New resource discovery mechanisms

JENNY WALKER

Executive Vice President Marketing, Xrefer

The E-Resources Management Handbook – UKSG

E-access opens up great opportunities for people to find – and get access to – an increasing corpus of knowledge. Search engines such as Google, MSN and Yahoo! are now targeting the traditional library user; and libraries are under increasing pressure to develop and offer new paradigms for discovery that meet the changing expectations of end-users. The library needs to be where the users are, and to deliver quality services to the users where and when they need them – on the network and within their preferred work or study environment. This chapter looks at some of the tools and technologies currently being deployed in libraries and briefly reviews some of the new paradigms for resource discovery.

Introduction

Nothing but Google?

Previous chapters of UKSG's *The E-Resources Management Handbook* have presented a good case for e-publishing and a clear indication that the number of e-resources grows apace. For many people, e-access opens up great opportunities with an increasingly significant corpus of information available at their fingertips through search engines such as Google[1], MSN[2] and Yahoo![3] and through sites such as Amazon[4] and Wikipedia[5]. Although most students today would cite Google as their premier source of research[6], to date, Google – and other such search engines – is unlikely to fully satisfy the researcher who seeks credible and unbiased scholarly content, which is not always available free of charge. Google is attempting to reach this market through recent initiatives such as Google Scholar[7] and Google Book Search[8], both in beta version at the time of writing. Even more recent initiatives include Yahoo! Mindset[9] (also in beta version at this time) and the offering of the Open Content Alliance[10], with members such as MSN, the Research Libraries Group (RLG)[11] and the University of California.

With the emergence of these new tools targeting the traditional library user, librarians are under increasing pressure to develop and offer new paradigms for discovery that meet the changing expectations of end-users and to deliver quality services to the users where and when they need them. The library needs to be where the users are – on the network and within their preferred work or study environment.

Libraries are starting to respond to these pressures. One example is the launch in January 2006 of a dramatically different online public access catalog (OPAC) at North Carolina State University (NCSU)[12]. This OPAC is built on a search engine from Endeca Technologies, Inc.[13] With the speed and flexibility of popular online search engines, the Endeca technology also introduces innovative features such as faceted browsing – or guided navigation – which helps a user refine a search. (For a discussion of faceted browsing, see the section 'New paradigms for resource discovery' in this chapter.) While the NCSU OPAC has received great plaudits for its innovative approach, it is, however, limited in terms of both the data types included and the user services offered.

Publishers are also under pressure in this changing world to ensure that in the plethora of information that is available, their resources are highly discoverable (that is, can be found online easily) and highly credible (recognized as coming from a trustworthy source). The changing landscape creates new opportunities for marketing, allowing publishers to reach deeper into established markets and to use web search engines such as Google and MSN to reach new audiences. It is important to scholarly publishers

that users recognize their resources as coming from a trustworthy source and that their resources be well integrated into a user's environment to allow seamless navigation through the user's information landscape. The many aspects of improving the 'discoverability' of resources include, for example, linking and metasearch; and adherence to standards to facilitate this discoverability is critical to success.

Resource discovery today

Libraries have developed standard internal practices for the management of metadata relating to physical collections, predominantly books and journals, and also sophisticated tools to assist in the management of these items. Furthermore, libraries have predictable ways of presenting services to users. In the management of physical collections, standards such as MARC, the Anglo-American Cataloguing Rules (the second edition, AACR2), and the International Standard Bibliographic Description (ISBD) are widely applied, with few exceptions or deviations. This is not the case in the management of digital collections, where:

- many different – and emerging – metadata formats are used to describe the resources and are not always applied consistently
- complex licensing requirements make for complex presentation of the resources and access to the resources
- each resource typically has its own interface and its own authentication method.

In short, as Lorcan Dempsey, vice president of research and chief strategist at OCLC, puts it, "The digital environment is one that lacks consistency; it is as if each book coming into the library was a different shape and had to be read in a different way."[14]

The management of physical items such as books and journals remains extremely important. However, such management must now take into account the multitude of new and growing formats.

