

Preprint of “Library technology in the next twenty years” *Library Hi Tech* 35, no 1 (2017): 5-10 published by Emerald Publishing.

<http://www.emeraldinsight.com/doi/pdfplus/10.1108/LHT-11-2016-0131>

Published version may differ slightly.

Library technology in the next twenty years

Michael Keeble Buckland

School of Information, University of California, Berkeley, CA 94720-4600

Abstract

Purpose - The purpose of this paper is to examine the agenda for library technology for the next twenty years.

Design/methodology/approach – A long-term historically based analysis of the evolving roles of librarians and library technology, especially the catalog.

Findings - The rise of standardized cataloging codes, communications formats, bibliographical utilities, and software for online searching constitutes a great triumph in universal bibliographical access for everybody. But each reader is unique and no-one is “everybody” so a uniform service is not ideal for all. The ideal librarian knows both the collection and the readers. The catalog is a guide to the collection and a surrogate for the librarian. The development of library technology will remain significantly incomplete until the uniqueness of each reader is accommodated. Some ways to do that are noted.

Research limitations/implications – Research and development should focus on relating the uniqueness of individuals to the uniformity of services provided.

Practical implications – Strategic directions are indicated.

Originality/value – Provides a perspective on the development of library service in terms of changing relationships between technology and librarians.

Keywords – Digital documents, Libraries, Librarians, Catalogues, Library technology, Networked resources.

Paper type – Conceptual paper.

Introduction

Our point of departure is the ideal of a friendly, knowledgeable librarian who knows the collection and is familiar with its readers. We can think of the scholar librarian of general libraries in the seventeenth and eighteenth centuries when collections were so much smaller than now and the knowledgeable librarian mediated access. “The gap between the primary museal function of the library and the occasional need of library visitors to actually locate an individual book was bridged by none other than the librarian, which is to say: by his local memory (*memoria localis*) and his polyhistoric knowledge” (Garrett 1999, p. 111).

There are, however, problems with a human librarian as a service provider. There are limits to how many topics can be mastered, how large a collection can be known well, and how many readers can be helped in a day. In engineering terms, a human librarian does not scale well and, as in other fields, only self-service scales affordably. Worse, a human librarian is prone to

catastrophic failure: the librarian may leave or die or become forgetful. Library “hi tech” and the origin of library science, by that name, can be seen as a response to these problems.

Around the end of the seventeenth century many monasteries in Europe were closed and their libraries confiscated. In Bavaria 200 monastic libraries were sent to Munich to be added to the royal library. The librarians were unable to cope with this flood of material until librarian Martin Schrettinger (1772-1851) understood that technical systems were needed to enable readers (as well as librarians) to find what they needed by themselves quickly and easily. It was for the technical guidelines that he developed that he coined the phrase *Bibliothek-Wissenschaft* (library science) in his textbook Schrettinger (1808, p. 11) which began “A ‘library’ is a large collection of books *whose organization enables every knowledge seeker* to use every treatise it contains without unnecessary delay according to his needs” (Transl. Garrett (1999, p. 116), emphasis added.).

The central idea was that with a unique identifier for each volume, a good catalog, and a link from the catalog record to the volume’s shelf location anyone could make satisfactory use of the collection whatever the arrangement of books on the shelves without requiring the help of a librarian. His particular solution was to provide shelf arrangement in subject clusters complemented by an author catalog, a shelf-list, and, later, a subject catalog (Garrett 1999; International dictionary 2001, v.1: 204-205).

The Library Catalog

A library collection of any significant size needs some kind of inventory in order to know where each document is and to avoid unintended duplication in purchasing. But here we are interested in the public catalog provided for readers. The traditionally accepted statement of the purpose of the catalog is the three objectives given by Charles A. Cutter (1904, 11-12):

- (1) To enable a person to find a book of which the author, the title, or the subject is known;
- (2) To show what the library has by a given author, on a given subject, or in a given kind of literature; and
- (3) To assist in the choice of a book as to its edition (bibliographically) or as to its character (literary or topical).

