Barriers and Obstacles to Use, Satisfaction, and Success:  
The Response of the Million Book Project

Denise Troll Covey  
Associate Dean, Carnegie Mellon University Libraries

User Problems and Academic Libraries

Human factors research conducted over the past few years clearly documents changes in the behavior of academic library users. The body of research available confirms that the Internet has precipitated changes in user needs, expectations, preferences and priorities. The studies detail what users want, their perception of library service, and what they are doing when the library does not meet their needs. While user behaviors might vary somewhat across institutions, libraries can learn a great deal from reading the literature and conducting user studies locally to clarify and prioritize problems and opportunities that distinguish their users and their environment.

This paper summarizes key findings in published library user studies conducted 2001-2003, augments these findings with research conducted by Carnegie Mellon University Libraries, and describes an international response to significant barriers and obstacles to library use, satisfaction, and success. This response, called the Million Book Project, endeavors to close some of the gaps identified in the research – gaps between what users want and what libraries provide. The Million Book Project aims to digitize one million books by 2007 and offer them free-to-read on the surface web. In addition to meeting user needs, expectations, preferences and priorities for easy, convenient, remote access to full text electronic resources, the Project will help address disparities in library collection sizes around the world, contribute to the democratization of knowledge, and provide a substantial test bed for digital library research.

Summary of Key Literature (2001-2003)

What Academic Users Want

Academic users want to be self-sufficient. To satisfy this desire, libraries must provide easy, convenient access to information. Roughly 90% of the students and faculty who participated in a survey of the scholarly information landscape sponsored by the Council on Library and Information Resources (CLIR) and conducted by Outsell Inc. indicated that ease of access is their second most important information need, topped only by the need for quality resources.[1]

Academic users associate easy, convenient access to information with remote access. According to a study conducted under the auspices of the Pew Internet and American Life Project [2], 69% of undergraduates live off campus and 59% of them use their home computer more than computers at school. A survey of college students conducted by the Online Computer Library Center (OCLC) reported that 90% of students access the web from their home computer. Four out of five students indicated that they sometimes use the library for web access, but only one in five prefers this access point; 80% prefer remote access. [3] Over half (54%) of the students and faculty participating in the Outsell study said they access library resources from their residence;
the percentage is higher (68%) for undergraduate students. Faculty said that they spend 10% of their information-seeking time in the physical library, graduate students 30%, and undergraduates 34%. Almost a third (31%) of the respondents indicated that having to go to the library was a problem.[4]

What Academic Libraries Are Providing

The results of the Outsell study indicate that 50% to 90% of faculty and students perceive a significant gap between their high priority needs – for example, for easy access, speedy delivery, and system functionality – and the service their library is providing. Fewer than half of the respondents indicated that libraries are adequately meeting their need for easy access.

Respondents rated ease of access to information as their third most significant problem, tied with lacking sufficient training to use the tools and resources available. Having enough time to do their work was their number one problem, followed by knowing what resources are available. Approximately 24% of the respondents indicated that they often cannot get information when they need it.[5]

A recent OCLC Task Force report entitled “Libraries and the Enhancement of E-Learning” highlights the need for better, more focused training for faculty, students, and librarians. The report strongly encourages the integration of library resources and services with course management systems, and urges libraries to embed training and library instruction within these systems. OCLC introduces a new paradigm shift for libraries: not just-in-case, not just-in-time, but “just-enough-just-in-time-just-for-me.”[6]

The OCLC survey of college students revealed that they encounter problems navigating and searching the library web site and online resources, and perceive these difficulties as barriers to library use. Students in the study perceive vendor licensing restrictions and password requirements as barriers to easy remote access to library resources. Their number one recommendation for libraries is to make it easier to access and use library resources.[7]

Problems with proxy servers and virtual private networks do doubt also present barriers to remote use of online library resources.[8]

Results from the Pew study [9] and a study conducted as part of the Evaluation of the Distributed National Electronic Resource (EDNER) Project [10] indicate that students believe it is easier to find resources using the Internet than it is using the library. The findings suggest problems with library web sites. Research conducted by Jakob Nielsen confirms that users encounter significant problems trying to perform tasks using web sites: they fail to accomplish their tasks an alarming 35% of the time.[11]