Often unaware that various discovery tools are available for the resources that the library has to offer, library users today become all too frequently frustrated that they cannot find what they seek. Those who are aware of the disparate tools – and 'silos' of information – often find the distinctions between the various resources to be arbitrary. Among the many resources available from a typical academic library are:

- *physical resources*, such as books, journals, and videos. These are commonly cataloged by librarians, and their descriptions (metadata) are available online through an OPAC
- *various abstracting and indexing databases,* for example, MEDLINE®, ERIC, Academic Search™ Elite, and ABI/INFORM®. These are available through a range of different interfaces offered by the information providers
- *various full-text databases,* such as JSTOR and Wiley InterScience
- *specialist databases,* such as geospatial databases, chemical structure databases, and databases for mathematical formulae
- *institutional repositories,* such as a database of electronic theses and dissertations (ETDs)
- *digital collections,* such as cultural heritage collections
- *subject gateways* as provided by the Resource Discovery Network (RDN)[15].

This list represents a range of items that constitute the library's collection. The collection can include items that the library manages, items that third parties manage, items held locally, items held remotely, items for a fee, and items that are free at their point of use. Libraries are challenged to find the best ways to present and promote their growing collections of heterogeneous resources and to do this in a way that:

- enriches learning and research
- provides timely, convenient access to relevant and appropriate resources
- exposes potentially valuable resources that otherwise might be overlooked

- enables users and the library to focus on a fruitful use of collections rather than on dealing with sometimes difficult aspects of access, navigation, and manipulation of the result sets[16].

Resources that the library makes available must be integrated with one another and within the library environment; and library services must support the learning and research behaviors of users. Furthermore, users often want to access and use items from more than one content provider, so they have no choice but to interact with various user interfaces. But even then, each service has a different user interface for discovery, with its unique set of 'presentation services' that the user must learn and understand. Although the results are readable (in HTML), they are difficult for users to compare, re-use, and manipulate; and users may need to deal with repeated authentication challenges.

A number of tools have emerged to help libraries meet the challenges in serving their users. These tools include:

- tools to assist in the creation and maintenance of an A–Z list of databases. (Such a list facilitates users' access to a range of databases)
- tools to assist in the creation and maintenance of an A–Z list of electronic journals. In some cases, such lists also show print journals
- metasearch tools to enable a user to search heterogeneous resources simultaneously through a single query form and to receive results in a consistent way that permits both the merging of the results and their general re-use, for example, onward linking from a retrieved citation to the full text of the cited item
- link servers to enable users to seamlessly move among disparate library resources, particularly from a citation to the full text
- library portals, which include many of the tools in this list. Among the key functions of a library portal is personalization, in the form of personalized alerts, personal resource lists, and saved search histories, for example.

A number of metadata standards have emerged – or are working their way through the standards processes – to support such tools: standards for linking (Digital Object Identifier, or DOI[17], and OpenURL[18]), collection description (Dublin Core Collection Description[19]), service access description (ZeeRex[20]), and the exchange of serials holding data (ONline Information eXchange, or ONIX[21]). These standards supplement the growing number of descriptive metadata schemas, such as MARC, Metadata Encoding and Transmission Standard (METS)[22], Metadata Object Description Schema (MODS)[23] and Dublin Core, each developed to meet a specific need. It is encouraging, however, to see that the committee tasked with reviewing the Anglo-American Cataloguing Rules (AACR) has proposed a revolutionary – and much needed – alternative to AACR2. The proposed new standard, first released as a draft in December 2005, is to be called Resource Description and Access (RDA)[24]. Designed from the ground up for the digital world, the standard will nevertheless provide a comprehensive set of guidelines and instructions on resource description and access covering *all* types of content and media.

