inter-lending of actual books or the provision of digital facsimiles continues to be a vastly under-exploited service. The proprietary library management system sector is a relatively insignificant player in the global IT market and therefore has very limited investment capital. However, much of this capital is squandered on customising their products for the benefit of librarians rather than their users.

Introduction

This paper aims to promote discussion of the strengths and weaknesses of traditional library cataloguing policies and practices and to identify some potential improvements for the future. Four major recent reports on the future of library catalogues were reviewed. These were produced by some of the

key international players in the library world, namely the Library of Congress (LOC), the National Library of Australia (NLA), the Research Information Network (RIN) and the Online Computer Library Center (OCLC). These reports are referenced below. The conclusions of all of these reports are remarkably similar and could provide libraries with some practical guidelines for improvement which enjoy a broad measure of international support in the library world.

Background

For over a hundred years library catalogues have led the way in supporting unmediated information discovery and access and, until the mid-1980s, they had few substitutes. The emergence of online information services followed by the internet has seen the library catalogue move from dominance to decline. Implementing some radical changes should help to restore its critical importance.

Challenges

The production of library catalogues is both labour-intensive and expensive. The systems on which they are based are extremely complex and difficult to master, even for librarians. Users rarely take advantage of more than a fraction of the added value elements of catalogues. Cataloguing also involves enormous duplication of effort across the library sector throughout the world. Yet less than 5% of researchers and students choose library websites as their first point for beginning a search.   Library catalogues are largely invisible in the arena in which these users conduct their information enquiries.

The contents of library catalogues are largely invisible to search engines. Because libraries were early adopters of digital technology their catalogues also tend to be automated versions of manual systems rather than products which fully absorb the potential of automation. Catalogues also do not generally include the enhanced features which users have come to expect from other search mechanisms. Catalogues still tend to concentrate almost exclusively on books and other printed materials, largely ignoring the vast range of information output in digital formats.

In spite of cataloguers high professional standards of information organisation and retrieval they are battling with the contemporary phenomenon of 'The Principle of Least Effort' – people do not just use information that is easy to find, they even use information which they know is of poor quality because it requires little effort to find.

At an early stage in the development of online library catalogues Charles Hildreth identified many failed searches, frustrating navigation, confusing search and retrieval methods and poorly organised search results as endemic problems of library catalogues (http://tinyurl.com/l6f22h) In the meantime

these failings have been addressed by web search engines but, unfortunately, not by library catalogues.

Inter-library loan and digital delivery services are vastly under-exploited, particularly given user expectations in a 'buy it now, get it now' world of instant access to electronic resources. Vast quantities of library materials remain unused when they could be either physically loaned or their contents digitally delivered with ease.

There are considerable opportunities for harnessing the untapped potential of library management systems and add-on products to improve our users' experience of our catalogues.

There is an urgent need for national and international leadership on these issues to produce a new integrated model of cataloguing fit for an online environment. This needs to be underpinned by rational, objective and radical debate on cataloguing issues within the library profession.

The recommendations of the reports reviewed are both radical and challenging and may initially appear to be totally unpalatable to professional librarians. However, considerable weight needs to be given to the fact that these have been produced by some of the leading organisations in the library sector and the remarkable degree of unanimity in their conclusions. Their suggestions are essentially aimed at improving the relevance and value of our catalogues to users and ensuring a prime position for libraries in a networked environment.

Rethinking the Model

There are convincing arguments in the reports for abandoning local catalogues in favour of unified catalogues based on sectoral, national, trans-national or subject parameters which would contain significant enhancements, have greater visibility in popular networked environments and eliminate inefficient duplication of effort. However essential this approach is it is unlikely to be realised in the short-term. However, the following changes of approach, which enjoy a significant level of support from the quoted sources, are probably achievable in the medium term. The recommendations listed below are prefaced by key quotations from the original reports reviewed.

