The People in Digital Libraries: Multifaceted Approaches to Assessing Needs and Impact
Gary Marchionini, Catherine Plaisant, and Anita Komlodi

Digital libraries (DL) serve communities of people and are created and maintained by and for people. People and their information needs are central to all libraries, digital or otherwise. *All efforts to design, implement, and evaluate digital libraries must be rooted in the information needs, characteristics, and contexts of the people who will or may use those libraries.* Like most principles, the devil is in the details—implementing and applying the principle to practical problems. Human-centered digital library design is particularly challenging because human information behavior is complex and highly context dependent, and the digital library concept and technologies are rapidly evolving. Two important aspects of human-centered design are assessing human information needs and the tasks that arise from those needs and evaluating how the digital library affects subsequent human information behaviors. Given the evolving nature of digital library development, solutions to these challenges must be process-oriented and iterative rather than product-oriented and summative. Given the complexity of human information needs and the uncertainty about the effects of new systems, a multiplicity of data views is essential to guide design and understand the impact of digital libraries. This chapter focuses on two elements of design--assessing information needs and ongoing evaluation of impact. Multifaceted approaches to needs assessment and evaluation of digital libraries are illustrated with three case studies, with particular emphasis on a users needs assessment conducted as part of a project to develop prototype interface designs for the Library of Congress National Digital Library Program. The human-centered design principle above links three clusters of constructs: people and their needs, characteristics, and contexts; design, implementation, and evaluation; and digital libraries.

Human information needs, characteristics, and contexts.
People is a term taken here to include the entire range of individuals, groups, and communities who have a stake in a digital library. The information needs of individuals have long been studied by researchers in marketing, education, and information science. There is a long history of studies of what types of information needs people bring to libraries (e.g., Krikelas, 1983; Marchant, 1991; Paisley, 1980; and Wilson, 1981). Dervin & Nilan's (1986) review of the information needs literature dichotomizes system-oriented and user-oriented approaches to determining information needs. They criticize the system-oriented approach as too narrow to actually identify user needs and propose an approach that attempts to directly assess people's information needs. Needs assessment research in information science recognizes that there are different levels of needs (e.g., Taylor, 1962, specified visceral, conscious, formalized, and compromised levels), that users may not be able to articulate their true needs (e.g., Belkin, 1980, noted that users often bring anomalous states of knowledge to a search task), and that needs change as information seeking progresses. Highly personalized needs must eventually be translated into executable tasks. Practical design aims to support a common set of these tasks. A fundamental goal of needs assessment is to identify large numbers of specific needs and map these unique needs into common classes of needs that may be met with standardized task procedures. A related goal is to develop systems that assist/guide people in mapping their personal needs into the system-supported tasks. One of the reasons for the popularity of hypertext selection and browsing mechanisms available in the World Wide Web (WWW) is that people are able to personalize these mappings experimentally albeit laboriously.

That individuals vary on a host of physical, mental, and emotional characteristics is a defining condition of humanity. Much of psychology is devoted to identifying the essential dimensions of human behavior (e.g., theories of multiple intelligence presented by Gardner, 1983; or Sternberg, 1985). A much-promoted but seldom realized aim of design in a democratic information society is universal access. It is axiomatic that designing for universal access is much more difficult than designing for specific populations because the entire range of human characteristics must be supported. Thus, assessing needs and designing a national digital library service requires examination of many communities and will likely lead to multiple system solutions.

In addition to the needs of individuals and groups who make use of information in DLs, the needs of the providers and managers also influence design and evaluation. There are many groups and individuals (e.g., digital librarians, taxpayers, political leaders, philanthropists, etc.) who have needs that must also be taken into consideration in digital library design. Individuals are embedded in many different communities and communities are embedded in larger social and cultural contexts. When it comes to human behavior, these contexts are inescapable and confound efforts to artificially isolate specific variables for assessment.

The variety of stakeholders and contexts exacerbates the inherent complexity of assessing human information needs and the impact of systems designed to meet them. Designers and evaluators who wish to take a human-centered approach are thus challenged to specify what people will be served, what levels and types of information needs will be supported, and what contextual influences will be at play. These challenges are summarized in Table 1 that expresses questions from the perspectives of designers and evaluators. Clearly, this is an overstated dichotomy-- designers are concerned with evaluation questions and evaluators must consider the design questions. In practice, design team members collaborate to address these questions and in some cases, individuals serve in both roles.

Table 1. Human-Centered Design and Evaluation Questions

Designers	Evaluators
Who are the users?	Who is impacted?
Who are the potential users?	Who and what may influence impact?
What are the common needs?	What are the indicators of impact?
How can those needs be mapped onto tasks?	How can indicators be measured?
How will the new system change needs (and tasks)?	How do impacts influence future generations and systems?

Design, implementation and evaluation.
Design, implementation, and evaluation are processes that marshal intellectual and physical capital to yield tangible, usable, and testable products. The design process is of primary concern to architects, engineers, and inventors and there is a significant body of literature devoted to theory (e.g., Braha and Maimon, 1997; Simon, 1996), history (e.g., Petroski, 1996), and practice (e.g., Brooks, 1975; Norman, 1988). Computer system designers have begun to consider the human factors (physical and psychological) associated with system usage and there is a growing recognition of the importance of user-centered design for systems used by general populations. This approach is increasingly adopted by practicing software designers. Hix &Hartson (1993) offer many practical suggestions for assessing the tasks that users bring to computer systems and Shneiderman (1998) provides a rationale for mapping user tasks and needs onto the syntax and semantics of interface designs. Marchionini (1995) provides a framework for mapping users, tasks, and information need settings onto interface designs.

Implementation issues follow design and much of the work libraries do to digitize collections, provide access, and insure interoperation advances practice by demonstrating "how to" procedures. In the ideal, the design specifications are perfect, the work goes smoothly, and the project comes in on time, on cost, and includes all and only the functionality defined in the specifications. In practice, systems contain a variety of workarounds, add new features, and do not include all functions in the specifications. In spite of efforts to build a science of design, iterative design informed by evaluation feedback is more typical.

