

Sadah, Tamar

Discovery and management of scholarly materials

ProInflow. 2011, vol. 3, iss. 2, pp. 4-22

ISSN 1804-2406

Stable URL (handle): <https://hdl.handle.net/11222.digilib/133714>

Access Date: 19. 02. 2024

Version: 20220831

Terms of use: Digital Library of the Faculty of Arts, Masaryk University provides access to digitized documents strictly for personal use, unless otherwise specified.

Tamar Sadeh

DISCOVERY AND MANAGEMENT OF SCHOLARLY MATERIALS: NEW-GENERATION LIBRARY SYSTEMS

Abstrakt:

S nedávným nástupem systémů nové generace, založených na oddělené architektuře, se změnil způsob shromažďování a spravování vědeckých prací. S využitím moderního uživatelského prostředí nabízí nové knihovní Discovery systémy širokou škálu materiálů dostupných skrze jednotné rozhraní daleko za možnostmi fyzických sbírek knihovny. Sbírkyně odborných prací tak dosahují globálního významu. Tyto systémy mají také vliv na množství dat shromážděných od institucí celého světa a zvyšují tak úroveň vyhledatelnosti odborných dokumentů. Nové systémy, určené od samého počátku ke správě všech typů odborných dokumentů, využívají technologického vývoje, sdílení metadat a spolupráce odborné komunity k optimalizaci knihovních služeb. Článek se zabývá některými současnými trendy v oblasti systémů nové generace a věnuje pozornost tomu, jak shromažďování informací (odborného obsahu, bibliografických metadat a uživatelských dat) v lokálním a individuálním kontextu buduje novou úroveň organizace a správy odborných dokumentů.

Klíčová slova: vyhledávací nástroj, žebříček relevance, index agregace, využití dat, recommender, oddělená architektura.

Abstract:

The discovery and management of scholarly materials have changed in recent years with the introduction of new-generation systems based on decoupled architecture. In addition to offering a modern user experience, new library discovery systems extend the scope of materials available through a single interface far beyond the physical collections of the library, reaching the wealth of scholarly collections of global significance. Such systems also leverage a body of usage data gathered from institutions worldwide to enhance the discoverability of materials. New management systems, built from the outset to manage all types of scholarly assets, harness technological advances, shared bibliographic metadata, and community collaboration to optimize library services. The paper examines some of the current trends in new-generation systems and focuses on the way in which collaboration among stakeholders and the aggregation of information—scholarly content, bibliographic metadata, and usage data—combine with the local and individual context to establish a new level of discovery and management of scholarly materials.

Keywords: search engine, relevance ranking, aggregated index, usage data, recommender, decoupled architecture

1 Introduction

The advent of metasearch technology, which was first presented by Ex Libris at the 10th CASLIN conference in 2001, marked the beginning of a new era for library systems. Metasearching and context-sensitive linking, which Ex Libris also introduced at about the same time, were the first forays by library-system vendors into an arena that until then had been the sole territory of information aggregators. Metasearch technology, which offers unified searching across heterogeneous information resources¹, and context-sensitive linking, which establishes library-controlled links between various components of the information landscape², helped libraries break down the traditional barriers between various silos of local materials and global content and added the institution's and the user's context to the information-seeking flow.

With the addition of electronic-resource management systems and digital-asset management systems shortly after, libraries gained the functionality required for managing the full spectrum of library content—print, electronic, and born digital content—and making discovery and access possible. However, the multiplicity and complexity of the library systems, along with the unparalleled scale of content and speed offered by non-library information systems (primarily Web search engines), triggered the drift of end users toward the latter. Librarians, on their part, were deploying less than optimal workflows for the management of all aspects of library services.

This paper addresses the way in which the discovery and management of scholarly materials have changed in recent years and examines some of the current trends. In particular, the paper focuses on the way in which collaboration among stakeholders and the aggregation of information—scholarly content, bibliographic metadata, and usage data—combine with the local and individual context to establish a new level of discovery and management of scholarly materials.

2 A Transformation in Progress

There is no doubt that the changing behavior of library users is part of an overall transformation that the society, economics, politics, and culture of our era are undergoing. The expectations of today's library users differ from those of the past, as the result of several factors, such as the immediacy of information and communication, the abundance of online activities in which people are typically

1 For a discussion of metasearching, see Tamar Sadeh, "The Challenge of Metasearching," *New Library World* 105, no. 1198/1199 (2004), doi: [10.1108/03074800410526721](https://doi.org/10.1108/03074800410526721).

2 Herbert Van de Sompel and Oren Beit-Arie, "Open Linking in the Scholarly Information Environment Using the OpenURL Framework," *D-Lib Magazine* 7, no. 3 (2001), <http://www.dlib.org/dlib/march01/vandesompel/03vandesompel.html>.

involved, and the effect of the social networks with which a great majority of users are engaged.

Web search engines, primarily Google, have had a profound influence on the way in which people find information. Started at the turn of the millennium as a means of reaching general information without mediators and without prior information literacy, Web search engines have shaped the way in which students seek scholarly information today. The ease of use, the immediacy, and the heterogeneous nature of information provided by Web search engines trigger expectations that library information systems were unable to match until recently.