The individual catalog record provides for the third objective, although to a rather limited extent. Arrangement of records by author, title and subject support the first two objectives.

The catalog, then, represents the books in the collection, individually and collectively. Two kinds of representation can be distinguished: indexical, when the representation points to the document; and surrogate, where the representation can serve as a substitute for the document, at least for some purposes, e.g. verifying the name of the author if the title is known or vice versa. Similarly, a printed bibliography is properly seen as an interface to the body of literature covered.

The catalog not only represents the collection, it also represents the librarian in two ways. First, a catalog record represents a brief summary of what the librarian knew and might remember about the document: author, title, topic, date, and call number. Second, by extension, the catalog as a whole is a representation of what an ideal librarian might know about the

collection as a whole but in practice would be unable to fully memorize. In this way the catalog is a partial surrogate for the librarian, with some advantages:

- (1) Accuracy: No human memory could retain all of these details;
- (2) Resilience: Librarians retire and die much more often than catalogs do; and
- (3) Service capacity: A catalog can serve many more readers simultaneously than a librarian can and it can remain constantly in use.

Of course, catalogs, despite their richness of detail do not record everything that a librarian may know about a document: That, for example, book A has been superseded by book B, that its publisher is a vanity press, or that it is currently out on loan.

Uniformity

During the nineteenth and twentieth centuries much progress was made in the development and standardization of cataloging codes and the sharing of catalog records. In Europe a few large libraries retained local, non-standard cataloging rules unto the second half of the twentieth century, but even there the record content was increasingly adapted from the *British National Bibliography* or other trustworthy sources. Bibliographies and the published catalogs of major libraries have long been available as a resource for catalogers and it is a very long time since librarians made catalogs entirely from their own original cataloging, if they ever did.

Gradually an impressive degree of universal bibliographic access has been provided, going a long way towards fulfilling the grand visions of Paul Otlet and others. It is a triumph of modernity, of standards, systems, collaboration, and efficiency to provide uniform catalog access to everybody. It is an achievement to be proud of.

This uniformity of universal access is largely centralized. Bibliographic access is distributed in manner reminiscent of television broadcasting: a standard version is distributed for all. And as MARC records dissolve into sets of linked open data links, the uniformity is accentuated with ever-greater efficiency achieved for everybody.

But there is a flaw: No reader is everybody.

No Reader is Everybody

No individual can know everyone else in the world, every place, every institution, every building, and every event. We cannot attend to every media outlet or publication. Each of us knows a lot less than is in principle knowable. Instead, we have a limited circle of friends, family, neighbors, and colleagues. We know, more or less, the neighborhood we live in, the routes we travel, and a work or school environment. Our personal world is a limited world even though it is complex and includes participation in multiple, different, overlapping communities. We all live in what Elfreda Chatman called small worlds (e.g., Chatman 1992). Each social group evolves its own microculture and use of language and each individual has an individual sphere of expertise. So all readers are geographically, socially, and intellectually limited and each individual reader is also a particular, a unique person.

So in theory we need multiple catalogs, one for each small world, or, better, one for each individual reader, or else a dynamic one that could adapt as needed to the personal context of each reader. With existing technology that was impractical, so it has been the librarian's role to

help diverse readers to cope with bibliographical uniformity. Later manifestations of the ideal librarian, such as the public library’s readers’ advisor, the corporate special librarian, and children’s librarians in school and public libraries, differed from the eighteenth-century scholar librarian in mediating with the library’s catalog as well as the collection.

In the late nineteenth century, some librarians tried to enrich their catalog for local needs by adding analytical entries for items within documents. An entry might be made in a college library catalog for a foreword written by that college’s president or in a public library’s catalog for an appendix relating to a landmark of only local interest. But the cost was too high.

Demand and privileging

Many librarians have believed that collection development was the most important professional task. Lionel McColvin's (1925) classic *The Theory of Book Selection for Public Libraries* begins “Book selection is the first task of librarianship. It precedes all other processes – cataloguing, classification, or administration – and it is the most important. No matter how thorough and efficient the rest of the work may be, the ultimate value of a library depends upon the way in which the stock has been selected” (McColvin, p. 9). It has long been a powerful part of libraries’ mystique that other documents can also be obtained, but interlibrary loan has played a very small role quantitatively.

