



Uniting the corpus of our collections through visualisation

Alex Byrne
State Library of New South Wales
Sydney, Australia

Session:

139 — Treating print in an increasing digital collection: issues, dilemmas and directions — Acquisition and Collection Development

Abstract:

This paper explores the challenges of visualising digital and physical collections to enable discovery and use with reference to new models for academic and public libraries as well as the particular challenges faced by libraries with great heritage collections such as the State Library of New South Wales. That Library is renowned internationally for its unique collections on Australia and the surrounding region. It and its international peers face the challenge to present their heritage collections, their digitised versions and their contemporary print, media and born digital materials to as a discoverable corpus for researchers, students and the public.

The physical corpus

At least since the fabled Great Library of Alexandria, libraries have long been assessed and compared in terms of their collections: how big, how old, how many incunables, how extensive, how specialist, and so on. Applying these Alexandrian measures, we view in awe the vastness of the holdings of the Library of Congress or the British Library, the depth of the Vatican Library, and the particularity of the Süleymaniye Library of Islamic manuscripts in Istanbul. Our mental images extend from the beauty of the famous Benedictine Library at

Melk in Austria to the serried rows of bookshelves in a great university library or behind the scenes in a major depository library such as the Deutsche Nationalbibliothek.

At the State Library of New South Wales, the premier library for the colonial history of Australia and its region, such measures are impressive. While those familiar with the library will immediately think of the books surrounding the grand Mitchell Library Reading Room, behind the scenes there are many stack floors and an external store, 25 km away. Our collections include:

- 5.5 million items
- 1.1 million photographs
- 11.2 linear km of manuscripts
- 234,000 prints, drawings, paintings and maps
- 114,000 architectural plans
- talking books, postage stamps, coins and books
- assessed value of AUD2.142 billion
- items significant to Australia's history, including 9 of the 11 known First Fleet journals and Joseph Banks' Endeavour journal
- the first book ever printed in Australia (*NSW General Standing Orders*, 1802) and the first Australian newspaper (*Sydney Gazette and New South Wales Advertiser*, 1803)
- the death mask, hat, pipe and pen of Henry Lawson, a major Australian writer

Laid end to end our shelves would stretch 120 km from the Library in the city centre to Mount Victoria, the pass in the mountains which it took the early colonists 25 years to cross. The collections grow by some 2 km a year, two lengths of the Sydney Harbour Bridge.

Housing collections of that size and growth rate presents a continuing challenge, one we share with many other great libraries – and, of course, our GLAM friends in galleries, archives and museums. To better store our collections, preferably on site, we are considering the use of a dense automated storage and retrieval system, emulating the University of Chicago among others, in addition to our use of compactus in staff-only stack areas.

Comprehending such extensive and varied collections, raises the question of how we and our clients can visualise them. How can we mentally ‘see’ collections spread between open access shelving, traditional stacks, dense storage and automated systems? And, how can we meaningfully comprehend their cultural and informational content across the myriad formats and topics represented in extensive collections?

The emerging digital corpus

Our challenge does not stop with our physical collections. As we digitise our heritage collections at the State Library of New South Wales, we will create some 12 million images or page views in addition to the approximately 500,000 we have at present. And, as we begin to capture born digital material relevant to our jurisdiction in the State of New South Wales, we will build up enormous datafiles which we will need to preserve indefinitely just as we preserve the paper, audiovisual, pictorial and other items which we have traditionally curated.

That digitalia is by its nature intangible. It consists of bits and bytes stored on transitory media such as CD-ROMs, tape or disc drives, or, increasingly, in the cloud. While digital

items are real and able to be viewed, manipulated, repurposed, they do not have a comprehensible physical existence which can be seen or touched. In fact, they frequently can be viewed only when the output is generated from stored data and can change momentarily as we experience computer games, digital artworks and other temporary media.

Visualising an aggregation of intangible and evanescent items as a collection is challenging and presenting them together with our physical holdings as an integrated corpus of cultural and informational materials stretches our imagination. But, as our digital collections grow, we are misportraying our collections if we do not assist our clients to visualise them in meaningful ways and in ways which join them to our physical collections to form a united corpus.

The challenge of visualising collections

So, how might we and our clients visualise those collections? And, even more vitally in this digital age, how can we bring together the tangible and the intangible to enable our collections to be visualised coherently?