Metasearch

Metasearch tools, sometimes referred to as federated search tools, emerged in the early 2000s and, since then, their development and deployment have been growing steadily.[25] For a list of metasearch tools for libraries, see the Library of Congress Portals Applications Group web site.[26]

A metasearch tool provides the interface between the user and the resources being searched, translating the queries and manipulating and displaying the results. The metasearch tool must 'understand' and work with a variety of search and retrieval methods and manage the following diverse access mechanisms:

- *Z39.50*[27]: For many years, library OPACs have supported the Z39.50 search and retrieve protocol, which is well proven in this environment. However, despite the complexity of this protocol and the fact that it was not designed for article-level citations, a number of major information providers, particularly the abstracting and indexing vendors, followed suit in implementing it.

- *Search/Retrieve via URL (SRU)*[28] *and Search/Retrieve Web Service (SRW)*[29]: SRU is a standard search protocol for Internet search queries, utilizing Common Query Language (CQL), a standard query syntax for representing queries. Queries are sent as URLs via HTTP to retrieve XML. SRW is a variation of SRU in which messages are conveyed from client to server by XML instead of a URL. Queries are sent over HTTP by means of a web services layer, which wraps the XML message in an XML envelope.
- *NISO Metasearch XML Gateway (MXG)*[30]: Currently available as a NISO draft standard for trial use, this protocol is a 'lightweight' alternative to the Z39.50 and SRW/U protocols and is based on SRU. MXG is targeted at content providers, especially newer players serving the library market who are not familiar with the library-based standards.
- *Proprietary XML gateways*: Some vendors offer proprietary XML gateways for searches using metasearch tools. Such gateways vary from vendor to vendor, each requiring custom programming.
- *'HTML parsing', also known as 'screen scraping'.* For those vendors who offer no standards-based or structured search and retrieval interface, it is possible – though not desirable – for the metasearch tool to use HTML parsing methods. These methods can and do work but are fragile and highly susceptible to interface changes by the vendor. Hence, maintaining programs or connectors for HTML parsing is labor intensive. HTML parsing is inefficient for the metasearch system provider, the content provider and, ultimately, the end-user.

Much has been written on the challenges of metasearch (see, for example, Tamar Sadeh[31], Judy Luther[32] and Andrew Pace[33]). An excellent bibliography can be found on the Library of Congress Portals Applications Group web site. To address some of the key challenges and to move toward industry solutions, NISO sponsored a Metasearch Initiative[34] in early 2004. The purpose of the initiative was to enable:

- metasearch service providers to offer more effective and responsive services
- content providers to deliver enhanced content and protect their intellectual property
- libraries to deliver services that are distinguished from those offered by Google and other free web services.

The NISO Metasearch Initiative

In September 2005, the NISO Metasearch Initiative published a set of draft standards and best practices in three key areas relating to metasearch: access management, collection description, and search and retrieval.

Access management

Although discussions of resource discovery do not immediately bring up access management – which covers both the authentication and the authorization processes for metasearch – it can be a key impediment to success. Relying on licensed information, on one hand, and committed to user privacy on the other, access management has a broader reach than metasearch in terms of academic institutions and other research organizations. Most institutions are already actively involved in deploying, or at least investigating, single-sign-on environments that will enable users to navigate between applications without being challenged to present their user credentials and yet will provide the appropriate levels of security for the systems and the data being accessed.

In a metasearch environment, authentication and authorization processes come into play between the user and the metasearch tool and also between the metasearch tool and the resources that are being searched. In each such case, a multistep process requires the presentation and checking of credentials and then the retrieval of attributes for use in the authorization process.

The NISO Metasearch Initiative recommends that institutions that are in the process of acquiring new electronic resources implement either IP authentication with a proxy server or a username/password authentication system to control access to their electronic resources. Such systems are the most widely supported by vendors, have the lowest implementation and maintenance costs, and are the simplest for smaller or less technically sophisticated organizations to install. They also ensure that remote (off-site) users can access the resources of the institution with little difficulty.

An initiative from the Internet2 community, Shibboleth[35], holds much promise in providing access management solutions that allow exchange of user information in a secure and privacy-preserving manner; but this method is not yet widely deployed.