In late 2005, an OCLC Online Computer Center survey of users' perceptions of library and information systems and services gave a clear picture of the changing user behavior³. OCLC has continued to monitor the behavior of users, including college students.^{4;5;6}

Already in 2005, library users had shifted from library services to Web search engines, online bookstores, and other Web-based services, such as blogs, online news, and e-mail, to satisfy their information needs without the help of the library. At that time, 89% of the undergraduate and graduate students that OCLC surveyed reported that they started their searches for information with Web search engines, whereas only 2% reported the library Web site as their starting point⁷. The 2007 OCLC survey showed an increase of usage of all Web services except one: the library Web site⁸.

The 2010 OCLC survey revealed changes that occurred with the emergence of social networks, Wikipedia, social sharing sites, ask-an-expert sites, and new communication channels such as Skype and Twitter. Although search engines clearly dominate information seeking (93% of college students use Web search

3 Cathy De Rosa, et al., *Perceptions of Libraries and Information Resources: A Report to the OCLC Membership* (Dublin, Ohio: OCLC Online Computer Library Center, 2005), <http://www.oclc.org/reports/2005perceptions.htm>.

4 Cathy De Rosa, et al., *College Students' Perceptions of Libraries and Information Resources: A Report to the OCLC Membership* (Dublin, Ohio: OCLC Online Computer Library Center, 2006), <http://www.oclc.org/reports/perceptionscollege.htm>.

5 Cathy De Rosa, et al., *Sharing, Privacy and Trust in Our Networked World: A Report to the OCLC Membership* (Dublin, Ohio: OCLC Online Computer Library Center, 2007), <http://www.oclc.org/reports/sharing/default.htm>.

6 Cathy De Rosa, et al., *Perceptions of Libraries, 2010: Context and Community* (Dublin, Ohio: OCLC Online Computer Library Center, 2011), http://www.oclc.org/reports/2010perceptions/2010perceptions_all.pdf.

7 Cathy De Rosa, et al., *Perceptions of Libraries and Information Resources: A Report to the OCLC Membership* (Dublin, Ohio: OCLC Online Computer Library Center, 2005), <http://www.oclc.org/reports/2005perceptions.htm>.

8 *Sharing, Privacy and Trust in Our Networked World: A Report to the OCLC Membership* (Dublin, Ohio: OCLC Online Computer Library Center, 2007), <http://www.oclc.org/reports/sharing/default.htm>.

engines to find online content), Wikipedia has gained considerable recognition by the college student population (88% of the students reported that they use Wikipedia to find information). Furthermore, social networks have become pivotal in the exchange of information—92% of college students use such networks, and two-thirds of them log on daily. The library Web site, on the other hand, was not mentioned as an initial starting point for information-seeking by any survey participants, although 57% of the students do use their library's site⁹ (a slight decline compared to 61% in 2005¹⁰).

As the result of a heavy reliance on social networks and social sharing sites, users have shifted from being consumers of so-called “objective” information systems—those that arrange a result list based on the degree of correlation between the query and the results, regardless of the specific user—to being the recipients of advice from other users, be they friends, other community members, or individuals whose path happened to cross that of the searcher. Although Web search engines do take a user's context into account to a certain degree, the user is likely to view a recommendation from a trusted person as more reliable than that of a Web search engine. Until recently, the providers of scholarly information systems took the opposite stand: they strived to remain objective (if one discounts librarians' selection of which materials to offer). Even though many systems have by now applied relevance ranking to result lists, the use of an assemblage of many criteria such as number of citations, number of downloads, and recency for sorting results (as opposed to sorting alphabetically or by date of publication) is a topic of debate among librarians.

Nevertheless, users still seek information that is authoritative and has been summarized by someone else, such as on ask-an-expert sites (for example, WikiAnswers). According to the 2010 OCLC survey, the number of respondents who use such sites increased by 136% from 2005, and the frequency of use increased as well. Online librarian-question services have become slightly more popular (10% of the respondents to the 2010 survey reported that they use such services, versus 8% in 2005), but they are still less popular than ask-an-expert sites. On the other hand, college students attribute greater trustworthiness and accuracy to library information systems than they did five years earlier (43% of the 2010 respondents indicated that information from library sources is more

9 Cathy De Rosa, et al., *Perceptions of Libraries, 2010: Context and Community* (Dublin, Ohio: OCLC Online Computer Library Center, 2011), http://www.oclc.org/reports/2010perceptions/2010perceptions_all.pdf.

10 Cathy De Rosa, et al., *Perceptions of Libraries and Information Resources: A Report to the OCLC Membership* (Dublin, Ohio: OCLC Online Computer Library Center, 2005), <http://www.oclc.org/reports/2005perceptions.htm>.

trustworthy than information from search engines, as opposed to 31% in 2005).^{11;12}

Commissioned by JISC and the British Library, the Centre for Information Behaviour and the Evaluation of Research (CIBER) at University College London (UCL) undertook a study aiming “to identify how the specialist researchers of the future, currently in their school or pre-school years, are likely to access and interact with digital resources in five to ten years’ time.”¹³ The investigation identifies some of the information-seeking behavior patterns that scholarly information systems will need to address: “in general terms, this new form of information seeking behaviour [digital information-seeking behavior] can be characterised as being horizontal, bouncing, checking and viewing in nature. Users are promiscuous, diverse and volatile and it is clear that these behaviours represent a serious challenge for traditional information providers, nurtured in a hardcopy paradigm and, in many respects, still tied to it.”¹⁴ The authors conclude that the information literacy of young people has not improved despite the exposure to technological tools from an early age. Furthermore, young people do not invest time in understanding their information need, developing search strategies, or evaluating the information that they find.