These questions go to the heart of our understanding of twenty-first century libraries as collecting institutions. They are integral to the challenge of moving public identification of the library/bibliothèque (bibliotek/biblioteca/bibliotheek/...) beyond the physicality of the book/liber/ βιβλίο (biblio). Solving the question of how we and they understand and visualise 'collection' has become central to our conceptualisation of contemporary and future libraries.

There are many contemporary arguments in favour of maintaining libraries. They include the library as the 'living room of the community' (Haskell 2010 for example), as a 'third space' which is welcoming and available to all (Bauer 2009 for example), as a 'learning commons' which facilitates interaction and both collaborative and individual study (Holmesglen 2009 for example). All are valid and their emergence testifies to the strength and adaptability of the idea of the library. But access to information and the preservation and transmission of culture must remain central. Without those core purposes, the library becomes just another space. A library may be without books but it cannot be a library without being a vehicle for transmitting knowledge. It is therefore crucial that we develop ways of describing and visualising the body of knowledge that a particular library transmits, in short its collection.

That collection might consist of items of enduring value which would never be discarded, as in the heritage collections of a library like the State Library of NSW. It might include items which will be held only as long as they are current, as in the reference tools of a special library. And they might include materials held only while they are of interest to and used by clients as in the regularly refreshed collections of a public library. The items might be in physical formats or digital. Some might be accessed by subscription to ejournals or aggregations or on a pay-as-you-go basis as ebooks often are, changing as vendors add or subtract content and with access limited by licensing conditions. Regardless of these considerations, the collection remains a collection in that is intentionally created. It is the processes of selection – and de-selection – coupled with organisation which make a collection.

The idea of collection has thus expanded beyond the narrowly defined and easily visualised arrangement of volumes or other items on shelves to encompass materials that are neither tangible nor able to be seen and which may be ephemeral. However, if we are to understand them as elements of the resources, the collection, which we offer to our clients, then we should be able to visualise them as a collection.

Traditional modes of visualising collections

Leaving aside earlier arrangements of clay tablets and papyri, and the particular needs of manuscript, numismatic and other special collections, the most immediate way of visualising a library collection is in an arrangement on bookshelves, whether the elegant shelves of Melk or Trinity College in Dublin or the utilitarian steel shelving of contemporary university and public libraries. Offering strong mental images, this physical arrangement is usually shaped by classification or other ordering systems to approximately collocate items on related topics. It supports the much prized practice of browsing along the shelves to instantaneously judge the usefulness of items and to make serendipitous discoveries. Even though a moment's thought exposes the weaknesses of browsing – not held, on loan, in use, missing, online – the practice is very popular and its loss through moving collections to closed access raises vociferous objections (Smith 2011 for example).

With very large, dispersed collections and closed stacks comprehending collections as shelf arrangements breaks down as the collections cannot be easily viewed or browsed as a whole even if it would be theoretically possible to walk the stacks. With an automated storage and retrieval system it becomes impossible to enter and view the items sitting in the system's totes even though the items are tangible and can be seen and handled when

retrieved from the system. With digital materials that becomes impossible, the items only becoming viewable when presented on screen, printed or converted to an audible or other format. Visualisation, in the narrow sense of 'being able to be seen with one's eyes', to stand back and 'see' the collection as one can with the great wall of books in the Library at Delft Technical University or the dramatic King's Library in the British Library, thus becomes impossible.

We must turn to other modes of visualising collections. Moving beyond simple arrangement on shelves, the most fundamental way in which we can visualise a collection is via bibliographic description. The sum of the bibliographic records and holdings statements offers an image of the collection which can readily include items sitting in stacks, remote locations, closed stores, automated library retrieval systems or even digital materials in the cloud. Bibliographic visualisation can be via an inventory in accession order, a shelf list in classification order, author-title listings, sorted by format or age, highlighted by popularity, or in any other way that the discovery system will support. It is limited only by the metadata which is available and the capabilities of the system to exploit it.