Evaluation may be a research genre aimed at assessing classes of techniques or methods (e.g., Suchman, 1967 for social programs; Flagg, 1990 for educational technology), or a systematic assessment of a specific product for the purposes of improvement (e.g., Nielsen, 1993). Summative product testing is another form of evaluation that is not applicable to complex and evolving concepts like digital libraries. A human-centered approach to design, implementation and evaluation is fundamentally complicated by the variability in human characteristics and behavior. See Harter & Hert (1997) for a recent review of evaluation research in information retrieval.

Evaluation of a digital library may serve many purposes ranging from understanding basic phenomena (e.g., human information-seeking behavior) to assessing the effectiveness of a specific design to insuring sufficient return on investment. Human-centered evaluation serves many stakeholders ranging from specific users and librarians to various groups to society in general. Additionally, evaluation may target different goals ranging from increased learning and improved research to improved dissemination to bottom line profits. Each of the evaluation goals may also have a set of measures and data collection methods. Finally, the evaluation must have a temporal component that can range from very short terms to generations. One approach to this evaluation complexity is presented in the Perseus DL case below.

Digital Libraries: Integrating People, Information, and Systems.
Digital libraries are the logical extensions and augmentations of physical libraries in the electronic information society. Extensions amplify existing resources and services and augmentations enable new kinds of human problem

solving and expression. As such, digital libraries offer new levels of access to broader audiences of users and new opportunities for the information science field to advance both theory and practice (Marchionini, 1998). There is a substantial body of literature related to digital libraries including many conference proceedings (e.g., Association of Computing Machinery annual DL series) and special issues of journals (e.g., Journal of the American Society for Information Science 1993, IEEE Computer 1995, Communications of the ACM 1995 and 1998, Information Processing & Management, 1999). Marchionini & Fox (in press) introduce one special issue by framing digital library design space with community, technology, service, and content dimensions and argue that most research and development to date have been devoted to the technology and content dimensions. Thus, the bulk of the work in DLs has been devoted to extending access beyond the physical walls of libraries and to extending citizen access to government-produced information. As extensions, we should be able to access more relevant information faster and with less expense. The augmentation of community and information services remain important challenges for the years ahead.

As DLs are actually developed, used, and improved, design guidelines will slowly evolve through experience and reflection. Needs assessment and evaluation offer several special challenges. An inherent limitation in directly assessing the human needs for an innovation is the fact that potential users must imagine what the innovation can and will do for them. This is very difficult to do and innovators often justify adopting a "build it and they will come" (BITWC) policy based on their own imaginations of needs and applications. If the engineering is good and the marketing successful, people will recognize the system's value and adopt it. Information technology history is filled with cases of top-down BITWC design success and failure. Alternatively, designers can study users continually and involve them at all stages of the design and evaluation process, thus insuring a ready-made market. This systematic, bottom-up approach can produce lowest-common denominator solutions, and in the worse cases may exhaust time and resources before any solution can be built. Clearly, some middle ground is needed for DL design. High-level visions guided by astute observations of human behavior coupled with systematic and iterative assessments seems the right approach.

Because digital libraries are extensions and augmentations of physical libraries, needs assessment and evaluation may be initially modeled generally on physical libraries. Starting with the general goals, stakeholders, methods, and outcomes of physical libraries and related information technology services, designers and evaluators must be alert to new applications and goals, new user communities that may emerge, changing needs and abilities of existing user communities, new technological developments, changing information processes and capabilities, and new possibilities for data collection and manipulation.

The design and evaluation of DLs is driven by high level visions but is mainly a bottom-up process that synthesizes specific instances and cases based on systematic probes of authentic environments with results from controlled investigations in simulated environments (e.g., laboratories). These approaches are equally expensive and mutually complementary. Authentic environments are context-rich but are therefore complex and not under the designers' control; additionally the environments must exist—there must be a working design to study. Thus, to design new environments, comparable built worlds must be investigated and results applied to new designs. Prototypes and laboratory studies offer good control over specific technical variables but only give glimpses of authentic environments. It seems clear that multifaceted approaches to determining user information needs and evaluating DLs must be used and the results integrated to inform design as an ongoing process. The integration is not algorithmic—it is, however, systematic, interpretive, and driven by the high level goals. This integration is analogous to medical imaging techniques (e.g., CAT scan) that aggregate a plethora of data slices so that diagnosticians may interpret holistic organ status. Of course the data cases in DL design and evaluation are less precise and come from several different sources, making the interpretations and conclusions more time consuming and more dependent on inference. For complex phenomena such as DLs where the complexities of human characteristics, the world's knowledge, and sophisticated IT and social systems intersect, it is clear that principles and guidelines are synthesized over time rather than hypothesized and demonstrated.

Design and evaluation must be customized because every DL is situated in a context defined by community policies, human needs and characteristics, and technical constraints. Aristotle noted in his *Ethics*: " In practical science, so much depends on particular circumstances that only general rules can be given." To identify some of the general rules for DL design and evaluation and to demonstrate how multifaceted data streams can be synthesized, three cases are presented. One case (Perseus) is summarized to illustrate an iterative, longitudinal, and multifaceted approach to evaluation of DLs. The case illustrates how a community develop, how usage patterns evolve with

technology , and what long-term impacts DLs can have on a multiplicity of stakeholders. The second case (Baltimore Learning Community) is described briefly to highlight specific high-level goals that drive design decisions and illustrate that good design in not sufficient to community building. The third case (Library of Congress) is discussed in detail to illustrate the multifaceted approach to user needs assessment when the user community is ill-defined and there are no precedents for the system.

The Perseus Project Evaluation.