But now the steady acceptance of network-accessible digital resources has been slowly replacing traditional library infrastructures. In a paper environment local collections effectively determined the service provided. Local collections served two purposes in an integrated way: catering to readers’ demands, and (2) the privileging of resources that the librarian selected as being most beneficial for the readers given that particular library’s mission. But with electronic resources these two purposes separate. Assuring that copies are physically available when needed is essentially an engineering issue beyond the responsibility or competence of collection development librarians. Their role of selecting which sources to privilege, however, remains just as important, but it is no longer done on a document by document basis. It is increasingly arranged wholesale, not retail, by licensing access to very large corpora and supporting the development of open access repositories such as the Internet Archive’s Open Library or consortial, membership programs such as JSTOR. So the local collection and maintaining the local catalog play progressively smaller roles relative to search, discovery, and selection. (For a more detailed discussion of demand and privileging see Buckland (1995).)

Subject Indexes

Since readers come from different backgrounds they do not have the same interests or vocabulary, so in principle there should be different subject indexing for each group of users even for the same resources.

Language evolves within communities of discourse. Every such community has its own more or less specialized, stylized practice of language. Attempts at controlled or stabilized vocabulary must deal with multiple and dynamic discourses and the resulting multiplicity and instability of meanings. Most bibliographies and catalogs have a single topical index, but include material of interest to more than one community. Since each community has slightly different linguistic practices, no one index will be ideal for everyone and, perhaps, not for anyone. In ordinary discussion of health, for example, the terms *cancer* and *stroke* are commonly used, but in a professional medical writing *neoplasm* and *cerebrovascular accident* are preferred names. It

is not, however, only a matter of linguistic variation, but also of perspective. Different discourses discuss different issues or, when the same issue, from different perspectives.

Imagine three doctors – an anesthesiologist, a drug therapy specialist, and a geriatrician – who each wanted recent literature on cardiac arrest (a medical term for a heart attack). “Cardiac arrest” itself is not a heading used in the standard Medical Subject Headings (MeSH) vocabulary, so what would be the most effective MeSH headings to use? The three doctors are specialists. They do not have the same kind of interest in cardiac arrests. Each inhabits a different medical subculture. Each would not be interested in (and might not understand) the specialized literature of interest to the others even when relating to the same topic. The ideal librarian would adapt each reader’s query accordingly. How can technology help with that?

Uniformity and Diversity: Summary

- (1) Access to local collections has been delegated to remotely derived uniform cataloging.
- (2) Local collections are increasingly being displaced by reliance on access to remote repositories of institutions and publishers which provide quite varied search support with indexing languages likely to be unfamiliar to all but experienced specialized readers.
- (3) There is a trend towards big data for which access is not, in principle, different from other areas of bibliography, but for which bibliographic description is less well-developed (Buckland 2011) and data sets have been prone to inadequate contextual metadata (Bowker 2005). Similarly, bibliographic coverage of micropublications (e.g. blogs and working notes) is incomplete.
- (4) So the task for the librarian who once providing knowledgeable access to a known, local collection now has the task of mediating between very diverse individuals and uniform, remotely broadcasted (and mostly unfamiliar) metadata languages.
- (5) The same imperatives that led to reliance on public catalogs – reliability and efficiency – now point to a need for investment in mediating mechanisms between readers and remote bibliographical services.

Conclusions

With hindsight we can see a kind of dialectic between library and technology. Publications increased through social pressures and improved technology. As local collections grew, they exceeded the capacity of the librarian and so catalogs increasingly provided the needed mediation between reader and collection. Further technical developments are now eclipsing the role (and limitations) of the local collection, rendering its catalog progressively less relevant in the overall challenge of bibliographical discovery. What reader knows or cares whose electronic memory provides a desired digital document?