Regardless of the methods that prevail, it behooves the community to heed the comments of Clifford Lynch: "Access management needs to be routine and easy to implement; once a contract is signed, lengthy technical negotiations between institution and content supplier should not be necessary before users can have access. In a world of networked information resources, access management needs to be a basic part of the infrastructure, and must not become a barrier to institutional decisions to change or add resource providers."[36]

Collection description

It is critical to effective metasearching that end-users be able to find relevant available resources. Also, librarians need help to find resources, describe them to the user, and arrange resources to facilitate end-user access by means of the metasearch tool.

Collection descriptions, be they via the metasearch tool, from an A–Z list, or from a library portal, aid users in selecting the most relevant resources among those accessible. Therefore, data elements are required that describe a collection for the purpose of resource discovery. Collection descriptions allow the owner of a collection to disclose information about the collection's existence and its availability to interested parties so that they can find and use the contents.

Metadata schemas for collection description are required to make the following actions possible:

- the discovery of collections that meet a specified set of criteria
- the identification of a desired collection
- the selection of one or more collections from a number of discovered collections
- the discovery of the services that provide access to the collection – for example, search and retrieval or data harvesting of that collection.

Much work has been done in this area to date, particularly by the Research Support Libraries Programme in the UK (RSLP)[37] and the Dublin Core Collection Description Working Group.[38] The NISO Metasearch Initiative has published a draft standard for collection description that builds on the Dublin Core Collection Description Application Profile (DC CD AP).

Associated with a collection is a set of services – one or more functions of interest to an end-user or a software application. Applications use service descriptions to determine how to access remote services. Examples include structured network services such as those based on a particular protocol (for example, Z39.50, SRU, SRW, Open Archives Initiative Protocol for Metadata Harvesting [OAI-PMH], and FTP). In the metasearch context, applicable search services could be offered through protocols such as Z39.50, SRU, SRW, and NISO MXG.

The NISO Metasearch Initiative has published a draft standard for service access description using ZeeRex. The origins of ZeeRex in the 'Z39.50 Explain' facility are clear, in that ZeeRex stands for 'Z39.50 Explain, Explained, and Re-Engineered in XML.' ZeeRex makes possible the exchange of records that describe databases and how they can be accessed.

With the potential of standardization for collection description and service access description, the problem arises of how metasearch tools, librarians and end-users can discover whether such descriptions are already available and, if so, how to obtain them. Conversely, if such descriptions have previously been created, how can they best be shared with the community? Registries such as the UK Joint Information Systems Committee's (JISC) Information Environment Service Registry (IESR) project[39], and the OCKHAM project[40], funded by the National Science Foundation (NSF), are beginning to address this need.

Search and retrieve

The search and retrieve methodology sits right at the core of metasearch. Metasearch tools today use a variety of search and retrieve methods, including the Z39.50 and SRW/U standards as implemented by information providers. However, many of the interface providers with whom metasearch tools should interact do not yet support such protocols, and metasearch tools must rely on proprietary XML gateways

or HTML parsing scraping methods. Tools do exist to facilitate the implementation of Z39.50 and SRW/U, yet interface providers still consider these protocols too 'heavyweight.'

Many of the interface providers, especially the newer players serving the library market, are not familiar with the library-based standards. A simpler solution lies in the proposed NISO Metasearch XML Gateway (MXG) protocol, which is based on the NISO SRU protocol. The first to implement NISO MXG was the Berkeley Electronic Press ('bepress')[41] and it is hoped that many other vendors will follow suit.

Linking

Web resources are more 'discoverable' when other sites link to them. This principle applies equally to scholarly information resources such as e-journals, e-books and articles; the most common form of linking among these resources is from a citation to the item that the citation references, typically an article or a book. However, many other linking services can help leverage the inherent value of a resource.

Linking solutions are now available – with more and more libraries deploying them – and are supported and assisted by standards such as DOI and OpenURL. Linking solutions built around these standards include CrossRef[42] and link resolvers such as SFX®, Discovery Resolver[43] (formerly called LinkFinder*plus*), and Article Linker™[44].