The Perseus Project began developing a corpus of multimedia materials and tools related to the ancient Greek world in 1987. The mission of this project was driven by the perceived needs of students and faculty to have improved access to primary source materials and to have juxtaposed linguistic and visual resources to better learn and understand culture. This evolving digital library (www.perseus.tufts.edu) began as a HyperCard-based CD-ROM library of Greek texts and English translations, images of vases, sculpture, and sites, maps and drawings of Greek sites, and a variety of retrieval and philological tools. The DL transitioned to the WWW in 1995. From the first days of the project, an evaluation team worked to address a set of research questions related to learning, teaching, scholarly research in the humanities, and electronic publishing. The evaluation effort has continued for more than a decade and there are many published reports on the project as well as the evaluation (several evaluation reports are able at http://www.perseus.tufts.edu/FIPSE/ see also, Marchionini & Crane, 1994).   The purpose here is to provide an overview of how the evaluation was initially framed and how it evolved over the years as the digital library was developed, used, and expanded.

In the original four-year evaluation plan, four evaluation goal sets were identified: learning, teaching, system (performance, interface, electronic publishing), and content (scope, accuracy). Three characteristics of the computational medium that we believed would add particular values were identified for particular emphasis: access, learner control, and collaboration. Based on these goal sets and media characteristics, a hierarchical set of 94 questions was developed to guide the overall evaluation[1]. Four sets of stakeholders were identified for study: students, instructors, project staff, and classics researchers. Based upon these goals and stakeholders, a set of data collection methods were adopted. These methods fell into three general classes: observations, interviews, and document analysis. Observations included baseline notes made by evaluators during classes or in laboratories, structured checklists and forms completed in these same settings, audiorecordings of people thinking aloud while using the system, and transactions logs of Perseus usage. Interviews included one-on-one and group interviews and discussions that were audiotaped and transcribed, and various written questionnaires completed by students, instructors, and scholars. Document analysis included examination of the system software and documentation; syllabi and assignments created by instructors and student responses to assignments (essays, journals, and hypermedia paths). Observations and interviews were done at a variety of university sites that included major research universities (both public and private) and small liberal arts schools, and in selected high schools and museums.

Over the first few years of the evaluation the scope of work and resources available caused the evaluation team to decide to invest most efforts in the learning and teaching goals and on students and instructors as stakeholders. This is a practical example of prioritizing evaluation goals. Additionally, additional data collection techniques were added (e.g., written comments of visitors to a museum exhibit at the National Gallery of Art in Washington where Perseus was available as an adjunct to a sculpture exhibit). As the Perseus corpus grew in size and especially as it migrated to the WWW, the evaluation team was able to examine more longitudinal effects. By the late 1990's evidence was gathered to support the notion that Perseus was having systemic effects on the field of classical studies.

---

[1] The number of questions were as follows: learning had four general questions, twenty four subquestions, and two more specific questions, teaching had three general questions and nine subquestions, system had three general questions, thirteen subquestions, and twenty-six more specific questions, and content had three general questions and twelve subquestions. As an example, the first general learning question was: What tactics and strategies do students employ at particular junctures in the Perseus environment? A subquestion under that question was: What proportion of time do students spend in the primary text?

The main findings over the first eight years of evaluation can be summarized in four categories: amplification and augmentation, physical infrastructure, conceptual infrastructure, and systemic change.

Amplification and Augmentation. The Perseus DL amplifies and augments teaching and learning. Amplification takes two forms: a) more texts (including some not available for students in print forms) and more images and maps than most department slide and map libraries offer; and b) an integrated corpus that allows text-oriented courses to easily add image-based context and vice versa. Another type of amplification often noted is that content may be accessed more quickly and easily than physical versions in libraries (a kind of mechanical advantage). Augmentation is evidenced by instructors who introduce new activities that are otherwise impossible. For example, the philological tools allow instructors in class to illustrate points with word analyses or to visually and easily correlate geographic characteristics and textual passages. In addition, entirely new courses were created that integrated the varied and multiple resources in the DL. More importantly, Perseus empowered new kinds of student learning such as sophisticated philological investigations by students who knew no Greek, visual investigations of vase painting themes, and new discoveries by students alone or as part of a class.

Physical Infrastructure. Using the Perseus DL requires substantial physical infrastructure investment. At every site evaluated, hardware and network problems caused frustration for faculty and students, substantial economic and human resources were necessary to make Perseus available in classrooms and laboratories, and laboratory staff had to be trained to support faculty and student access. These challenges are faced by all educational institutions introducing technology into instruction and reflect the larger learning curve investments taking place in all disciplines at the close of the twentieth century. One important effect apparent when Perseus was used through the WWW was the relative ease of use when compared to the HyperCard-based version--students did not have to learn a new system but used the mostly familiar interfaces of web browsers. Thus, the common infrastructure of the WWW, especially the general purpose client-side browsers that faculty and students use for many different courses and purposes, mitigated some of the technical requirements for Perseus use. One result that recurred over the years is that self-reports on the system interface and learning effects are highly correlated, whereas demographics, computer experience, and frequency of Perseus use are not statistically correlated with self-reported learning effects.

Conceptual Infrastructure. The Perseus DL demands new conceptual infrastructures for teaching and learning. Instructors must learn to teach with the DL and students must learn how to learn with it, and both these requirements involve substantial amounts of time. Many instructors noted the large investments in time required to create assignments and Perseus-augmented lectures. Likewise, some students complained about the amounts of time it took to learn to use and access the system and to find the information needed. Several instructors noted that students took longer to complete assignments than anticipated and that classroom use often took longer than planned since interesting alternatives or additional examples could easily be pursued with the system. Instructors should take into consideration that the novelty and amount of work an innovation demands may lower student course evaluation results during the early years of adoption. Several instructors noted that Perseus raised their levels of expectations about the scope of material accessible to students. Likewise, students at schools where Perseus was used in multiple courses came to expect that such resources would always be available for use in their courses.

A number of opportunities and challenges related to teaching emerged. The traditional dilemma of how best to mix open-ended and guided instructional activities is exacerbated by the many possible uses that Perseus offers. Likewise, how to best use class time and what content is displaced when Perseus-based content is introduced into a course or curriculum must be considered? Perseus allowed instructors to model how they do their own research and the risks and time required to model research should be considered. Instructors must learn to evaluate electronic assignments. One instructor noted that more extensive feedback was made possible by having assignments and his comments in electronic form since he could leverage all the advantages of word processing while grading. Certainly, instructors and administrators must understand that iterative planning and implementation cycles are required over years rather than weeks or semesters, and appropriate allowances, resources, and rewards must be available.

