

ABUNDANCE, ATTENTION, AND ACCESS: OF PORTALS AND CATALOGS ¹

by Sarah E. Thomas, University Librarian, Cornell University Library

The world's information resources are abundant, but time is a scarce commodity. The ideal discovery tool, therefore, is one which consults omnivorously, but which returns a selection of relevant results in rapid sequence. Searchers find what they need promptly without having to wade through a vast assortment of tangentially related, inaccurate, or otherwise deficient data. It costs little to build and operate, and it yields a high degree of user satisfaction because it delivers reliable information in a timely manner with relative ease. Such a tool is still imaginary, although it could become a reality in the near future if librarians organize themselves appropriately and commit the resources to design it.

One precursor of this discovery tool is the library catalog. Over the past century, the catalog has achieved a reputation as a dependable starting point for people seeking particular kinds of information: published information appearing chiefly in books or journals that has undergone a formal review and editing process. Libraries have further filtered these publications through their selection practices, in which collection policies guide the acquisition of texts to meet the needs of a particular community. The titles chosen by bibliographers represent quality because they have been at least twice vetted, once by professional editors and once by library subject specialists. Once acquired, the volumes receive added value through their integration with other related materials held by the institution. Through the cataloging of these chosen items, they gain in significance as trained experts distill their essence into a standard bibliographic description. Subject headings and classification contribute further to the access of these materials. Widespread adoption of the *Anglo-American Cataloguing Rules*, Library of Congress Subject Headings, Library of Congress or Dewey Decimal Classification, and the MARC format has enabled the creation of virtual union catalogs with a number of superior attributes. The catalog yields highly codified citations to quality publications in a predictable and dependable fashion. Names are usually authoritative and consistent. These features instill a strong degree of user confidence in the results of their searches. Libraries enhance this trust by ensuring that the materials cited are readily available for consultation, since titles cataloged are either locally held or obtainable through resource-sharing agreements. Furthermore, the library's commitment to preserve the documents in its custody guarantees enduring access.

Over the past few years, library management systems have matured and now present catalog users with sophisticated online public access capabilities that include web access, improved keyword searching, relevance ranking, ability to limit searches by date or other fields, and reference linking.

The catalog's trustworthiness comes at a price. It depends heavily on human intelligence to apply principles of organization to bibliographic works, and the complex rules needed to exact consistency have proven expensive to apply. Original cataloging can cost \$50 to \$100 per title. (Of course, so can the answer to a single reference question.) Although libraries have reengineered their processes to take greater advantage of cataloging copy and automated assistance in the creation of bibliographic records, the world's output of publications has outstripped the resources available to control them using traditional cataloging procedures. Compounding the situation, there has been a rise in interest in other formats, such as films, recordings, photographs, manuscripts, maps, and, now, digital objects. At the same time that the aperture of scholarship has been widening, libraries have been diverting resources devoted to cataloging to other functions, as ARL statistics reveal: the percentage of the total professional workforce in ARL libraries employed as catalogers has dropped by 25% from 1990 through 1998. ²

The declining role of the catalog was highlighted in a *New York Times* article headlined "Choosing

Quick Hits Over the Card (sic) Catalog. " While conceding that "libraries are organized and easily navigated, " the author reports "students prefer diving into the chaotic whirl of the Web to find information. " ³ More and more students and faculty turn to the Internet as their first point of contact for information. Even deans at top-ranked universities have confessed that they get satisfactory results from "Ask Jeeves, " or that they question the need for physical libraries since they find all the references they require for their writing available online. Over the past five years, there has been a strong trend in many quarters of the university to prefer electronic resources because of the ease and speed of access. More recently, the concept of portal has emerged as a unifying device for the wide variety of data made available on the web. The portal serves as a starting point for searching, and it usually offers an array of associated services such as news, related products, and reference tools, often clustered around a theme or discipline.