DOI and CrossRef

CrossRef is a co-operative effort among publishers that uses a digital object identifier, or DOI, for citation linking in journals. A DOI is an alphanumeric name (for example, 'doi:10.1101/gr.10.12.1841') for a digital object such as a book, journal article, chapter or image. Paired with the object's electronic address, or URL, the DOI is stored in a central directory that is easily updated. The DOI is published in place of the URL to ensure that users will be able to find the content even in cases where, for example, the publisher migrates content from one production system to another (preprint to postprint) or content moves from one publisher to another (if a journal or the publisher itself changes ownership.) In such cases, the DOI never changes, so all the hyperlinks to the content that have already been published and disseminated still function. Hence, one key advantage of the DOI system is persistence. The other is 'actionability': like the URL itself, one click on a DOI takes users to the location of the content that they want.

CrossRef is an official DOI registration agency, appointed by the International DOI Foundation[45]. Although CrossRef began with a specific focus on linking journal articles in scientific, technical and medical fields, it now covers DOI-based linking of additional scholarly and professional literature, including a variety of content types and disciplines. Up-to-date statistics on the coverage and use of CrossRef are available on the CrossRef web site[46].

CrossRef provides content visibility through linking; publications are registered in the CrossRef database, thousands of participating organizations (publishers, intermediaries and libraries) programmatically pick up links to that content, and the resource becomes more discoverable and hence more accessible.

OpenURL and link resolvers

The OpenURL is a standardized way to transfer metadata that describes a record from the interface that displays that record to a link server under a library's control. The library, therefore, can determine the 'appropriate' services (or links) to offer to the user. After adopting the OpenURL linking framework, libraries are able to set up links between the resources to which they subscribe – and in a manner that the librarians deem helpful for their users, not necessarily in the manner determined by the vendors. Once implemented, the OpenURL and link servers enable users to navigate seamlessly and in a meaningful way among the resources that librarians have selected for them.

For vendors, this mechanism is an effective way of ensuring that their databases become fully incorporated in the library's interlinked environment. According to Walt Crawford, senior analyst at RLG:

> "If properly implemented, OpenURL is a win-win situation. Good abstracting and indexing services become valuable by linking to local resources. Licensed resources and print holdings see more use

because the link from the identification to holdings is fast and easy. None of this requires fancy new numbers; the information is already there – ISSN, journal and article titles, year, volume, and so on."[47]

An information resource can be both a link resolver *source,* through the implementation of the OpenURL, and a *target,* by enabling users to link to the resource as seamlessly and at as deep a level as possible, that is, directly to the journal article rather than to the journal issue's table of contents page.

New paradigms for resource discovery

Librarians who have invested much in the care and management of their collections tend to expose their work to end-users, assuming that users are sufficiently interested that they will come and search for information. However, as aptly expressed in the popular maxim of Roy Tennant of the California Digital Library, "only librarians like to search; users like to find."[48] Moreover, users want to obtain the information as quickly and easily as possible.

Librarians and the vendors that support them would do well to learn a few principles from popular web search engines such as Google, and community web sites such as de.licio.us[49] and flickr.com [50]:

- Google is the first port of call for most users. Therefore, ensure that library resources are well integrated into Google and, indeed, other search engines. The integration of the OpenURL in Google Scholar[51] is a good example.
- The behaviors and expectations of today's users are changing, so make the effort to learn about them. For example, users take for granted, and expect, relevance ranking; they like the user experience to be intuitive, without compelling them to learn how to use the system; and they expect to be able to interact with the system by contributing, for example, reviews and tags that will help them and others find the information in future.

But Google – and even Google Scholar, which is specifically targeted at the research community – falls short in serving this community. Researchers are mostly associated with an institution and can benefit from the collection development policies and the range of services offered by their institution.