Much of the available research addresses the information-seeking behavior of academic library users, primarily undergraduates and graduate students. When the behavior of only graduate students and faculty members is examined, the findings are slightly different^{15;16;17;18;19}: although the great majority of academic

-
- 11 Cathy De Rosa, et al., *Perceptions of Libraries, 2010: Context and Community* (Dublin, Ohio: OCLC Online Computer Library Center, 2011), http://www.oclc.org/reports/2010perceptions/2010perceptions_all.pdf.
 - 12 Cathy De Rosa, et al., *Perceptions of Libraries and Information Resources: A Report to the OCLC Membership* (Dublin, Ohio: OCLC Online Computer Library Center, 2005), <http://www.oclc.org/reports/2005perceptions.htm>.
 - 13 Centre for Information Behaviour and the Evaluation of Research (CIBER), *Information behaviour of the researcher of the future* (London: CIBER, 2008), <http://www.ucl.ac.uk/slais/research/ciber/downloads/ggexecutive.pdf>.
 - 14 CIBER, *Information behaviour*, 9.
 - 15 Research Information Network, *Researchers and discovery services: Behaviour, perceptions and needs* (Research Information Network, 2006), <http://www.rin.ac.uk/our-work/using-and-accessing-information-resources/researchers-and-discovery-services-behaviour-perc>.
 - 16 Bradley M. Hemminger et al., “Information Seeking Behavior of Academic Scientists,” *Journal of the American Society for Information Science and Technology*, 58, no. 14 (2007), doi: 10.1002/asi.20686.
 - 17 Anne Gentil-Beccot et al., “Information Resources in High-Energy Physics: Surveying the Present Landscape and Charting the Future Course,” [arXiv:0804.2701v2](https://arxiv.org/abs/0804.2701v2) [cs.DL] (arXiv, 2008), doi: [10.1002/asi.20944](https://doi.org/10.1002/asi.20944).
 - 18 H. R. Jamali and D. Nicholas, “Information-Seeking Behaviour of Physicists and Astronomers,” *Aslib Proceedings*, 60, no. 5 (2008), doi: 10.1108/00012530810908184. Version used for this study: <http://eprints.rclis.org/16127/1/JAMALi-FINAL-preprint.pdf>.
 - 19 Laura L. Haines et al., “Information-Seeking Behavior of Basic Science Researchers: Implications for Library Services,” *Journal of the Medical Library Association*, 98, no. 1 (2010), doi: [10.3163/1536-5050.98.1.019](https://doi.org/10.3163/1536-5050.98.1.019).

users employ Web search engines, graduate students and faculty members tend to use discipline-specific information systems to satisfy some, or even most, of their information needs. However, it is clear that even in research communities, users are drawn to the simplicity, comprehensiveness, and ease of use of Web search engines. The CIBER study suggests that “it would be a mistake to believe that it is only students’ information seeking that has been fundamentally shaped by massive digital choice, unbelievable (24/7) access to scholarly material, disintermediation, and hugely powerful and influential search engines. The same has happened to professors, lecturers and practitioners. Everyone exhibits a bouncing/flicking behaviour, which sees them searching horizontally rather than vertically. Power browsing and viewing is the norm for all.”²⁰

Because most scholarly materials are discoverable through multiple interfaces, users may well be able to obtain the same materials through Web search engines and academic systems, although Web search engines provide an easier and faster route to these materials. However, Web search engines come with their own drawbacks, particularly the limited search and filtering options available to users and a search scope that comprises a universe of materials of unequal quality. Furthermore, with the growing amount of available data, even Web search engines have lost some of their attraction, as revealed by the 2010 OCLC survey: “only” 83% of college students begin their searches using search engines, as opposed to 89% in 2005.^{21;22} In addition, library-driven services such as bibliographic tools and citation analyses are not available through Web search engines. Hence, most searchers rely on more than one type of information system and typically use both scholarly information systems and Web search engines.

In a report commissioned by the Bibliographic Services Task Force of the University of California, the authors conclude that their “users expect simplicity and immediate reward and Amazon, Google, and iTunes are the standards against which we are judged. Our current systems pale beside them.”²³ The challenge for libraries, therefore, is to determine how they wish to portray themselves to their users and how they can best serve the institutions to which they belong.

Looking for ways to retain their users and maintain their hegemony as information providers, libraries have started considering new approaches—user-

20 CIBER, *Information behaviour*, 8.

21 Cathy De Rosa, et al., *Perceptions of Libraries, 2010: Context and Community* (Dublin, Ohio: OCLC Online Computer Library Center, 2011), http://www.oclc.org/reports/2010perceptions/2010perceptions_all.pdf.

22 Cathy De Rosa, et al., *Perceptions of Libraries and Information Resources: A Report to the OCLC Membership* (Dublin, Ohio: OCLC Online Computer Library Center, 2005), <http://www.oclc.org/reports/2005perceptions.htm>.

23 University of California Libraries Bibliographic Services Task Force, *Rethinking How We Provide Bibliographic Services for the University of California* (University of California, 2005), <http://libraries.universityofcalifornia.edu/sopag/BSTF/Final.pdf>.

centric solutions that replace their traditional online user interfaces. However, in order to address users' expectations regarding the interface, the breadth and relevance of services, and the comprehensiveness of the body of information available through the system, libraries had to undergo a major conceptual shift.²⁴

3 The Challenges for Library Management Systems

While worrying about the drift of users to other information systems, librarians must also meet the challenge of efficiently managing their assets. Because the various systems in the library were developed over time to support specific needs as they arose, the overall library environment became complex and workflows became cumbersome. Finally, the economic crisis of the last decade coupled with the resulting budget cuts led to the realization that to retain their role as the providers of quality information, libraries would have to operate in a different way.