The Association of Research Libraries’ LibQUAL+ survey results from spring 2002 and 2003 confirm that, from the users’ perspective, libraries are not providing quality service. In all four dimensions assessed by LibQUAL+ – affect of service, library as place, access to information, and personal control – there is a gap between what users want from their library and what they perceive they are getting. The largest gaps are in the areas of access to information and personal control, which are also the areas that users perceive to be the most important. Speed and convenience are key characteristics in both areas. Users want convenient business hours and
timely document delivery and interlibrary loan. They want easy, convenient, remote access and easy-to-use tools and web sites that enable them to find the information they need without assistance. Though users are generally satisfied with the treatment they receive from librarians and staff, they are much less satisfied with the support they receive. [12]

What Users are Doing

The Outsell study revealed that though students and faculty trust the library more than the Internet, they turn to a popular Internet search engine like Google to satisfy their daily information needs. Almost half (48%) of the students and faculty participating in the study indicated that they start their search for information using an Internet search engine. Only a third start with the library web site. The figures are significantly different for undergraduates: almost twice as many undergraduates (54%) start with an Internet search engine as start with the library web site (28%). Though most (86%) of the Outsell survey participants indicated that their campus library meets most of their information needs, almost as many (80%) said that the Internet has changed their use of the library. Over a third (35%) said they use the library less now than they did two years ago. Most (87%) indicated that they find more relevant information on the Internet now than they did two years ago.[13]

The OCLC survey [14], the EDNER Project study [15], and the Pew study [16] confirm the Outsell finding. Students turn to an Internet search engine first when they need information. They use the Internet more than the library – probably because Internet search engines are more convenient and easy to use. According to the Pew study, 73% of graduate and undergraduate students use the Internet more than the library. Roughly 80% of the college students who participated in the OCLC survey said they use the library fewer than three hours per week. Approximately 40% of the students indicated that they use an Internet search engine for every class assignment, while only 11% said they use the library web site for every class assignment. Almost all (96%) of the students in the OCLC study believe that the information they find on the open Internet is good enough to use in their coursework. Almost half of them (46%) believe that other sites have better information than the library web site.[17] The Outsell survey revealed that 11% of undergraduates are not concerned with the authoritativeness of the information they find, and 5% do not verify the information with other sources.[18] According to the EDNER study, efficiency, or the amount of time and effort required to find information, appears to matter more to students than the relevance of the information found.[19] A research report published by the Library and Information Commission suggests that user satisfaction is multi-dimensional, occurring within a framework of expectations comprised of the information-seeking task, the functionality of the retrieval system used, the assistance provided, the user’s own abilities and immediate goals, and (finally) the information retrieved. Efficiency and the user’s experience of interacting with the retrieval system – in short, the ease and speed of finding information – can be equally as important in satisfying the user as the utility or appropriateness of the information found.[20]

Almost all (94%) of the student and faculty participants in the Outsell survey indicated that they are comfortable with electronic resources. The results suggests that they prefer electronic resources: users indicated unmet needs for electronic resources more than twice as often as
unmet needs for print materials. Students and faculty continue to rely heavily on print resources for research, teaching, and learning. Fewer than half of the faculty and graduate students indicated that they can find most of the information they need for teaching (30%) and research (45%) online. Most graduate students (71%) also rely heavily on print resources to do their coursework. In contrast, almost half (48%) of participating undergraduate students indicated that they use online resources, not print, “all” or “most” of the time to do their assignments. No more than two-thirds of the students and faculty participating in the Outsell study indicated that they get at least half of the information they need for research, teaching, and learning from their campus’s physical or digital library. [21]

Only 56% of the international faculty participating in a study of factors in student success sponsored by McGraw-Hill Ryerson in 2002 indicated that the library was a valuable resource for student success. They positioned the library on a par with tutoring. In contrast, most of the faculty (83%) rated new technology among the three key factors in student success, just behind course preparation and faculty training and development.[22] A somewhat parallel pilot study of student perception of learning success, conducted by Dieter Schönwetter, indicates that students also believe technology is critical. When asked what the library could do to enhance their success, 46% of the students responded provide more technology, 30% said provide more assistance, but only 24% requested more materials.[23]


Research conducted by Carnegie Mellon University Libraries reveals that our users are similar to other academic users as profiled in the studies summarized above. Our users want to be self-sufficient. Easy, convenient access to information is a high priority, topped only by the need for quality resources. Using the physical library is inconvenient. Using the library web site is problematic. Carnegie Mellon users prefer electronic resources and remote access. They want the Libraries to provide more books and journals, preferably in electronic format, and easier-to-use online tools. They also want a customizable, personalizable interface to library resources and services.