Opportunities and challenges related to learning were also varied. Students are certainly motivated by Perseus, especially by the images. Students were observed to work harder and better when their assignments were put on the web--the persistence of the assignment beyond the end of the course and the "publication" of the work are likely explanatory factors in this regard. Some students reported being overloaded by the amount of content available. Likewise, some students were overwhelmed in lectures that included many Perseus examples and multiple verbal themes. The learning curve necessary to use Perseus tended to be more problematic in large, general studies courses

than in advanced courses for classics majors who tended to recognize the time it took to learn to use the tool as an investment to be amortized over multiple courses.

Systemic Change. Perseus is bringing systemic changes to fields of classics. In 1998 there were more than 50 courses included on the Perseus web site representing more than a dozen colleges and two dozen instructors. These courses use Perseus in a variety of ways and illustrate the penetration of Perseus into the classics curriculum internationally. In some universities, multiple instructors use Perseus for several of their courses. Several instructors noted that Perseus use has led students, faculty, and administrators to see classics as technologically "plugged-in" leading to more recognition and resources on campus. While it is too soon to generalize, the new courses created based on Perseus tend to integrate textual and visual materials and illustrate ways to break down barriers between distinct areas such as philology and art history. New classics faculty position announcements list computer experience as requisites. Popular textbooks now include Perseus companion paths, and a Spin-off company: Classical Technology Systems provides training and support. The Perseus project continues to attract funding for expanded work, and workshops and papers related to Perseus are standard fare at professional conferences in the multiple fields of classics as well as education. By 1998, the Perseus DL was responding to approximately 25,000 hits per day (unique requests excluding GIFs) and has become an electronic gateway to a suite of electronic resources well beyond the original ancient Greek culture corpus and tools. Many of the tools and techniques have been integrated into new projects at Tufts University (e.g., Rome, ancient science) as well as in other projects. The Perseus DL impacts a large and diverse community beyond university classics courses. It serves as a stable and authoritative resource for other publications (e.g., several commercial online encyclopedias refer to the Perseus DL) as well as for distance education and other non-traditional education venues.

These results address many of the high level goals related to teaching and learning that were set out in the original evaluation plan more than ten years ago. The evaluation adapted to new technologies and content and to the new ways that instructors and learners found to apply this DL to their needs. It was opportunistic in that it took advantage of new venues (museums and high schools) and new data collection techniques (e.g., the transaction logging scripts built for HyperCard version and the less user-specific but more broad-based transaction logs of the WWW). It took a multifaceted approach to evaluation that was bottom up in that it integrated many specific data collection efforts guided by the high-level general questions that guided it. The longitudinal dimension allows the evaluation team to extend the evaluation space to a temporal dimension. Figure 1 illustrates this expansion by framing a wide range of stakeholders across time (Marchionini, 1995). Such a framework gives a "big picture" flavor to digital library evaluation research and can easily be adapted for other stakeholder sets or evaluation goals. In the figure, time intervals are: immediate (days or weeks), annual (year), short term (two-five years), mid term (five-fifteen years), generational (20-30 years), and very long term (50+ years). The icons in each cell are balance beams that represent tradeoff balances between tangible and intangible costs and added values of the DL. In the figure, all balance beams are equivalent. Specific DL evaluation efforts can aim at determining what cost-benefit tradeoffs are evident for any stakeholder/time subset. In the Perseus example, we might depict the immediate benefits to learners and instructors as more costly than traditional resources but show better payoffs in the short or mid-term rows. Such situations would be visualized by balance beams tipped in opposite directions in different rows and columns.

Baltimore Learning Community

The Baltimore Learning Community (BLC) Project aims to leverage DL resources and technology to facilitate a community of teachers who create and share outcome-oriented instructional modules. Beginning with a seed collection of video material from Discovery Communications Inc., the BLC project team worked with teachers to create a small repository of multimedia materials, and tools for accessing those materials and building and using instructional modules. Teachers are encouraged to share the modules they create, together with reflections on how the modules were used in classes. Additionally, teachers may contribute specific resources to the BLC repository.

Figure 1. Temporal Framework for Evaluating DL Effects on Educational Stakeholder
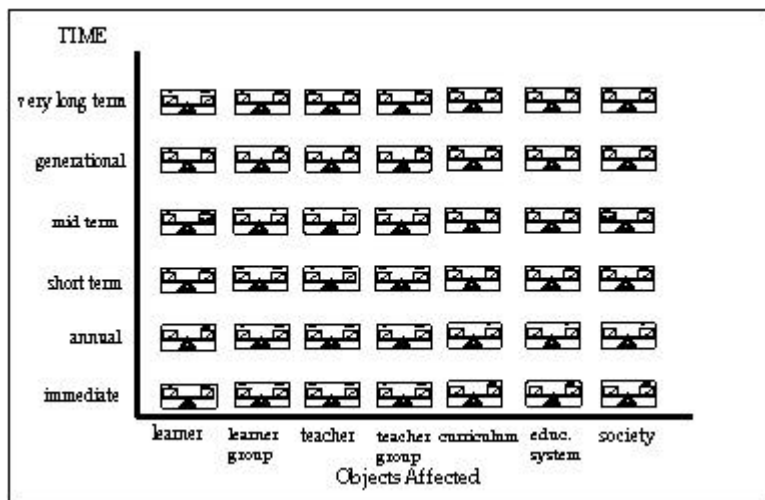
Together with face-to-face meetings, requiring multiple participants in each school, and Internet communication facilities, this sharing of ideas, raw materials and modules were central components for the vision of community building and sustainability. The BLC DL provides support for teachers to:
- Find multimedia resources in the database by a variety of entry points such as topic, instructional standards, media, and source;
- Use those resources and others found in the WWW to create instructional modules keyed to state of Maryland instructional outcomes;
- Present the modules in classes;
- Contribute new resources to the BLC repository, including project-defined metadata;
- Contribute instructional modules to the repository; and
- Communicate with other teachers in the BLC community.