Recent reports indicate that libraries are undergoing considerable changes.²⁵ New trends include the following:

- The nature of collection development is changing and is becoming much more driven by patrons' requests. Shifting to e-materials, libraries are expanding their collections and dismantling the traditional distinction between local and remote content—be it owned by other libraries, under subscription, available on demand, or open access. The types of materials

24 For a discussion of user preferences, see Tamar Sadeh, "Time for a Change: New Approaches for a New Generation of Library Users," *New Library World* 108, no. 7/8 (2007), doi: [10.1108/03074800710763608](https://doi.org/10.1108/03074800710763608); "User-Centric Solutions for Scholarly Research in the Library," *LIBER Quarterly* 17, no. 3/4 (2007), http://liber.library.uu.nl/publish/issues/2007-3_4/index.html?000215; "A Model of Scientists' Information Seeking and a User-Interface Design" (PhD thesis, City University London, 2010).

25 ACRL Research Planning and Review Committee, "2010 top ten trends in academic libraries," *College & Research Libraries News* 71, no. 6 (June 2010), <http://crln.acrl.org/content/71/6/286.full>; Association of College and Research Libraries (researched by Megan Oakleaf), *Value of Academic Libraries: A Comprehensive Research Review and Report* (Chicago: Association of College and Research Libraries, 2010), <http://www.acrl.org/value/>; Council on Library and Information Resources, *No Brief Candle: Reconceiving Research Libraries for the 21st Century* (Washington, D.C.: Council on Library and Information Resources, 2008), <http://www.clir.org/pubs/reports/pub142/pub142.pdf>; L. Johnson, A. Levine, and R. Smith, *The 2009 Horizon Report* (Austin, Texas: The New Media Consortium, 2009), <http://wp.nmc.org/horizon2009/>; L. Johnson, A. Levine, R. Smith, and S. Stone, *The 2010 Horizon Report* (Austin, Texas: The New Media Consortium, 2010), <http://wp.nmc.org/horizon2010/>; Matthew P. Long and Roger C. Schonfeld, *Ithaka S+R Library Survey 2010: Insights from U.S. Academic Library Directors* (ITHAKA, 2010), <http://www.ithaka.org/ithaka-s-r/research/ithaka-s-r-library-survey-2010/insights-from-us-academic-library-directors.pdf>; James Michalko, Constance Malpas, and Arnold Arcolio, *Research Libraries, Risk and Systemic Change* (Dublin, Ohio: OCLC Research, 2010), <http://www.oclc.org/research/publications/library/2010/2010-03.pdf>; David J. Staley and Kara J. Malenfant, *Futures Thinking for Academic Librarians: Higher Education in 2025* (Association of College and Research Libraries, 2010), <http://www.acrl.org/ala/mgrps/divs/acrl/issues/value/futures2025.pdf>.

are becoming more varied, and libraries are likely to be involved in dealing with new material types, such as data sets.

- Because of continuous budget cuts all over the world, libraries are becoming more vulnerable and have to demonstrate measurable value for the institutions that they serve. To stay relevant, libraries are trying to engage in new tasks, traditionally not within the library domain, such as teaching and conducting research. Topics such as scholarly communication and intellectual property are increasing in relevance with the involvement of libraries in the publishing of locally created materials, such as theses, dissertations, and course materials. Libraries are also active in digitization projects and are likely to become involved in the preservation of digital materials.
- Driven by the need to increase efficiency and adopt cost-effective processes, libraries are seeking assessment tools and analytics that can provide more insight into library activities and help library executives in their decision making. True to their long tradition of collaboration, which is now supported by technological advances, libraries are collaborating more with other groups in their institution and beyond—primarily suppliers and other libraries.
- The rapid pace of technological changes and the ongoing library investment in hardware, operations, and skilled staff is motivating library decision makers to seek an infrastructure that can ease the burden of managing information technology or exempt libraries altogether from those tasks. Cloud-computing environments, now offered by many vendors, are already gaining momentum and are likely to continue attracting libraries.

While many such changes trigger discussions about the role of the library, it is clear that no matter what new missions libraries undertake, existing software solutions fall short in providing optimal support for all current and future library activities because of architecture that is neither flexible nor scalable enough. New software solutions addressing the new needs are likely to take the lead.

4 The Introduction of Decoupled Architecture

One of the main challenges in adapting library system environments to better serve end users is the systems' focus on library staff workflow and the tight bond between the management of the library assets and the way in which the library makes these assets available to end users. The librarian-centric focus originates from the fact that librarians perform most of the tasks in the library; furthermore, library vendors sell to librarians, not to end users.

Two main reasons make it difficult for information systems to both fulfill the expectations of users and accommodate librarians' need for optimized systems. First, to date, each system in the library has tended to deal with only a single type of material—print, electronic, or digital—whereas end users expect to find everything they need in one place, regardless of format; given the architecture of the existing systems, it is not likely that one such system can be extended to cover all materials. This “silo” approach to managing materials is also challenging for librarians who attempt to understand what is going on across the library's collections. Second, the traditional model of library workflows—each of which was developed for a single type of material—coupled with a lack of efficient technical channels for collaboration among libraries and other stakeholders does not lend itself to an efficient, cost-effective infrastructure.