That we have the ability to use these remote collections is a powerful tribute to the success in establishing universal and uniform bibliographical access for everybody through standards and collaborative cataloging. But there remains the perennial challenge of selecting the most suitable documentary means for our individual purposes. McColvin’s insistence on the

importance of selection is still just as valid, but it is no longer limited to the local collection because the local collection no longer determines what is locally available. With the rise of networked access the need and the challenge are much greater.

Readers are not uniform. The ideal librarian knew both the collection and the readers. The catalog has acted as a powerful surrogate for the librarian in mediating between the reader and the local collection, but there is no evident surrogate for the librarian’s knowledge of the reader and ability to mediate between reader and catalog, let alone between the reader and the increasingly more relevant other kinds of bibliographies.

Some strategies for improvement can be identified:

- (1) Diagnoses of why searching fails and determination of corrective measures (e.g. Berger 1994).
- (2) Richer indexing, especially more links between headings not only for semantic relations but also for functional relationships (e.g. horse and cart) and, of course, from non-preferred terms to preferred terms.
- (3) Mapping between related terms in different vocabularies and especially those connecting different facets. For example, if the library subject heading for, say, castles, were linked to the geographical feature type for castles in a place name gazetteer, then literature about castles can be linked to actual instances of castles and vice versa. This is no more than scholars have done tediously by hand in reference libraries (Buckland 2017, pp. 127-133).
- (4) Search term recommender services which are sensitive to different contexts can be created drawing on training sets derived from bibliographical records carefully tailored to represent the distinctive discourse of specialists. This provides, in effect, multiple separate catalogs for readers in diverse specialties (Petras 2006). Perhaps a similar approach based on the individual scholar’s research notes might also be feasible.
- (5) Resuming the late nineteenth century interest in providing analytical entries for parts within documents.
- (6) As technology and scholarship evolve, other kinds of documents with seriously inadequate bibliographical access, notably datasets and videos, are becoming progressively more important.

There is ample agenda for library technology in the next twenty years.

References

- Berger, Michael G. (1994). *Information-seeking in the online bibliographic system: An exploratory study*. Ph.D. dissertation. University of California, Berkeley. (UMI AAI9504745).
- Bowker, Geoffrey. (2005). *Memory practices in the sciences*. Cambridge, MA: MIT Press.
- Buckland, M. K. (1995). What will collection developers do? *Information Technology and Libraries* 14, no 3: 155-159. <http://escholarship.org/uc/item/2v2258mk>
- Buckland, M. K. (2011). Data management as bibliography. *Bulletin of the American Society for*

M. K. Buckland, *Library technology in the next twenty years*” *Library Hi Tech* 35, no 1 (2017): 5-10 7

Information Science and technology 37 no. 6 (Aug-Sep 2011): 34-37.

http://asis.org/Bulletin/Aug-11/AugSep11_Buckland.pdf

Buckland, M. K. (2017, forthcoming) *Information and Society*. Cambridge, MA: MIT Press.

Chatman, Elfreda A. (1992). *The Information World of Retired Women*. New York: Greenwood Press.

Cutter, Charles A. (1904). *Rules for a Dictionary Catalog*. 4th ed. Washington, DC: Government Printing Office.

Garrett, Jeffrey. (1999). Redefining order in the German library, 1775-1825. *Eighteenth-Century Studies* 33, No. 1, (Fall, 1999): 103-123 . <http://www.jstor.org/stable/30053317>

International Dictionary of Library Histories, edited by David H. Stam. Chicago : Fitzroy Dearborn, 2001.

McColvin, Lionel. (1925). *The Theory of Book Selection in Public Libraries*. London: Grafton.

Petras, Vivien. (2006). *Translating dialects in search: Mapping between specialized languages of discourse and documentary languages*. Ph.D. dissertation. University of California, Berkeley. <http://www.sims.berkeley.edu/~vivienp/diss/vpetras-dissertation2006-official.pdf>

Schrettinger, Martin. 1808. *Versuch eines vollständigen Lehrbuches der Bibliothek-Wissenschaft*. Munich: Author. Available at http://archive.org/details/bub_gb_x2qePg9yKNkC

Stam, D. H. (Ed.) (2001), *International Dictionary of Library Histories*. Chicago: Fitzroy Dearborn.