Bibliographic visualisation, however, suffers from a fundamental flaw: it substitutes an artefact, the bibliographic record, for the desired collection item. That record is a tightly defined construct which palely reflects some key characteristics of the item. Bibliographic records can conjoin items irrespective of format so that ejournals can be grouped with eBooks, printed books, maps, compact disks, etc but they offer only very limited glimpses of the items. For example, although a standard MARC record may include spine height, it does not convey the physicality of a book: its mass, colour, texture, smell, the quality of its paper,

its typography, its binding. Nor does a record attempt to convey the ideas in an item and, far less, to evaluate their quality. The record aims simply to provide sufficient information to enable discovery and a preliminary assessment of relevance and usefulness.

A catalogue of bibliographic records can represent a collection, and a union catalogue such as Libraries Australia or WorldCat can represent the intersection of many collections and can bring together holdings irrespective of format so that digital and physical coexist. They can thus enable us to visualise the collection or collections holistically but only to the extent that the necessarily limited metadata will permit. They can support browsing of the collection via lists of the bibliographic artefacts sorted in various ways but they fail to convey the richness of the collections because of the inherent limitations of those artefacts. Although the tools are powerful and strongly support discovery and the evaluation of items, they ignore the sensory aspects of assessing collections, the ways in which we bring our eyes and other senses to assist us in navigating physical collections, assessing the relevance of collection areas to our interests, evaluating the items in those areas and surprising ourselves through serendipity.

Many attempts are being made to overcome the inherent limitations of catalogue based discovery and visualisation. The inclusion of thumbnail images of book covers in catalogues is now almost commonplace. The image of the cover recalls the appearance of the book and, to some extent, evokes its physicality. Icons derived from logos, titles or in other ways can be used for digital items so that they too have a concrete presence during discovery. Faced with the challenge of representing the immaterial data files which make up ebooks as attractive titles for readers, the sup suppliers of ebooks and ebook readers have adopted

imagery that evokes the physicality of libraries. The Kindle Fire for example, displays its holdings via an image of a bookshelf (<http://www.pdfstoepubmac.com/resources/kindle-fire-review.html>). The iPad's virtual bookshelf to display downloaded eBooks has a timber appearance which reminds us of traditional bookcases.

These approaches help us to give substance to the items we discover and thus to comprehend them as elements of collections.

Web 2.0 to the rescue

Other methods of visualising collections have emerged from Web 2.0 and the rapid adoption of social media. Many map books, music and sometimes other materials as collections constructed socially by their users. LibraryThing (<http://www.librarything.com>), for example, claims to be a community of 1,500,000 booklovers which “connects you to people who read what you do” and enables you to “Catalog your books from Amazon, the Library of Congress and 690 other world libraries ... Find people with eerily similar tastes... Find new books to read... “. Similarly, Shelfari (<http://www.shelfari.com>) from Amazon.com is promoted as “a community-powered encyclopedia for book lovers” to “Create a virtual bookshelf, discover new books, connect with friends and learn more about your favorite books – all for free”. At the discovery level it presents the books with a thumbnail of the cover and a brief synopsis and enables them to be sorted in various ways. Users can then create their own virtual bookshelves of read and yet-to-be-read books. Currently in abeyance, zoomii.com from Canada offered a powerful way of selecting books online via Amazon (Thiessen 2012). Users could zoom from the individual book to collections of thousands to visualise collections at various levels of aggregation. Other sites offering similar functionality include Shelfluv (<http://www.shelfluv.com>) and Picclick

(<http://picclick.com/Books>). Goodreads (<http://www.goodreads.com>) offers recommendations based on members' associations between books. Although yet to be extended to whole library collections, some of these capabilities are being built into contemporary library catalogue discovery layers.

More visual approaches to discovery include Picclick which offers its results as a basic display of thumbnail images but, disappointingly, it can be searched only by keyword, not visually. Liveplasma (<http://www.liveplasma.com>) has a basic search interface but, interestingly, shows books, movies or music as items within a web of associations so James Joyce's *Ulysses* is linked to *Dubliners*, *Finnegan's Wake* and a biography of Joyce but also Blake's *Songs of Innocence and Experience*. This connections oriented technology prototypes a powerful method for understanding collections as multidimensional webs of associations. Thus, for items held by the State Library of New South Wales, we can see that images of the breadfruit plant are linked to maps of the island of Timor via Bligh's logbook which records his amazing feat of navigation after being abandoned at sea in a longboat with a few loyal shipmates. The logbook in turn links to newspaper reports of the recent floods in Brisbane through his descendant, Anna Bligh, who was the Premier of Queensland at that time and so on and on. This connections based approach offers the potential to unite the corpus of collections through the multilayered associations between the elements of the collections.