In December 2005, the bibliographic services task force of the University of California consortium of libraries published a report on rethinking bibliographic services for their member institutions[52]. Among the key recommendations in the report for enhancing search and retrieval are the following:

- Provide users with direct access to the item. This recommendation assumes that a system is able – at least for the majority of users – to correctly anticipate what service users want.
- Provide an 'I want this' button when appropriate that would always offer a fulfilment option.
- Provide recommender features that mine the information in the bibliographic and holdings records, aggregated use data and more, to suggest other works of interest.
- Support customization and personalization so that users can, for example, define the set of resources that they want to search simultaneously.
- Offer alternative actions for failed or suspect searches – for example, to compensate for spelling errors or to suggest a preferred search term.
- Offer better navigation of large sets of search results:
  - the grouping of similar results, such as different editions of the same work
  - guided navigation by means of 'facets', which allow users to drill down in a particular area, such as subject or author.
- Provide relevance ranking, which is now commonplace in web search engines. However, relevance ranking in the library environment needs to factor in many criteria such as the type of material, the query, the context and the user.
- Deliver bibliographic services to the users' location, be it their institutional portal, virtual learning environment, favorite web search site or community web site.

Other recommendations in the report focus on the need to support new cataloging practices that are better suited to the new demands for search and retrieval.

Web 2.0 and Library 2.0

The University of California recommendations are consistent with community thinking on the resource discovery process, as evidenced by industry discussions regarding Web 2.0[53] and Library 2.0[54]. According to the Wikipedia entry for Library 2.0, 'Library 2.0 is a model for a library service that reflects a transition within the library world in the way that services are delivered to library users. This redirection will be especially evident in electronic offerings such as OPAC configuration, online library services, and an increased flow of information from the user back to the library. The concept of Library 2.0 borrows from that of Web 2.0, and follows some of the same philosophies underpinning that concept. Proponents of this concept expect that ultimately the Library 2.0 model for service will replace outdated, one-directional service offerings that have characterized libraries for centuries.'[55]

Tagging

A January 2006 article in *The Wall Street Journal* entitled 'The Next Big Thing in Searching' indicates that tagging is now moving into the mainstream. 'Tagging sites are increasingly transitioning beyond places individuals go to for retrieving their favorite web pages to sites they visit first when they want to search the Internet. That means they are beginning to compete directly with search behemoths such as Google and Yahoo!.' [56] Yahoo! recently purchased two of the early tagging web sites, de.licio.us and flickr.com. Others that have emerged in recent years include reddit.com, wink.com and shadows.com. Tagging is now also available on sites such as Amazon.com.

But what exactly is tagging? Tagging addresses a commonly heard complaint from Internet users that searching is clumsy and inefficient, they have to trawl through many pages of search results to find what they are seeking, and often must repeat this exercise to get back to previously retrieved results. Tagging allows web surfers to save their favorite web pages under key words; and unlike a folder of bookmarks on a single computer, tagged pages are stored on the web and accessible from any computer. Users can then base a search on tags that they have previously assigned and can publish these tags for a closed community or for the entire web public. Other users can then limit their searches to an edited set of web pages that others have already tagged as interesting or helpful.

Although the initial use of tagging was for web sites, tagging is now increasingly available for specific information on web sites – for instance, items found on Amazon.com. Users expect this type of functionality, perhaps combined with other social-networking features such as reviews and recommendations, when viewing scholarly articles. Tagging may not be precise, but it certainly reflects the way people view items and enables them to share their views. For example, students might tag materials relevant to a course in a way that helps all their classmates take advantage of each other's research results.

Faceted browsing

Faceted browsing, mentioned in the University of California recommendations, warrants further discussion. Facets allow users to drill down in a particular area – 'facet' – for example, subject. This approach of guiding a user through a set of results ensures that the user progressively refines the search in the most appropriate manner, selecting only from valid values. This method takes the guessing out of the standard refinement function, whereby a user must know or guess the specific term and the correct spelling to use in refining the search. Furthermore, guiding the user through a result set ensures that there is no 'dead end' or null result.