Although the original vision may be characterized as BITWC informed by state and local curriculum guides, the design process involved teachers after funding was secured and this involvement has increased as the project evolved. The basis for design is rooted in assumptions about teaching practice and needs. Teacher needs are actualized in curriculum standards and guidelines--it is unmistakable that teachers in public schools are strongly influenced by local and state curriculum guidelines and assessment procedures. Our intention from the start was to assist teachers in addressing these needs. The mechanism was providing teachers with access to high-quality multimedia materials tied to curriculum rubrics and encouraging them to leverage these materials to build community where teaching practice and lessons are shared to everyone's benefit. This approach to meeting common needs through shared resources and experiences is the basis for the sharium concept--a distributed electronic problem solving space with rich information resources and tools where people can work independently or collaboratively to solve problems (www.ils.unc.edu/~march/sharium/sharium1.1.html).

The BLC evaluation plan aims to document how teaching practice evolves in such an environment and eventually how student learning is affected. See www.learn.umd.edu for background on the BLC, including a demonstration of selected system components and a set of reports on the project. See Enomoto, Nolet, & Marchionini (1998) for preliminary evaluation results.

The system design used a selection and form based approach to lesson construction, alternative interface options for teachers to facilitate personal styles as well as working at home or at school, a dynamic query interface style to facilitate resource exploration, and different video surrogates to save teacher preview time. The system design evolved as teachers expressed their needs more sharply as they gained experience. Based on teacher feedback, several changes to the design have been made:

- A simple slide-show presentation component was added to the original presentation component that included details on pedagogy (e.g., objectives, groupings, instructional strategy, etc.);
- the full Maryland School Performance Assessment Program Outcomes Model and Content Indicators for social studies and science were made available from the module construction tool;
- a feature for printing locally-required lesson plan formats from the modules was added; and
- the resource explorer tools were embedded in the module construction tool rather than remaining separate functions.

In the fourth year of acquisition, indexing, interaction, and use, the system continues to evolve and there are a few indicators that a community is forming. Overall, progress has been slow and arduous.

- a few teachers have contributed original materials (e.g., photographs of the Baltimore harbor were taken, indexed, and contributed);
- more than 100 modules have been created and added to the repository although teacher reflections and commentaries on usage have seldom been included;
- most teachers have made significant strides in incorporating the Internet into their classrooms, in most cases without using the BLC repository directly (although they often add create modules that incorporate WWW materials they find); and
- a fourth middle school and group of more than a dozen teachers have joined the project.

On the research and development side, the advanced interfaces have been incorporated into a web-based application and a series of empirical studies of video surrogates have been conducted (e.g., Ding et al., 1997; Komlodi & Marchionini, 1998; Slaughter et al., 1997). It is clear that the process of community building in schools and across schools is a very slow process. Building the DL has proven to be much easier than building a community of instructional practice. Although some of the difficulties teachers have in creating and sharing materials are due to technical issues (e.g., slow access, system crashes, multiple platform incompatibilities), the challenges of teaching in inner-city middle schools, administrative pressures to raise test scores, and lack of incentives for out-of-class time investments deter forward progress. Few teachers have actually used the modules created by other teachers, there is a flurry of activity immediately after the annual summer institute and less activity through the school year. With the exception of a few teachers who are actively creating and using modules, most teachers ignore email and choose not to attend after-school development sessions (small stipends are provided).

It seems clear that the demands of students, administrators, and the school system climate get much more attention than the project. It is likely this is as it should be--DLs and technology are not panaceas to solve the problems of public education. It will take time for communities to form around the resources and tools the DL research and development community provide. This case illustrates that content and technologies are not sufficient to create sustainable communities in highly structured environments such as public schools. Funding bodies and designers must show good patience in assessing impact in such environments. Designers and evaluators must exhibit reasonable expectations in such settings, recognizing that the people who use the DL are concerned with the immediate value to their work rather than with the design and evaluation.

Library of Congress National Digital Library Program

This case study reports on the Users Needs Assessment conducted for the Library of Congress National Digital Library Program in order to develop prototype user interfaces for the National Digital Library (NDL). The design process and the various prototypes are described in Plaisant, Marchionini, Bruns, Komlodi, and Campbell (1997), and Marchionini, Plaisant & Komlodi (1998). The focus here is on the extensive information needs assessment that was undertaken before designs were begun. Based on the perceived needs of the populace to have better access to the Library of Congress' (LC) treasures, a multifaceted and flexible approach to estimating the needs of the citizenry was taken.

This users needs assessment of the LC NDL was conducted over a relatively short period of time (four months) so the flexibility proved especially important in identifying potential users and tasks using techniques that were practical in the allotted time period. For example online surveys of users of the existing NDL web interface were not used for two reasons: first they only reach self selected users who have already found the NDL and chose to fill the questionnaire, and second they required a long process of authorization, review and implementation which was prohibitive in this case.

As we collected information about user characteristics and their needs, we examined corresponding user interface challenges. A taxonomy of users and tasks was proposed. It is highly likely that this taxonomy will change as the NDL evolves since it is rooted in data that come from a mix of current users of the physical LC, early adopters of the nascent implementation, and non-users asked to speculate about using it. This taxonomy and a general list of interface challenges became important reference materials for the team designing the prototypes for the NDL.

The LC NDL Context and High Level Goals
The Library of Congress has taken a leadership role in the development of a National Digital Library. LC's role in this effort is to digitize 5 million items from 200 Americana collections over a five year period (http://memory.loc.gov/ammem/amabout.html). As part of this initiative, the NDL Program planned activities for preparing, digitizing, archiving, and providing access to these historical and contemporary materials and for managing the overall project and collaboration with other institutional partners. The challenges of providing access are addressed by efforts to develop an easy to use yet powerful human-computer interface. It is our view that a poor user interface is actually worse than denying equitable access since it frustrates people and wastes their time. To guide the development of the interface, a user needs assessment was undertaken to determine what tasks and search strategies users bring to the Library today, what new types of users, tasks, and strategies the NDL will attract, and how these user needs may best be served by the NDL interface. Because the NDL Program was new, there was no extant user community. To assess the needs of potential users a variety of data collection approaches were taken and a flexible approach to integrating and interpreting data was adopted.