Portals differ from the catalog in a number of significant ways. They provide access to a wide range of materials, much of which would fall in the "unpublished" category. Harvesting of data occurs through the use of algorithmic programs, and there is little human cataloging effort. The largest catalog departments process 250,000 volumes annually, while Google' s search engine indexes millions of web pages weekly and claims to hold links to more than a billion URLs. To cope with the effect of large response sets, portals have moved to include relevance ranking. Still it is common to experience duplicates, false drops, and dead ends in many searches conducted with the discovery tools popular on the web. In addition, the free service they provide in locating information is often subsidized through advertising, which sometimes even affects placement in search results. Proprietary information, such as licensed databases to which libraries often subscribe, is inaccessible through most commonly used portals.

Despite these deficiencies, portals offer many compelling features. They yield up-to-the-moment information, and they deliver a vast quantity and rich variety of resources, including full-text, images, and sound. The web' s ability to support linking enables the searcher to move with ease from document to document, and to capture material for his own use and facile manipulation. Customization and even personalization of searching and retrieval are other popular characteristics of search engines and portals. Looney and Lyman observe that " portals gather a wide variety of information resources into a single ' one-stop' web page, helping the user to avoid being overwhelmed by ' infoglut' or feeling lost on the Web. " ⁴

The enthusiasm with which people embrace the web, and the skepticism which many of these same digerati express about traditional libraries, has been sobering for some librarians. Information professionals note the flaws in search engines and in digital " library" content, which is often unstable, of dubious quality and authenticity, and incomplete. Their cautionary observations go unheeded among the hype and the reality of the flexibility of the web and the wealth of resources it encompasses. It is unlikely that the catalog--in its present incarnation--can ever be the tool that provides the principal mode of access to information for students and scholars.

An alternative to both the dot-com portals and the catalog is the scholars portal. A concept developed by a number of library leaders in 1999, the scholars portal, described at length by Jerry Campbell, would promise high-quality content of the type consistent with the support of research and scholarship, and it would incorporate the suite of services, such as reference, familiar to library users. ⁵It would build on the collaborative strength of libraries to create a destination for scholars that would convey the traditional values of careful selection, expert organization, skilled retrieval and delivery, interpretation, and endurance that have characterized libraries over the past century.

The scholars portal has not yet come to fruition, but the energy and debate surrounding the proposal are healthy signs of the realization that libraries must transform themselves and create new services in the

near future. One area that is ripe for review is the proportion of time dedicated to the cataloging of books and journals. In the past decade and more, catalogers have become increasingly productive, using copy cataloging and automated workstations to lower the cost of cataloging. Library administrators have seized resources previously devoted to cataloging and reallocated them to other, higher priority areas, such as technology development. There is a chronic imbalance between the amount of work to be done and the resources available to do it, however, and in addition to backlogs of printed publications such as books and journals, the level of bibliographic control over sound recordings, photographs, films, and archival materials such as manuscripts remains poor. To this growing population of resources, which is increasingly considered important research territory, we now add endlessly proliferating electronic files. User expectations are expanding for timeliness, for in-depth access at the word or image level, for rapid retrieval, and for linkages. Expert management of all this content is essential because, as David Levy astutely observed, " There is a growing awareness of attention as a highly limited resource, stemming in part from the realization that an abundance of information, good though it is in many ways, is also a tax on our attention." [6](#)

To serve their clients well, libraries must blend the features of the catalog with the virtues of the portal. This will require the use of a sophisticated search engine to deliver the quantity of resources, the hyperlinks, the customization and personalization, and the instantaneous access that provides the user with convenience, flexibility, and immediacy. Libraries can add value by promoting filtering and ranking which would prefer resources produced by universities, governments, and other sources that meet a set of established criteria, such as having a strong likelihood of authenticity, accuracy, or endorsement by others of standing. Added to the content retrieved by the search engine should be material contributed by libraries through a reengineered cataloging process. This reformed activity should favor timely access to a wide variety of formats. To achieve this, libraries will have to reallocate funds presently devoted to describing books and journals to materials that are proportionately underrepresented in today' s catalogs, such as films, music, photographs, and digital objects. This cannot be accomplished by fine-tuning, but will rather require significant compromise and change.