Visualisation beyond the catalogue

More adventurous and artistic approaches to visualising collections include Chris Gaul's *Book babble* (<http://www.chrisgaul.net/utslibrary/prototypes/book-babble>) which gives voice to the words in books on shelves by using RFID and Google Booksearch coupled with

text to speech technology. It is a very powerful work which strongly conveys that libraries are houses of stories which individuals hear in their own ways.

Alicia Martin's installation at Madrid's Casa de America startles the visitor with an avalanche of books pouring from a window of the institution (<http://cubeme.com/blog/2012/03/19/alicia-martins-biografias-installation>). One of three in her Biografias series, it presents books as a torrent of ideas and lives. Bronia Sawyer's book sculptures are more delicate and intricate (<http://www.thisiscolossal.com/2011/06/book-sculptures-by-bronia-sawyer>). *Recycled Library: Altered books* is a travelling exhibition tracing a history of altered books in Australia through the work of seventeen artists including books presented as sculptures, wall-based collages, artists' books, and photographs (<http://www.sl.nsw.gov.au/events/index.html>). Other re-presentations of books include their refashioning as notebooks, furniture and building features.

These and other artistic interpretations of library collections give us pause. How might we more imaginatively effectively enable our clients to visualise our collections and the information and ideas they convey?

Conclusion

The varied approaches to visualising collections and their elements – books, journals, eresources, ephemera, realia, etc – extend from the traditional images of serried shelves to the surprises of conceptual art. They employ the metadata we apply to bibliographic records and the powerful capabilities of social media. By applying them imaginatively we can unite

the corpus of our collections, digital and physical, permanent and ephemeral and however and wherever stored.

At the State Library of New South Wales, we face enormous challenges to both expose our collections and to assist our clients to perceive the connections within them. We have nearly completed a very large project to generate records for all of our holdings so as to enable discovery. We are vastly expanding our drive to digitise our valuable heritage materials. We are working to better house the physical collections in dense storage and, hopefully, an automated system. We have significantly expanded our digital resources which we deliver statewide and we are working to increase our capture of born digital materials. We have started to explore the methods – including artists' ideas – that will enable us to unite this large and culturally valuable corpus.

References

Bauer, Louise (2009), 'Public Libraries – Third Places, Social Spaces', Submission, 2009 ALIA Public Libraries Summit, Australian Library and Information Association, <http://www.alia.org.au/governance/committees/public.libraries/summit09/pdfs/Public%2520Libraries%2520Third%2520Places,%2520Social%2520Spaces.pdf>, accessed 20 April 2012.

Haskell, Daniel (2010), 'The living room of the community', Haskell Architects, http://www.libraries.vic.gov.au/downloads/The_Public__Library_Design_Working_from_the_Inside_Out_and_the_Outside_In_Seminar/danielhaskel_thelivingroomofthecommunitypt129oct2010.pdf, accessed 20 April 2012.

Holmesglen College (2009), 'About us', http://www.holmesglen.edu.au/services/learning_commons_library/about_us, accessed 20 April 2012.

Smith, Adam Jasper (2011), 'The optimization of Fisher Library', *Ampersand Magazine*, Issue 4, Spring, <http://ampersandmagazine.com.au/feature-articles/1438-2>, accessed 2 April 2012.

Thiessen, Sonia (2012), 'What happened to zoomii.com?' <http://www.quora.com/What-happened-to-zoomii-com>; Latechna (2012), 'One of the best ways to view Amazon books' <http://latechna.com/blog/99/zoomiicom-one-of-the-best-way-to-view-amazon-books>; YouTube (2011), 'Tour of zoomii books' <http://www.youtube.com/watch?v=bHuMbNzFKmw>; all accessed 12 April 2012.

Biographical Note

Dr Alex Byrne is the State Librarian and Chief Executive of the State Library of New South Wales following posts in library and university management at several Australian universities. Alex served for a decade in leadership positions with the International Federation of Library Associations and Institutions including President from 2005 to 2007.

Theme: Examining old vs new – blending the traditional collections with the increasingly digital