An information provider can, based on its user constituency, determine what facets to apply to a resource and how to map the data in the resource to create these facets. Many options for scholarly content exist. For example, users who live far away or have to submit the paper early the next day might want to see only materials that are offered online. Other users may want to see only articles or books by a particular author, only articles in French, or articles published in the current year. Today's systems need to provide multiple options for narrowing down the search and helping users who do not define their search criteria precisely enough; such users, when they see the results and the system's analysis of the results, can easily follow the system's suggestions. The system can build a list of topics based on the subject information that it found in the retrieved records and display this list to the end-user, who now easily chooses the most relevant subject. Similar facets can be built based on author names, material types, language and dates.

Conclusion

New search engine technologies, such as Endeca, FAST[57], and Lucene[58], can transform the user experience with their speed of retrieval and faceted browsing capabilities. For these technologies, the data must be preharvested, premassaged and preindexed by the search engine – a 'just-in-case' approach. This approach requires upfront preprocessing of the information to be searched but saves precious time for the user who is waiting for the information and enables enrichment of the information for the best user experience. Preprocessing of results is not possible for resources that are not under the library's control. These resources must be provided and processed 'just in time' – that is, following the user's query. Finding a way to combine the best of the just-in-case and just-in-time methodologies to meet changing user expectations is the current challenge for librarians and their vendors.

References

1. Google
   http://www.google.com
2. MSN
   http://www.msn.com
3. Yahoo!
   http://www.Yahoo!.com
4. Amazon
   http://www.amazon.com
5. Wikipedia
   http://www.wikipedia.org
6. OCLC, *Perceptions of Libraries and Information Resources*. A report to the OCLC membership, 2005, Dublin, Ohio.
   http://www.oclc.org/reports/2005perceptions.htm
7. Google Scholar
   http://scholar.google.com
8. Google Book Search
   http://books.google.com/
9. Yahoo! Mindset
   http://mindset.research.Yahoo!.com/
10. Open Content Alliance
    http://www.opencontentalliance.org/
11. RLG (Research Libraries Group)
    http://www.rlg.org/
12. North Carolina State University (NCSU) OPAC
    http://www.lib.ncsu.edu/catalog/

13. Endeca Technologies, Inc.
http://endeca.com

14. Dempsey, L., The recombinant library: portals and people. Co-published simultaneously in *Journal of Library Administration*, 2003, 39(4) 103–136; and in *Improved Access to Information: Portals, Content Selection, and Digital Information*, Ed. Lee, S.H., Binghamton, NY, The Haworth Press. Available at
http://www.oclc.org/research/staff/dempsey/recombinant_library/

15. RDN (Resource Discovery Network)
http://www.rdn.ac.uk/

16. Dempsey, L., *ibid*.

17. The Digital Object Identifier (DOI) System
http://www.doi.org

18. The OpenURL Framework for Context-Sensitive Services
http://www.niso.org/committees/committee_ax.html

19. Dublin Core Collection Description Application Profile
http://www.ukoln.ac.uk/metadata/dcmi/collection-application-profile

20. ZeeRex: The Explainable 'Explain' Service
http://explain.z3950.org/

21. ONIX for Serials
http://www.editeur.org/onixserials.html

22. METS: Metadata Encoding and Transmission Standard
http://www.loc.gov/standards/mets/

23. MODS: Metadata Object Description Schema
http://www.loc.gov/standards/mods/

24. Joint Steering Committee for Revision of Anglo-American Cataloguing Rules
http://www.collectionscanada.ca/jsc/index.html

25. Boss, D., Library Portal Sales Soar, *Information systems report*, 2004, 3(5)

26. The Library of Congress Portals Applications Issues Group
http://www.loc.gov/catdir/lcpaig/

27. Z39.50 Maintenance Agency
http://www.loc.gov/z3950/agency/

28. SRU: Search/Retrieve via URL
http://www.loc.gov/standards/sru/

29. SRW: Search/Retrieve Web Service
http://www.loc.gov/standards/sru/srw/index.html

30. NISO Metasearch Gateway (MXG) Implementors Guide
http://www.niso.org/standards/resources/MI-MXG_v0_3.pdf

31. Sadeh, T., The Challenge of Metasearching, *New Library World,* 2004, 105(1198/1199), 104–112. Available at
http://www.exlibrisgroup.com/resources/metalib/The_Challenge_of_Metasearching.pdf