Users

A large and growing number of students and scholars routinely bypass library catalogs in favor of other discovery tools and the catalog represents a shrinking proportion of the scholarly information universe… users don't get the idea of the catalogue, they just want results…. less than 5% of users begin their search with library web pages (LOC)

Conventional libraries form a diminishing part in what is now a global networked information environment, based around the web. Global services such as Google Books, Google Scholar and Open Library give users free access to a huge collection of resources, including bibliographic data… much research focuses on information sources (e.g. books or newspapers) and systems (e.g. catalogs) rather than on the needs, motivations and behavior of information seekers (OCLC)

An unranked search result that produces more than 10 results is annoying; an unranked search result which produces more than 100 results is virtually useless (NLA)

Libraries should define their user groups, their needs and those of any potential new groups. The fast turnaround and delivery of library materials to the users should be the standard of quality service, not the fullness of catalogue data. Catalogues should implement interactive features such as user tagging, comments, reviews and recommendations.

Cataloguing

Library of Congress Subject Headings require too much behind-the-scenes understanding to be useful….investing in automatic classification rather than LCSH may get you something more competitive in the Google world and get better subject access too…subject access will be and has already been marginalized by keyword searching…today's catalogs are put together mainly by humans and….this approach doesn't scale…we should be building or expanding the scope of catalogs by using automated methods… the pace of standards development and implementation is painfully slow, because the process requires competitors to collaborate. When standards are approved compliance is voluntary, so the community ends up with many variants of the same standard (LOC)

In a digital environment, top down approaches, where expert cataloguers decide what a book is about, and place it in a pre-ordained scheme of knowledge, may be less appropriate: users can reshape pre-ordained categories in ways that they find more useful for their own purposes…how libraries choose to organize their data and collections has no canonical authority: it is just one option among others (OCLC)

The greater proportion of the material to which libraries provide online access under licence agreements, are not included in their catalogues….our key finding is that the current arrangements for producing and distributing bibliographic data for both books and journals involve duplications of effort, gaps in available data, and missed opportunities (RIN)

As cataloguing has grown, so too have the number and complexities of standards we expect cataloguers to know well...'it takes two years to make a good cataloguer' – and that time is on top of formal study... Google Books and Google Scholar provide access to lots of resources without having ever catalogued them, and this is an approach we need to consider... It seems sometimes that as cataloguers we base our professional pride on being able to find out what we want in a forest of unfriendly systems.... Just as library catalogue systems aren't very user-friendly, neither are cataloguing tools..... If the library catalogue ...could reference some sort of rating system it would be much better. People need to go to Amazon to see what books are most highly recommended.... LibraryThing now has more than 16 million tags applied to its bibliographic records and libraries should think of linking with LibraryThing to use their impressive collection of tags... Information literacy is also harmful because it encourages librarians to teach ways to deal with the complexity of information retrieval, rather than try to reduce that complexity... libraries should create systems that eliminate the need for instruction... Subject guides and tags, are a way of providing basic reference help in a generic way – so we can tag items as suitable for beginners or domain experts ...subject guides in particular would make a catalogue a preferred destination for many users (preferred over Google that is) because it would offer them an authoritative introduction to the topic and links to follow for more information. Some of these links would take searchers on to online resources such as Wikipedia and specialist web pages, but other would be to resources available from their library (NLA)

Catalogue records should be simplified to a set of basic elements to support discovery, browsing, identification, delivery, resource sharing, linking and inventory control. As much as possible, libraries should obtain or re-use data available at the point of selection, or automatically generate this information. Manual data creation should be reserved for ordering, receiving, claiming and cataloguing for those situations in which it is the only viable approach. Local customization and record editing practices should be identified and eliminated in favour of accepting as much cataloguing copy as possible without review or modification. Libraries should abandon the attempt to do comprehensive subject analysis manually with LCSH in favour of subject keywords. Customised cataloguing effort should be concentrated on rare and unique special material.

Catalogues

Ideally research libraries should retire their own catalogs and concentrate instead on segments of users… the demand for the product and the capacity to produce it are out of balance… librarians are beginning to question the

prevailing model of catalog creation and maintenance that is, replication of data in shared cataloging systems and in thousands of local catalogs (LOC)

There are more than 160 university libraries in the UK. RIN recommends a shared catalogue for all of these with dynamic links to local holdings…catalogues seldom include records of the contents of the institution's repository…both Copac and OCLC have announced that they are making use of a Google API to enable users to link from an OCLC or Copac record to the full text (or part of the text) made available through Google Books (RIN)

Libraries should enrich the catalogue with services such as 'more like this', 'get it' options, new book lists and reviews. They should enable best-match retrieval, avoiding search dead-ends and use 'see also' and 'see instead' suggestions. Relevancy ranking should be provided for search results.