A survey conducted in March 2001 revealed that over 78% of the participating students and faculty used the library web site at least once a week – despite problems they encounter finding information on the site, navigating the site, and using the online library catalog, databases and e-journals. The survey results suggest that the information that undergraduates find on the library web site is somewhat less useful to them in their work than the information found by faculty.[24] Redesigning the library web site alleviated some of the problems, but significant difficulties remain. User protocols conducted in February 2002 revealed that the redesigned web site was somewhat less cluttered and more user friendly than its predecessor. However, users still reported problems using the online library catalog and databases. They had difficulty finding links to other library catalogs and locating digital reference services. They could find electronic resources specified by name, but they could not easily find resources specified by subject. Following the user protocols, the web site design was modified again to improve the visual distinction of interactive objects or text and to foreground important links. Unfortunately, improving the design and functionality of the online catalog and licensed electronic resources are beyond the control of the University Libraries.[25]
Results of Carnegie Mellon’s LibQual+ survey, conducted April 2002, revealed significant gaps between what the University Libraries are providing and what users want. The overall service adequacy gap – the difference between user perceptions of minimally acceptable service and the service the University Libraries is providing – was more significant for graduate students than undergraduates and faculty. The overall service superiority gap – the difference between what users want (desire) and their perception of the service the Libraries provide – was more significant for undergraduate students. As with other institutions that did the LibQual+ survey, the most significant problems were in the areas of personal control and access to information. Carnegie Mellon users indicated that personal control was their highest priority and the area most in need of improvement.[26]

In response to the results of the LibQual+ survey, the University Libraries conducted focus groups with students and faculty to gain insight into their issues and experiences. Students and faculty indicated that they use the library both remotely and onsite. They prefer and rely heavily on remote access to full text electronic resources. Focus group participants reported problems and frustration with the proxy server, the design of the library web site, and the design and functionality of the online library catalog and databases. They expressed the need for assistance in determining which resources are appropriate for their purposes, and a convenient way to discover and keep current with available library resources, services, and equipment. They expressed a strong desire for more electronic resources, including back issues of journals, and a desire for a customizable, personalizable interface to the library.[27]

The University Libraries took several steps to address these concerns. The proxy server was replaced with an IP address extension service (virtual private network). An Automated Resource Finder was developed to help users locate appropriate online resources. Online and printed inventories of library equipment were prepared. Online and printed guides to library resources and services are being developed. Librarians are being encouraged to develop online instructional modules. User requests to be able to customize or personalize the library web site are being addressed through the development of a library portal as part of the campus portal. The pace of development is controlled by Computing Services’ implementation of features and functionality in the portal software. Planning is underway to develop library portlets that will facilitate access to electronic reserves and user services associated with the library catalog, help users identify appropriate electronic resources, access digital reference services, and submit interlibrary loan requests. By third quarter 2004, the University Libraries should be able to push discipline specific information to the appropriate students and faculty.

Most recently, the University Libraries conducted a brief survey to better understand how graduate students find and obtain the information they need. The survey was precipitated by concerns about the impact that escalating journal prices, serials cancellations, and changes in access privileges to resources at neighboring University of Pittsburgh are having on graduate students, and the LibQual+ results that indicated graduate student perception of library service quality as barely meeting their minimal needs. Undergraduate students are served primarily by the Libraries’ book collections and the materials faculty put on academic reserves or in BlackBoard. Faculty have salaries that enable them to purchase journals to which the Libraries do not subscribe or to use fee-based document delivery services, and they have colleagues who
can facilitate their access to information. Faculty often have lengthy research projects, so the two to three weeks required for a typical interlibrary loan transaction need not hamper their work. The graduate student experience is different from the undergraduate student and faculty experience. Like faculty, graduate students need access to journals, but unlike faculty, they don’t have the financial resources or personal networks to easily acquire resources that the Libraries do not provide. Like undergraduate students, graduate students often have short deadlines, so the turn-around time required for interlibrary loan can seriously impede their work. The survey of Carnegie Mellon graduate students revealed the following behaviors and preferences:

- When they need information, 82% of graduate students use an Internet search engine “most of the time.” Only 25% indicated that they use the library web site most of the time. Few (9%) indicated that they use the physical library facility most of the time.
- More than twice as many graduate students indicated that using an Internet search engine was easy and convenient as indicated that using the library web site was easy and convenient. Few students indicated that using the physical library facility was easy and convenient.
- Almost all (93%) of the graduate students identified the quality of information retrieved as the most important factor in determining what search tool they use. The next most significant factors were convenience (67%), speed (46%), and ease of use (41%).
- Most graduate students use an Internet search engine first. If that doesn’t work they use the library web site. If that doesn’t work, they turn to a human resource – a professor, classmate, or librarian. If that doesn’t work, they go to the physical library.