32. Luther, J., Trumping Google? Metasearching's Promise, *Library Journal,* 2003, October 1. Available at
http://www.libraryjournal.com/article/CA322627.html?display=searchResults&stt=001&text=trumping+google

33. Pace, A., Much ado About Metasearch, *American Libraries*, 2004 June/July. Available at
http://www.ala.org/ala/alonline/techspeaking/techspeak2004/Junejuly2004muchado.htm

34. NISO Metasearch Initiative
http://www.niso.org/committees/MS_initiative.html

35. Shibboleth
http://shibboleth.internet2.edu

36. Lynch, C.A., Access Management for Networked Information Resources, *ARL Newsletter* 201 (December 1998), pp. 3–7. Available at http://www.arl.org/newsltr/201/cni.html

37. The RSLP Collection Description http://www.ukoln.ac.uk/metadata/rslp

38. Dublin Core Collection Description Working Group http://dublincore.org/groups/collections/

39. The Information Environment Service Registry (IESR) http://www.iesr.ac.uk/

40. The OCKHAM Initiative http://www.ockham.org/

41. Berkeley Electronic Press (bepress) announces support for NISO MXG http://www.bepress.com/press111505.html

42. CrossRef http://www.crossref.org

43. Discovery Resolver http://www.endinfosys.com/software/resolver.html

44. ArticleLinker http://www.serialssolutions.com/articlelinker.asp

45. International DOI Foundation http://www.doi.org

46. CrossRef usage statistics http://www.crossref.org/01company/crossref_indicators.html

47. Crawford, W., OpenURL: Standards Can Be Fun! *American Libraries*, 2002 33(7). Available at http://www.ala.org/al_onlineTemplate.cfm?Section=2002columns&Template=/ContentManagement/ContentDisplay.cfm&ContentID=1787

48. Tennant, R., Digital Libraries Cross-Database Search: One-Stop Shopping, *The Library Journal*. October 15, 2001. Available at http://libraryjournal.com/article/CA170458

49. del.icio.us http://del.icio.us/

50. flickr.com http://www.flickr.com

51. Google Scholar: Support for Libraries http://scholar.google.com/scholar/libraries.html

52. The University of California Libraries bibliographic services task force, Rethinking how we provide bibliographic services for the University of California. December 2005.

53. O'Reilly, T., What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. 2005. Available at http://www.oreillynet.com/lpt/a/6228

54. Wikipedia entry for *Library 2.0* http://en.wikipedia.org/wiki/Library_2.0

55. Crawford, W., Library 2.0 and 'Library 2.0', *Cites and Insights* 2006 6(2) Available at http://cites.boisestate.edu/civ6i2.pdf

56. Vascarello, J., The Next Big Thing in Searching, *The Wall Street Journal*, January 2006.

57. FAST
http://www.fastsearch.com

58. Lucene
http://lucene.apache.org/

■ **Jenny Walker**
E-mail: jenny.walker@xrefer.com
Tel: +44 (0)20 7566 6380
Web: http://www.xrefer.com/

Correspondence address:
UK Office
Xrefer Ltd
36–37 Featherstone Street
London EC1Y 8QZ, UK

Biographical note

At the time of writing, Jenny Walker was Corporate Vice President Marketing for Ex Libris, based in London, UK. However, Jenny moved to Xrefer in June 2006 where she is Executive Vice President Marketing. She draws on her international experience in the library automation and information marketplaces. As Director of Technology Product Management at SilverPlatter (London) Jenny was involved in the early SFX research work at Ghent University which led to the development of the OpenURL Framework of Linking and the OpenURL standard. Jenny is a member of the NISO Standards Development Committee (SDC) and co-chair of the NISO Metasearch Initiative.