In 2006, Ex Libris, a provider of library automation solutions, introduced “decoupled architecture” as the cornerstone of its future offering. Other vendors, such as OCLC and Innovative Interfaces, have demonstrated the same vision. This architecture, which separates the user experience from the management of the library collections and the library services, is based on data exchange between the discovery layer and the management layer of the library systems; each such layer is designed around the needs of its users—end users and librarians, respectively.

Because retaining their target audience was an urgent task for most libraries, the first new systems based on decoupled architecture were discovery systems. With such systems in place, the libraries' current solutions, such as integrated library systems and digital-asset management systems that were tailored to the needs of the librarians, continued to fulfill their administrative functions.

Embraced by all stakeholders, the decoupled architecture model has already been implemented in discovery systems developed by library system vendors, other software vendors, information providers, and open-source communities since 2007. While library system vendors have been addressing both components of decoupled architecture, others have been offering only the discovery layer.

Decoupled architecture provides libraries with several benefits:

- **Flexibility:** Because it is not tied to the library's management systems, the user experience is free of constraints related to specific data structures and administrative workflows and hence can be designed and tailored to the needs of end users.
- **Unification:** Data coming from heterogeneous, harvestable information resources—both local and remote—form one source of information for end users. The fact that data elements originate from separate sources has become irrelevant to the discovery process.

- **Data enhancement:** Once it harvests the data from the administrative systems, the new system easily enhances that data with additional elements from other resources, typically from third parties. Such data elements include images of book covers, tables of contents, descriptions, and reviews.
- **Integration:** Adherence to the newest standards of interoperability enables a library to easily integrate its new interface into an existing environment and the various systems in the user's workspace, such as course management systems, social network sites, and reference management tools. Furthermore, discovery interfaces are now available for mobile devices and hence have become more accessible and practical.
- **Analytics:** Providing access to both local and remote collections through a single interface enables a library to easily gather information about user behavior and analyze the information for purpose of development planning and the adjustment of library collection and services.
- **Step-by-step evolution:** With a system that deploys decoupled architecture, libraries have started offering a new, improved user experience without disturbing the behind-the-scenes administration. Similarly, libraries can add, modify, and replace components of their administrative systems without affecting the user experience.

Decoupled architecture lays the groundwork for a robust discovery service for end users, on the one hand, and a unified resource-management system for librarians, on the other hand. The first service already exists: end-user-centric discovery solutions, including Primo from Ex Libris, appeared on the market in 2007 and have been embraced by national, academic, and public libraries everywhere (Primo alone has been selected by more than 800 institutions worldwide). Evidence shows that, indeed, the number of searches in library-based systems increases dramatically when the library introduces a solution whose user experience is tailored to the library community's needs.²⁶

²⁶ According to statistics cited by New York University in an e-mail exchange to the NGC4LIB listserv in October 2009, the average number of search sessions after the implementation of Primo almost tripled compared to the average number before the implementation, and the average session time decreased from 15 minutes to six minutes, indicating that users are finding what they want much faster. At a user group meeting in early 2010, Yonsei University (in Seoul, South Korea) reported a similar increase in the number of search sessions after the implementation of Primo. In addition, the Primo search session time decreased to 10%, on average, of the previous system's session time for a particular set of queries. The difference can be attributed to the display of all the titles on the first page in Primo as opposed to on the second to fifteenth page in the previous system.

5 Current Discovery Challenges

No longer considered new, discovery systems resemble each other in many respects. All such systems are fast and offer heterogeneous scholarly materials, a modern user interface, one search box for searching all types of materials, faceted browsing, and relevance-ranked result lists. However, one significant difference between discovery systems lies in the way in which they deal with global offerings in a local context. This difference is manifested in the breadth and quality of the materials offered, the customizability of the search scope, the ease of integration with local library services, the sophistication and adaptability of relevance ranking, the inclusion of recommendations as search aids, and the provision of functionality for local branding.

There is no doubt that the content available for scholarly discovery systems sets the point of departure for the discovery process. On the one hand, the content should be as far-reaching as possible and include all types of materials in every possible discipline. On the other hand, if discovery systems provide only relevant, high-quality scholarly results, academic users are more likely to turn to them (as opposed to Web search engines).

Because collections are no longer bound to space and location, the whole spectrum of scholarly information is open for discovery, and the boundaries set by institutional physical or digital holdings are no longer relevant. Furthermore, in today's academic environment, information needs can rarely be satisfied by local resources alone. However, the integration between global searching and the local physical holding is crucial in many disciplines and for specific user communities.

The biggest challenge of discovery systems is how to provide users with the most relevant items in the immense landscape of available content. Thus, new tools have been added to such systems to help users find specific items that they are looking for or items that will satisfy a broader search query. For example, faceted navigation helps users quickly refine their result list and focus on subsets of it,^{27;28} and a display of recommendations based on other users' selections draw one's attention to items similar to a given item. However, familiarity with Google and other search engines leads users of discovery systems to scan only the first results; hence, relevant items can easily remain unnoticed if they are not displayed on the first page. Relevance ranking, whose purpose is to highlight what the system deems the most appropriate materials for the particular query, has become a major factor in satisfying user needs, together with immediate delivery (or, in the

27 Tamar Sadeh, "Multiple Dimensions of Search Results" (paper presented at the Analogous Spaces Interdisciplinary Conference, Ghent University, Belgium, May 15-17, 2008).

28 Marti A. Hearst, "Clustering Versus Faceted Categories for Information Exploration," *Communications of the ACM*, 49, no. 4 (2006), <http://people.ischool.berkeley.edu/~hearst/papers/cacm06.pdf>.

case of physical items, immediate OPAC-based services), and in increasing the value of the library for the user and for the institution.