The survey revealed the following about how graduate students obtain information after having identified what they want or need:

- Almost three times as many graduate students obtain information “most of the time” from the open web (70%) as obtain information “most of the time” from full text electronic resources provided by the Libraries (27%).
- They rated obtaining information from the web as at least twice as easy and convenient as obtaining information from online library resources.
- About half of the graduate students indicated that they sometimes obtain information using interlibrary loan. About half of those who use ILL indicated that it was inconvenient and difficult to obtain materials this way.

Graduate students indicated that most of the time they can get the materials they need for teaching and coursework, but their comments indicated frustration with the amount of time it takes because the library web site and online resources are not easy to use. Their comments reiterate findings from earlier user studies conducted by Carnegie Mellon University Libraries, and in addition, express concerns about collection size, the availability of books and journals, and the difficulty of acquiring old journals and out-of-print books. Graduate students reported that the limitations of the University Libraries’ collection and the turn-around time required for ILL can impact their selection of a research topic, the quality of their work, and their grade point average.[28]

**The Response of the Million Book Project**
The Million Book Project is an international response to the key findings of the user studies described above:

- Students and faculty turn to an Internet search engine first when they need information because Internet search engines are easier to use than library web sites, databases, and catalogs. Internet search engines meet their high priority needs for speed, convenience, and ease of use.
- Students and faculty prefer remote access to full text electronic resources. They need online access to backfiles of journals and out-of-print books.
- Given that undergraduate students follow the path of least cognitive resistance, lack of quality resources on the surface web is having a negative impact on the quality of student learning.
- Given the time constraints and competitive environment in which they work, lack of speedy access to quality resources is having a negative impact on the timeliness and success of faculty and graduate student research.

The Project also looks beyond the boundaries of these problems to the tremendous disparity that exists across the nation and around the world in the size and accessibility of library collections. Some single institutions, like Harvard and Yale, have more books in their libraries than some entire states have in all of their libraries combined. In our rapidly changing world, lifelong learning and access to books have become essential to employment, health, peace, and prosperity. The democratization of knowledge and empowerment of a global citizenry require equitable access to information. Meanwhile, old books and journals are abandoned by their copyright holders and disintegrate on library shelves. From the perspective of stewarding our cultural and intellectual heritage, achieving the mission of higher education, and engendering democracy, something is seriously awry with the current situation.

Carnegie Mellon School of Computer Science and the University Libraries are leading an initiative to digitize a million books by 2007 and offer them free-to-read on the surface web. The scanning is being done in India and China, with labor funded by the respective governments, following preservation-quality standards developed by the Institute for Museum and Library Services and endorsed by the Digital Library Federation. The National Science Foundation (NSF) funded the initial collection development meeting in 2001, and in 2002 provided $3.6 million to purchase equipment and support administrative travel. With NSF’s approval, some of the travel dollars were allocated to fund the pilot shipment of books to India for scanning. University libraries in the United States have joined the Million Book Project, as have the Online Computer Library Center (OCLC) and the Internet Archives.

Scanning is well underway, with a target of 100,000 books to be available by 2004. Scanning follows established standards for bibliographic metadata and file formats to ensure interoperability with existing systems and migration to new technologies in the future. Plans are to register the books in the Million Book Collection in OCLC’s digital registry to avoid redundant scanning of the same books.

The Million Book Collection will be a multi-lingual collection of collections, including indigenous Indian and Chinese cultural heritage materials, government documents, out of
copyright works, and copyrighted materials. Copyright law is being strictly adhered to. Full time staff have been hired to seek permission from copyright holders to offer their works free-to-read on the surface web. The goal is to have 500,000 in-copyright, out-of-print books and journals in the Collection.

The Million Book Collection will be replicated on servers around the world. The Internet Archive and OCLC have agreed to host the collection. The books will be indexed by popular Internet search engines and freely available on the surface web – accessible anywhere, anytime, to anyone with an Internet connection. Any school, public, or academic library will be able to link their library catalog records to books in the Million Book Collection. The Collection will support education, research, and lifelong learning worldwide. Additional information about the Million Book Project can be found at http://www.library.cmu.edu/Libraries/MBP_FAQ.html.

Notes


[9] Jones and Madden, ibid.


Marcum and George, ibid.


Jones and Madden, ibid.


Marcum and George, ibid.


Marcum and George, ibid.


[28] Analysis of raw data. Report not yet available. Contact Denise Troll Covey (troll@andrew.cmu.edu) or Carole A. George (cgeorge@andrew.cmu.edu).