Although certain titles may still receive full bibliographic description and analysis such as is conducted today using the *Anglo-American Cataloguing Rules* and the Library of Congress Subject Headings, the need to increase greatly the number of resources of interest to scholars that can be located through the library' s access tool will result in the application of a different standard to many materials. For digital documents, it should be possible to derive key metadata using automation informed by and combined with human intelligence. The emphasis should be placed on identification of many new resources of value to the scholar and researcher, rather than on the cataloging of only a few, relatively speaking, new items.

By working collaboratively, libraries will ensure that they avoid redundancy, and they can aggregate their efforts to create a large-scale portal (" portalog. "?) that offers access to a large quantity of high-quality resources of current interest. One of the salient distinctions of this portal will be that it will bridge the analog and digital worlds in a far more comprehensive way than most Internet search engines do, calling attention to the 96% of the world' s published knowledge that does not exist electronically. In addition, it will draw on two defining characteristics of a library: the ability to provide access to the materials cited or displayed and to offer this access across time through the commitment to preserve and safeguard its collections. A critical aspect of the library portal is that for an individual library to serve its users successfully, it must connect and ally itself with other libraries and developers of commercial search engines in a highly integrated fashion.

Libraries should seek to partner with developers of portals and search engines to share expertise in a constructive way, drawing on the best each has to contribute to the goal of effective access to information. Traditional libraries have much to learn from the commercial portals about attracting and

satisfying users. At the same time, libraries can call attention to the value they have offered and continue to offer for today's and tomorrow's scholars. The result will be a superior service for the world of higher education and beyond.

Endnotes

1. Stanley Wilder, "The Changing Profile of Research Library Professional Staff, " *ARL: A Bimonthly Report on Research Library Issues and Actions from ARL, CNI, and SPARC*, no. 208/209 (February/April 2000): 4. [Also available online.](#)
2. This article is inspired by the author's paper, "The Catalog as Portal to the Internet" (Contributed to the Library of Congress's upcoming "Bicentennial Conference on Bibliographic Control for the New Millennium, " 15-17 November 2000). The [full text of the paper is also available.](#)
3. Lori Leibovich, "Choosing Quick Hits Over the Card Catalog, " *New York Times*, 10 August 2000, G1.
4. Michael Looney and Peter Lyman, "Portals in Higher Education: What Are They and What Is Their Potential, " *EDUCAUSE Review* (July/August 2000): 30.
5. Jerry D. Campbell, "[The Case for Creating a Scholars Portal to the Web: A White Paper.](#) " (Paper delivered at the 136th Membership Meeting of the Association of Research Libraries, Baltimore, Md., 18 May 2000).
6. David Levy, "I Read the News Today Oh Boy: Reading and Attention in Digital Libraries, " in *ACM Digital Libraries '97: Proceedings of the Second ACM International Conference on Digital Libraries, Philadelphia, PA, July 23-26, 1997*, ed. Robert B. Allen and Edie Rasmussen (New York: Association for Computing Machinery, 1997), 202.

--Copyright (c) 2000 Sarah E. Thomas

[Table of Contents for Issue 212](#) | [Other Current Issues](#) | [Other Access and Delivery Services Articles](#) | [Other Scholarly Communication Articles](#) | [Other Networked Information Articles](#)

ARL policy is to grant blanket permission to reprint any article in the Bimonthly Report for educational use as long as full attribution is made. Exceptions to this policy may be noted for certain articles. This is in addition to the rights provided under sections 107 and 108 of the Copyright Act. For commercial use, a reprint request should be sent to ARL Director of Information Services, [Julia Blixrud](#).

© The Association of Research Libraries

Maintained by: [ARL Web Administrator](#) Site Design Consultant: [Chris Webster](#) **Last Modified:** 17 October 2000