6 What Does the Library Offer the User?

If in the past a library was judged by the number of volumes it held, today scholarly information is broad and borderless. However, one of the main roles of librarians—the selection of appropriate resources—is no less applicable today than in the past. Libraries need to differentiate themselves from Web search engines by ensuring the quality and breadth of the local and remote information that is provided and by making the information easy to use through preprocessing (such as the detection and grouping of duplicates) and by integrating scholarly search functionality into the user's environment (by supporting, for example, institutional single sign-on and embedding search boxes in various institutional and external systems). Given the huge quantity of available data, libraries can reduce information overload by setting an initial search scope that is more appropriate to their communities and by using techniques such as the grouping of similar materials. Although discovery systems match Web search engines when it comes to ease of use and speed, speed is not measured only by the amount of time that elapses until a result list is displayed. Much more important is the amount of time that an information system takes to satisfy an information need and provide the user with the desired outcome. In this respect, the control that libraries have over the search scope and the relevance-ranking algorithms, the deployment of services such as recommendations, and the immediate delivery of the materials significantly decrease the time that users take to find items of interest and amplify the value of the library's services for its users.

When defining the search scope, libraries should be addressing the “long tail” of information resources that are of utmost importance to some of their users. While it is likely that almost all information needs of undergraduates can be satisfied by the most popular information resources, many researchers require more specialized information that might dictate the adoption of various searching technologies. A discovery system that is based on one technology and is not flexible enough to provide access to information resources of all types cannot become the ultimate search entry point for many users.

Depending on a user's information need—an exploratory search for items on a particular topic or a search for a specific item—an information system might have to use more than one method of selecting the most appropriate results. When a user is looking for a specific item, the system should display that item at the top of the first page of results. However, an exploratory search is more complex, because the broader the search is, the greater the quantity of relevant results.

Furthermore, in an exploratory search the information need is not necessarily well defined; the user might not be sure what is needed, and the way in which the user phrases the query might not be clear. Because undergraduates—who are typically less adept at phrasing their information needs—tend to conduct exploratory searches²⁹, addressing such searches adequately is of great importance. In addition to identifying the items that are most likely to fulfill the user's needs and putting them at the top of the result list, the system should draw the user's attention to items that are likely to be relevant although they might not be on the result list.

In a manner that is similar to human interaction between two parties, in which each person adjusts to the other in tone, language, and content, information systems need to “understand” the user's context as well as the value of the information that they offer, regardless of the specific query. While the usefulness of the information available through a system lies in the aggregation of global data—both the content itself and measures that are associated with it, such as the journal impact factor and usage statistics—the user is always an individual who is part of a local and wider community and has a specific information need at a given moment. “Awareness” of the user context, such as the person's discipline, can help an information system adjust the relevance ranking so that items related to the user's discipline are ranked higher. Similarly, the academic level of the user may indicate the degree of applicability of items that are more general.

Usage data has proven to be a most valuable resource for information systems and management systems in the scholarly arena.³⁰ From such data, a system can generate metrics for evaluating the significance of items and for associating items with each other; then the system can feed the results back to the user through recommendations (such as those provided by the bX article recommender service) and relevance ranking and can aid librarians in collection-development decisions. However, the gathering of usage data is most meaningful when the data are aggregated across institutions rather than related to the few individuals that happen to be at one institution.

Recommendations from the bX service expand the search results to include items that are not retrieved by the query yet are clearly relevant. Such recommendations are highly valuable for cross-disciplinary research and for research in an area with which the user is less familiar. A user who does not know the applicable

29 According to a survey by University of Minnesota Libraries, over 70% of undergraduate searches are exploratory; as researchers become more knowledgeable in their area, they tend to search more for known items. *Discoverability: Phase 2 Final Report* (2010), <http://conservancy.umn.edu/bitstream/99734/3/DiscoverabilityPhase2ReportFull.pdf>

30 Johan Bollen and Herbert Van de Sompel, “An Architecture for the Aggregation and Analysis of Scholarly Usage Data (paper presented at JCDL '06, Chapel Hill, North Carolina, June 11–15, 2006, ACM 1-59593-354-9/06/0006), http://public.lanl.gov/herbertv/papers/jcdl06_accepted_version.pdf.

terminology is likely to miss relevant results; however, recommendations that are displayed along with an item on a result list aid such users by highlighting materials that are relevant even though they may not share the same keywords.

The system's evaluation of the user's context serves the entire information-seeking process, not just searches. Services such as those related to evaluating materials, accessing them, integrating them into the user's space (for example, enabling the user to save the citation or bookmark the item), and accessing other relevant materials should be part of the institutional context. In addition, because more and more users carry a mobile device, they can identify and make use of services that are relevant in their current location.

The user's context brings up issues of privacy, which is of great concern to libraries; however, gaining more information about a user is the key to tailoring the system's behavior to that user, just as in human interaction.

7 New-Generation Library Management Systems

While the new-generation discovery systems, available since 2007, were the first manifestations of decoupled architecture, corresponding library management systems started to emerge toward the end of the last decade. No such system is in full-scale production yet, although two systems—Alma from Ex Libris³¹ and Web-scale Management Services (WMS) from OCLC³²—have been made available to selected libraries for specific functionalities.