The discovery of books, journals, journal articles and digital resources should be federated. The user should be linked to full text wherever possible. The catalogue should also provide summaries and excerpts. Summaries should be highly visible, ideally in the search results, in the form of a short blurb as in Google. Catalogues should understand the common use of terms and weigh results towards these.

It should be possible to refine a search through faceted navigation, filtering your results in multiple ways. Libraries should develop recommender systems like those provided by Amazon or at least link to the latter. Catalogues should link to LibraryThing to take advantage of their 16 million user generated tags.

Libraries should also develop online subject guides as an authoritative introduction to a subject with links to follow for further information such as Wikipedia, specialist web pages and library resources. Items should tagged according to their level of complexity e.g. suitable for beginners, domain experts etc. Links to citations and reviews should be incorporated. Catalogues should assist the searcher by providing hints and suggestions in a commercially disinterested way.

Remote Document and Content Delivery

Will interlibrary loans be based around local catalogs or union catalogs such as WorldCat and what is the role of on-demand services like salesforce.com? (LOC)

In the US ILL accounts for only 1.7% of circulation and Dempsey suggests: 'we are not doing a very good job of aggregating supply (making it easy to find and obtain materials of interest wherever they are). The flow of materials from one library to another is very low when compared to the overall flow of materials within libraries...the mechanism which allows both parties (libraries and borrowers) to benefit from the wider circulation of materials can be

constructed using an operational model loosely based on the NetFlix approach – posting or streaming DVDs and Blu Ray (NLA)

Libraries should review their policies on inter-library loans and re-assess their role in both digital delivery and more general lending of stock. Current technology allows for cheap and immediate delivery of facsimile content but this service is far from the norm in the library world.

Serials

For serials, librarians..may be the only users who think in terms of the container [the serial title] (OCLC)

Readers of journals are less interested in information about journal titles than in getting direct access to the text of the articles.   Metadata in library catalogues typically relates to titles and holdings, rather than specific articles… Users find metadata about journal articles through…abstract and indexing databases, publisher websites, Google Scholar and so on…catalogues rarely provide any information about scholarly journal articles, the single most important category of information resource for researchers… a recent survey found only ten small not-for-profit publishers who did not make their journals available online. Such journals are not now attractive to authors; for articles not exposed to the web are much less likely to be read and cited (RIN)

Research libraries need to take user demand for serial content more seriously and incorporate access to both article references and content into their catalogues.

Library Management Systems

In the library management system market ….there are few vendors, poorly capitalized, and libraries are a small and demanding market with, relatively speaking, little to invest in new ventures…'libraries want a ton of customization: this is ridiculous and must stop'…library management system vendors should …not ask librarians what systems should do, but find out what libraries need to do for their users (and forget the long enhancement lists from librarians) (LOC)

Libraries would spend years asking vendors for systems that solved our problems but did little to serve our users (NLA)

The focus of library management system development should shift to being user rather than staff orientated and a cross-sectora,l international, co-operative approach would support this.

Sources

1.1. Calhoun, K ' Changing the nature of the Catalog and its Integration with Other Discovery Tools', Library of Congress, 2006 (tinyurl.com/3dfpjt)

1.2. Calhoun, K, Cantrell, J, Gallagher, P, Hawk, J ' Online Catalogs: What Users Want and Librarians Want' , OCLC 2009 (tinyurl.com/yaonxwb)

1.3. Chad, Ken 'Creating Catalogues: bibliographic records in a networked world', Research Information Network 2009 (http://www.rin.ac.uk/files/Creating_catalogues_REPORT_June09.pdf)

1.4. Dellit, A, Fitch, K 'Rethinking the catalogue', National Library of Australia 2007 (**http://tinyurl.com/y8lfxf9**)