Needs Assessment Stakeholders.
Three general user communities were considered at the beginning of the project and others emerged as the needs assessment progressed. The three original groups were:

> LC staff. This community is characterized as experts in both library system and content, as heavy users of the existing systems, as knowledgeable about articulated (formalized) user needs, and able to devote short (usually less than two hours) amounts of time to individual's needs.

> Scholars and other visitors to LC. This community is characterized as having high levels of domain expertise and knowledge of library systems. They are also able to devote large but specific amounts of time (days or weeks) to their individual information needs.

> Visitors/Prospective users. This community is characterized as having varied levels of domain expertise and low levels of library systems knowledge. They are able to devote only short periods of time to their information needs or explorations.

Data Collection Procedure
Three types of data collection were undertaken: selected LC Reading Rooms were visited and staff hosts were interviewed; written questionnaires were created and distributed to important prospective user communities; and LC documents were examined.

Reading Room Visits and Interviews. This was considered to be the most informative component of the assessment since the librarians in the reading rooms are not only themselves important users of the existing system but also have extensive knowledge about current user needs at LC. In addition, many of these staff members are either directly or indirectly involved in the development and maintenance of the NDL and have stakeholder needs that must be captured. Two or more project team members participated in visits to nine different reading rooms. For each visit, information was gathered with regards to the three facets of the user needs assessment and a fourth

category was added for miscellaneous data. These facets were:

> Content: What materials are housed in the reading room, what indexes or finding aids exist, and which materials were scheduled for inclusion or possible inclusion in the NDL.

> Users: The types of users who typically used the reading room, what types of information needs they brought to the Library, what levels of searching skill they typically have, how these users might benefit from the NDL, and what new users might be attracted by the NDL.

> Strategies: How do users conduct searches, what types of search tools are available in the reading room, and how does the reading room staff assist users.

> Other: Any notes on interface implications, additional notes specific to that reading room

The notes were transcribed by one of the interviewers and sent via email to the other team members who participated in the visit. These team members augmented and edited the electronic notes based on their own written notes and the final electronic version was sent for comment (member checking) via email to the reading room staff member who had hosted the visit. After staff had made comments, corrections, or addition, a final set of notes was prepared.

Questionnaires for Potential Users. Because K-12 teachers and students are featured prominently in plans for the NDL, a limited survey of educators was conducted. Rather than an expensive and time-consuming random sample, we took advantage of existing contacts to survey a convenience sample of educators. There were two groups so selected. The first group was a set of 24 teachers from different areas of the U.S. who participated in NDL orientation sessions. The second group of educators was the set of 27 school library media specialist supervisors in the state of Maryland. Both of these groups were administered a questionnaire including questions on types of materials used, search strategies applied, system characteristics preferred and possible future applications of a NDL collection. The difficult task of assessing the needs of the general citizen was addressed by a survey of parents and workers in a large daycare center. This facility serves a large and diverse community of blue-collar and white-collar workers in a Midwestern U.S. city.

Document Analysis. Documents represent an institution's expression of procedures, policies, and responses to user needs. Three types of LC documents were examined as part of the needs assessment to better understand user and system needs:

> Reading Room handouts and brochures. During reading room visits, many handouts were obtained and these handouts reinforced the notes taken during visits in that in some cases they illustrate user needs that are so pervasive that special publications or finding aids are created.

> User study reports prepared by LC staff.   These two studies reported on the experiences of different user groups in different environments with the pilot American Memory project, a precursor to the NDL using similar materials. These reports identified many interface issues.

> User email commentaries or inquiries about the NDL sent by early users of the infant NDL, as well as staff responses, were analyzed and categorized.

Results of Reading Room Visits and Interviews

Results are organized by content and by users and strategies.

Content. The NDL is not an academic exercise that begins with well-articulated user needs and then adds content that is tailored and organized to meet those needs. Instead it is a large, real-world effort that is rooted in a complex cultural context. Content for the NDL is selected based on a variety of economic, legal, social, and political exigencies. As the digitization plan continues to evolve, we can anticipate that materials added to the NDL will be as varied as the holdings of the Library itself. This implies that interfaces to the NDL must be driven by content decisions as well as users' needs.

Challenges are posed by the variety of content digitized from the collections of the different reading rooms. For each visit we used the notes to compile a list of content-related interface challenges to be addressed in the prototypes. Table 2 gives samples extracted from notes and lists of challenges:

Table 2. Samples from reading room notes related to the "Content" facet of the assessment, and corresponding interface challenges

Examples of visits	Extracts from notes	InterfaceChallenge Examples
American Folklife Center	The collection is largely uncataloged, much of it is unpublished. Originally consisting of folk songs, it now includes cultural documentation in various media. Most collections are multi-format. Most include recorded sound.	Browsing of materials will be more important than catalog search. Accessing audio (searching capability as well as download times); showing between to multi-format items (i.e., coordinating a sound track and textual field notes).
Map reading room	Items include panoramic maps (cataloged), county atlases (cataloged), railroad maps, East European maps (not cataloged, heavily used by genealogists), Civil War maps, Sanborn fire insurance maps (about 700,000 items), and selected maps of American history	Very large maps to be seen at varying resolutions. Specification of areas/regions on maps; Problem of place name ambiguity
Prints and Photographs Room	Items for NDL include: Washington, Lincoln, Jefferson, papers; WPA life histories; Whitman notebooks; Margaret Mead collection; some finding aids for other collections. Finding aids very important, but vary in detail.	Full-text search. Specification of limits of search within a manuscript, across a collection, etc.; Important to distinguish finding aids levels and primary materials.