Designed from scratch rather than as an extension of existing products, the new-generation management systems have the privilege of presenting an optimal infrastructure that is likely to serve libraries for the next decade and more. Ideally, such systems address the following aspects:

- The unification of data structures and consolidation of processes, regardless of the type of material, its location, and its type of ownership. Such unification and consolidation open up opportunities to build more efficient library services by creating optimized workflows that are supported by comprehensive information gathering. Furthermore, unified data structures enable business processes to be automate and facilitate the sharing of data among libraries.
- Collaboration among stakeholders—other libraries, parent institutions, suppliers, and the scholarly community—to optimize processes and collections. With such collaboration, community efforts can be leveraged to

31 For information about Alma, see http://www.exlibrisgroup.com/category/URM_ResourceCenter.

32 For information about Web-scale Management Services, see <http://www.oclc.org/webscale>.

reduce library-specific investment in generating or obtaining information (for example, metadata) that is already available to the community, thus supporting streamlined processes such as acquisition on demand. Furthermore, collaboration can be extended to activities such as joint collection building, with physical items shared among institutions.

- Network-level architecture based on cloud services. This type of architecture increases the return on investment and reduces the total cost of ownership of the library's infrastructure while providing easy access to shared data and services.
- The leveraging of intelligence gathered by the system to generate analytics that support both informed decisions regarding all library activities (primarily collection development) and the tailoring of a personal context for users to improve the discovery and delivery process
- Integration of library services within the institutional environment to facilitate strategic library support of the institution's mission statement and of institutional activities such as teaching and research, thus demonstrating the library's value within its environment
- A core system that supports the building of software extensions by libraries. Such extensions enable libraries to add services and local adaptations.

The handling of metadata demonstrates, once again, the way in which aggregation provides a springboard to greater efficiency. While the optimization of metadata handling is tightly bound to the capability of a system to leverage large aggregates of metadata shared by libraries, it is crucial that the system operate in the context of the specific library and balance the global sharing with the local library's characteristics and needs. The individuality of a library—primarily, its unique collections—should be combined with the global metadata repository shared by many libraries to achieve optimal flexibility while supporting the efficiency of processes.

8 Conclusions

The past decade has seen a fundamental change in the way in which libraries have been providing services to their users. During that time, libraries expanded their services to offer a much greater volume and variety of materials through multiple systems. Yet, because of global changes outside the boundaries of libraries, users drifted to other spaces and libraries found themselves looking for ways to remain relevant.

Decoupled architecture, through which discovery systems support the needs and expectations of end users while administrative systems serve librarians, has been embraced by industry stakeholders in recent years. Hundreds of libraries have adopted discovery systems that were developed by library software vendors, such as Ex Libris and Innovative Interfaces; information providers, such as EBSCO and Serials Solutions (a ProQuest company); the open-source community (from which comes the VUFind portal, for example), and other providers, such as Endeca. The new-generation systems based on this architecture—both discovery systems and management systems—leverage technological advances and the aggregation of content, bibliographic data, and usage data to deliver library services on a new scale.

While enabling libraries to expand their offerings to their users, on the one hand, and optimizing administrative processes, on the other hand, software systems need to help libraries maintain their individuality and set the appropriate context for their users. By doing so, libraries can better serve their users and add greater value to their institutions.

References

1. ACRL Research Planning and Review Committee. “2010 top ten trends in academic libraries.” *College & Research Libraries News* 71, no. 6 (June 2010): 286-292. <http://crln.acrl.org/content/71/6/286.full>.
2. Association of College and Research Libraries (researched by Megan Oakleaf). *Value of Academic Libraries: A Comprehensive Research Review and Report*. Chicago: Association of College and Research Libraries, 2010. <http://www.acrl.ala.org/value/>.
3. Bollen, Johan, and Herbert Van de Sompel. “An Architecture for the Aggregation and Analysis of Scholarly Usage Data. Paper presented at JCDL '06, Chapel Hill, North Carolina, June 11–15, 2006. ACM 1-59593-354-9/06/0006. http://public.lanl.gov/herbertv/papers/jcdl06_accepted_version.pdf.
4. Centre for Information Behaviour and the Evaluation of Research (CIBER). *Information behaviour of the researcher of the future*. London: CIBER, 2008. <http://www.ucl.ac.uk/slais/research/ciber/downloads/ggexecutive.pdf>.
5. Council on Library and Information Resources. *No Brief Candle: Reconceiving Research Libraries for the 21st Century*. Washington, D.C.: Council on Library and Information Resources, 2008. <http://www.clir.org/pubs/reports/pub142/pub142.pdf>.