Next all challenges were summarized into a set of general content-related interface challenges. First, there are a variety of materials that will become part of the NDL. Not only do the materials vary by topic, size, and format, but also by degree of cataloging. Some materials are cataloged at the item level, some are cataloged only at the collection level, and some are uncataloged. This state of affairs presents huge challenges to LC staff as they work to serve user needs and challenges to users who are trying to find information in the NDL. It is highly unlikely that new cataloging efforts can be undertaken to catalog those items not currently cataloged (e.g., in the Geography and Map Reading Room alone, only about one-quarter million items are cataloged out of 4 million.). It is also unlikely that catalog records for specific items in collections cataloged at the aggregate level can be created, or that existing catalog records can be edited to reflect any special requirements of the NDL (e.g., geo-referencing for maps). Thus, the first challenge is to develop a conceptual interface design--an organizational framework with appropriate rules for applying the framework. This framework must characterize for users the granularity, size, and nature of objects in the NDL. This challenge cuts across all the reading rooms and has several facets. The interface must communicate to the user the following:

- What is contained in the entire NDL,
- The level of representation for a displayed object (bibliographic record to collection, series, or item).
- The alternative levels of representation available for a displayed object (bibliographic record only, thumbnail or other extract, primary object),
- The nature of a displayed object (secondary or primary; format(s); concomitant or linked objects)

This challenge must be met within the current levels of indexing. However, just as the Library takes advantage of user browsing behavior in uncataloged collections or in collections that are described only at a group level to create

records subsequently for individual items selected by patrons and therefore presumed to be "high demand" (e.g., photographs found and copied by patrons in Prints and Photographs), the NDL presents an opportunity to add additional cataloging information to items that users find while browsing the digital collection. Another element of this challenge is to integrate searching of catalog records with searching of SGML finding aids available for some objects such as manuscripts.

A second challenge is to communicate to the user what items are NOT in the NDL. At one level this places the NDL within the context of the entire LC itself (e.g., copyrighted materials, 3-D objects, etc.), and in the context of the world of information available in other institutions and on the entire Internet. On another level this assists users in planning for visits or becoming aware of related information at LC or other institutions.

A third challenge is to support users in the NDL without human intervention. Although some level of reference service will be necessary in the NDL, every patron cannot expect the level of human support current visitors to LC need, receive, and expect.

A fourth challenge is to create an interface that is accessible to user with state-of-the-market technology. The interface (and underlying retrieval system) cannot be based on assumptions about state-of-the-art hardware (e.g., in 1995 terms: 20 inch high-resolution displays, very high-speed connections, huge RAMs, specialized I/O devices) or software (e.g., the latest operating system or web browser)[2]. Just as the overall system must leverage compression/decompression and advanced retrieval algorithms, the interface must inform the user about temporal demands for data transfer, provide posting information about result set sizes, and provide some level of explanation for ranked results. In addition, the interface must be "growable" to keep pace with the evolution of hardware and software.

A fifth challenge is to invent new techniques to search for multimedia objects and to integrate those techniques into the interface (e.g., visual and audio query languages, surrogate viewers).

Together with the specific interface challenges listed in the reading room summaries above, the NDL content begins to define both a development and a research agenda for more general digital library interface design. Although these challenges were addressed at preliminary levels in the resulting prototypes, they all will remain long-term problems for the DL research and development field for many years to come (see http://www.cs.umd.edu/projects/hcil/ndl/index.html for reports and various prototypes).

Users and Strategies Facet. The Library of Congress is mandated to first serve Congress and its staff. It has traditionally been a library of last resort for other citizens in that its collections are so large and specialized that only well-prepared researchers can take full advantage of the Library. Although many casual visitors come to LC to see and experience it as an institution, it is a research library and operates to serve those prepared to work in a complex and scholarly environment (e.g., citizens under the age of 16 are not permitted to use the Library). The NDL is a fundamental change in the service mission of LC in that a much broader user community is addressed. Table 3 below gives samples from notes and interface challenges.

Table 2. Samples from reading room notes related to the "Users and strategies" facets of the assessment, and sample interface challenges

Examples of visits	Extracts from notes	InterfaceChallenge Examples
Geography and Map Reading Room	One-third of users (estimate) are looking for genealogical information. Estimate one-forth of users are contractors looking for environmental information. Specialized maps are often sought	Specifying areas; awareness of the potentially huge files to transfer and display.

---

[2] A more parsimonious position is to assume a state-of-the-installed-base which seems entirely too constraining for an evolving, forward-looking project like the NDL.

	(e.g., railroads, land use, hot topics). Users rely heavily on reference librarians to get started, and are often not cartographically nor geographically literate.	
Newspaper and Current Periodical Reading Room	Many users are first-time visitors to LC; about 20% are regular users who come for sustained scholarly visits. Reference interviews are important. Since much of collection is copyrighted, finding aids and a few specialized items will be in NDL. Holdings information is as important as pointer information.	Helping NDL users quickly understand that few of the primary materials are online.
Law Library	Primary users are Congressional staff. People in the wrong place are directed to their local libraries; there are good law materials on the Internet so an analogous strategy can be used in NDL. Novices pointed to law encyclopedias. Main entry points are time, country, type of material (law, court decision).	Determining when to point users elsewhere
Manuscript Reading Room	80-90% are academics visiting for a few days to more than a month. When visiting, part of the entry registration process is a reference interview with a librarian. Most users have to work through multiple levels of guides to get to primary materials and are heavily dependent on librarian assistance. Entry points are mainly name of person, some subject access that point to finding aids.	Helping NDL users quickly understand that few of the primary materials are online and that there are many levels of search to work through.

There are many interface challenges posed by the assessment of users and the strategies used in the reading rooms. The most fundamental challenge is: How to serve a wide range of users who will visit the NDL? This general problem has several facets.