6. De Rosa, Cathy, Joanne Cantrell, Diane Cellentani, Janet Hawk, Lillie Jenkins, and Alane Wilson. *Perceptions of Libraries and Information Resources: A Report to the OCLC Membership*. Dublin, Ohio: OCLC Online Computer Library Center, 2005.
<http://www.oclc.org/reports/2005perceptions.htm>.
7. De Rosa, Cathy, Joanne Cantrell, Janet Hawk, and Alane Wilson. *College Students' Perceptions of Libraries and Information Resources: A Report to the OCLC Membership*. Dublin, Ohio: OCLC Online Computer Library Center, 2006. <http://www.oclc.org/reports/perceptionscollege.htm>.
8. De Rosa, Cathy, Joanne Cantrell, Andy Havens, Janet Hawk, and Lillie Jenkins. *Sharing, Privacy and Trust in Our Networked World: A Report to the OCLC Membership*. Dublin, Ohio: OCLC Online Computer Library Center, 2007. <http://www.oclc.org/reports/sharing/default.htm>.
9. De Rosa, Cathy, Joanne Cantrell, Matthew Carlson, Peggy Gallagher, Janet Hawk, and Charlotte Sturtz. *Perceptions of Libraries, 2010: Context and Community*. Dublin, Ohio: OCLC Online Computer Library Center, 2011.
http://www.oclc.org/reports/2010perceptions/2010perceptions_all.pdf.
10. Gentil-Beccot, Anne, Salvatore Mele, Annette Holtkamp, Heath B. O'Connell, and Travis C. Brooks. "Information Resources in High-Energy Physics: Surveying the Present Landscape and Charting the Future Course." [arXiv:0804.2701v2](https://arxiv.org/abs/0804.2701v2) [cs.DL]. arXiv, 2008. doi: [10.1002/asi.20944](https://doi.org/10.1002/asi.20944). (Journal reference: *Journal of the American Society for Information Science and Technology* 60, no. 1 (2009): 150-160.)
11. Haines, Laura L., J. Light, D. O'Malley, and F. A. Delwiche. "Information-Seeking Behavior of Basic Science Researchers: Implications for Library Services." *Journal of the Medical Library Association* 98, no. 1 (2010): 73–81. doi: [10.3163/1536-5050.98.1.019](https://doi.org/10.3163/1536-5050.98.1.019).
12. Hearst, Marti A. "Clustering Versus Faceted Categories for Information Exploration." *Communications of the ACM* 49, no. 4 (2006): 59-61.
<http://people.ischool.berkeley.edu/~hearst/papers/cacm06.pdf>.
13. Hemminger, Bradley M. Dihui Lu, K. T. L. Vaughan, and Stephanie J. Adams. "Information Seeking Behavior of Academic Scientists." *Journal of the American Society for Information Science and Technology* 58, no. 14 (2007): 2205–2225. doi: [10.1002/asi.20686](https://doi.org/10.1002/asi.20686).
14. Jamali, H. R., and D. Nicholas. "Information-Seeking Behaviour of Physicists and Astronomers." *Aslib Proceedings* 60, no. 5 (2008): 444-462.

- doi: 10.1108/00012530810908184. Version used for this study:
<http://eprints.rclis.org/16127/1/JAMALIi-FINAL-preprint.pdf>.
15. Johnson, L., A. Levine, and R. Smith. *The 2009 Horizon Report*. Austin, Texas: The New Media Consortium, 2009.
<http://wp.nmc.org/horizon2009/>.
 16. Johnson, L., A. Levine, R. Smith, and S. Stone. *The 2010 Horizon Report*. Austin, Texas: The New Media Consortium, 2010.
<http://wp.nmc.org/horizon2010/>.
 17. Long, Matthew P., and Roger C. Schonfeld. *Ithaka S+R Library Survey 2010: Insights from U.S. Academic Library Directors*. ITHAKA, 2010.
<http://www.ithaka.org/ithaka-s-r/research/ithaka-s-r-library-survey-2010/insights-from-us-academic-library-directors.pdf>.
 18. Michalko, James, Constance Malpas, and Arnold Arcolio. *Research Libraries, Risk and Systemic Change*. Dublin, Ohio: OCLC Research, 2010.
<http://www.oclc.org/research/publications/library/2010/2010-03.pdf>.
 19. Research Information Network. *Researchers and discovery services: Behaviour, perceptions and needs*. Research Information Network, 2006.
<http://www.rin.ac.uk/our-work/using-and-accessing-information-resources/researchers-and-discovery-services-behaviour-perc>.
 20. Sadeh, Tamar. "The Challenge of Metasearching." *New Library World* 105, no. 1198/1199 (2004): 104-112. doi: [10.1108/03074800410526721](https://doi.org/10.1108/03074800410526721).
 21. Sadeh, Tamar. "A Model of Scientists' Information Seeking and a User-Interface Design." PhD thesis, City University London, 2010.
 22. Sadeh, Tamar. "Multiple Dimensions of Search Results." Paper presented at the Analogous Spaces Interdisciplinary Conference, Ghent University, Belgium, May 15-17, 2008.
 23. Sadeh, Tamar. "Time for a Change: New Approaches for a New Generation of Library Users." *New Library World* 108, no. 7/8 (2007): 307-316. doi: [10.1108/03074800710763608](https://doi.org/10.1108/03074800710763608).
 24. Sadeh, Tamar. "User-Centric Solutions for Scholarly Research in the Library." *LIBER Quarterly* 17, no. 3/4 (2007).
http://liber.library.uu.nl/publish/issues/2007-3_4/index.html?000215
 25. Staley, David J., and Kara J. Malenfant, *Futures Thinking for Academic Librarians: Higher Education in 2025*. Association of College and Research Libraries, 2010.
<http://www.acrl.org/ala/mgrps/divs/acrl/issues/value/futures2025.pdf>.

26. University of California Libraries Bibliographic Services Task Force. *Rethinking How We Provide Bibliographic Services for the University of California*. University of California, 2005.
<http://libraries.universityofcalifornia.edu/sopag/BSTF/Final.pdf>.
27. University of Minnesota Libraries. *Discoverability: Phase 2 Final Report*. University of Minnesota, 2010.
<http://conservancy.umn.edu/bitstream/99734/3/DiscoverabilityPhase2ReportFull.pdf>.
28. Van de Sompel, Herbert, and Oren Beit-Arie. "Open Linking in the Scholarly Information Environment Using the OpenURL Framework." *D-Lib Magazine* 7, no. 3 (2001).
<http://www.dlib.org/dlib/march01/vandesompel/03vandesompel.html>.