- The entire range of the U.S. citizenry
  - Users of different ages
  - Users with the entire range of education level
  - Users with a range of cultural and ethnic perspectives
  - Users with special physical and cognitive needs
  - Users who vary in their experience with computer technology
- The entire range of experience specific to their visit to the NDL
  - Users who vary in experience in the domain of the information problem they bring to the NDL
  - Users who vary in their experience with libraries and research collections

In addition to the challenge of individual user characteristics, the interface must support a wide variety of information needs that users bring to the NDL. Because we are focused on formalized and compromised needs, we use the term task in the following. From a user's need perspective, these tasks vary on five non-orthogonal dimensions.

- Complexity
  - The number of concepts involved

- How abstract are the concepts
- Specificity (ranging from a particular fact to interpretations)
- Quantity (the amount of information required or expected to meet the need)
- Criticality (how important it is to the user to meet the need)
- Timeliness (how long users are willing to spend or expect to spend in meeting the need.

The NDL interface must help users easily communicate some of these characteristics of their task to the system. Other general challenges echo the content challenges above.
- The interface should help users distinguish primary and secondary materials (including multiple layers of each).
- The interface should help users make links among items across different collections and reading rooms.
- The interface must capture the essential elements of the reference interview so that users can find what they need without human intervention.

In sum, the reading room interviews reinforced well-known general challenges for interface design but also raised concrete cases specific to multimedia DLs aiming to serve general populations. They also served to introduce the project team to the rich culture and collections of LC.

Questionnaire Results

Teachers and School Library Media Specialists. Responses from the teachers and school library media specialists were very similar and provide insight into the great diversity of technology penetration in U.S. schools in 1995. Thirteen of the 24 questionnaires (54%) sent to teachers were returned and eleven of the 27 questionnaires (41%) sent to school library media specialist supervisors in the state of Maryland were returned. The two-page teacher questionnaire had four general demographic questions, three scaled questions (Likert) on types of materials, five scaled questions on search strategies, five scaled questions on student search strategies, five scaled questions on computer system characteristics, three scaled questions on access to LC NDL, and one open ended question about desired useful materials for teaching. The school media specialist questionnaire was the same but media center was substituted for classroom and teachers in your school for students.

The data for the two groups showed very similar patterns. For preferred information-seeking patterns, both groups reported using all the strategies listed in the questionnaire often: browse list of potential terms; browse items with embedded links; navigate a hierarchical set of menus; search using controlled vocabulary; search using natural language. Both groups were similarly generous in estimating students' use of the entire variety of information-seeking strategies. Likewise, teachers and school library media specialists were unified in demanding a system that could be used in school and home, and is easy to learn and use. There was less unanimity within or across the groups about the need for access to large amounts of primary materials. Even less agreement was reported for the need for access to multimedia materials, perhaps reflecting concerns about the computing infrastructure in schools at that time. Both groups were far more concerned about professional use of the NDL than they were about personal uses.

Although both groups rated primary documents as important, teachers were more uniformly consensual in rating them as extremely important. Teachers rated finding aids as somewhat less important than school library media specialists (most of whom are trained as librarians). It is interesting to note that teachers rated teacher guides as somewhat less important than did the school library media specialists.

At the time of the survey (Fall 1995), there was a wide range of equipment available in these schools ranging from DOS and Apple II machines to Windows and Macintosh machines. This wide range of equipment was problematic, as significant numbers of workstations in these schools were not be capable of using the NDL. It was somewhat surprising that all but two respondents said that there was some kind of Internet access in their schools, most often in the school library media center. Perhaps even more surprising, eight of the thirteen teachers and six of the eleven school library media specialist supervisors reported having Internet access at home. Although Internet access was not defined as an IP-capable connection, some type of remote access capabilities is available either in school or at home to most of this very select group of educators.

The final open-ended question asked what types of materials respondents would find most useful for themselves or

their teachers. Not surprisingly, the teachers were more verbose in describing specific materials they could use in class, in many cases citing very specific NDL collections that presumably they had encountered in their LC training.

Examples of Teacher Suggestions:
> Images of the past & current interest linked to curriculum & interesting primary sources and documents; sounds of historical events and music.
> Major American figures from all areas of life--pictures, writing, inventions, personal items (diaries)
> Life in other times (Edison film of NY harbor, sharecroppers, immigrants, farm life, etc.)
> Documents that helped shape our nation.
> Links to related sites (archives, Smithsonian, etc.)
> Include photos of Depression, American Indian, Black studies, popular culture, World Wars
> Ease of access is vital.

Examples of School Library Media Specialist Supervisor suggestions :
> Download public domain information; indexes, finding guides, etc.
> Lesson plans for all subjects areas; support materials for lessons (text, sound, images); text will accompany images (that can be printed and used away from terminal)
> Social studies, American history documents, text, visual, census statistics
> Source information on specific subject area topics, teaching strategies & techniques for given objectives/outcomes

Day Care Center. The results of the day care center questionnaire stand in somewhat sharp contrast to the school-based data. The two-page questionnaire was administered in the fall of 1995, and had five demographic questions, five questions on computer usage, three questions on library usage, and twelve questions on expectations about the NDL. Eighty-five questionnaires were returned and provide a 1995 snapshot of the views of citizens remotely located from the Library of Congress. About two-thirds of the respondents were female and the largest number were in the 23-41 age group (another large group was composed of high school students who worked at the center). Figures 2 and 3 illustrate these data. This distribution is representative of working parents with day-care age children.
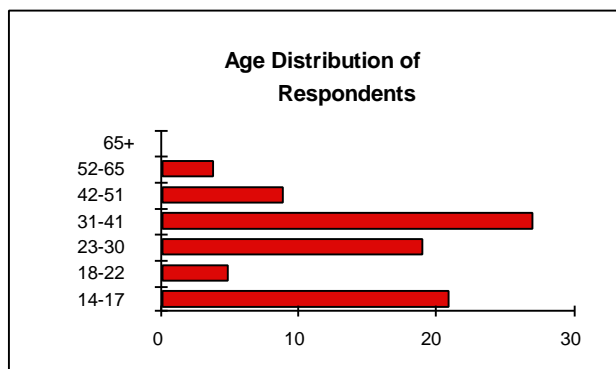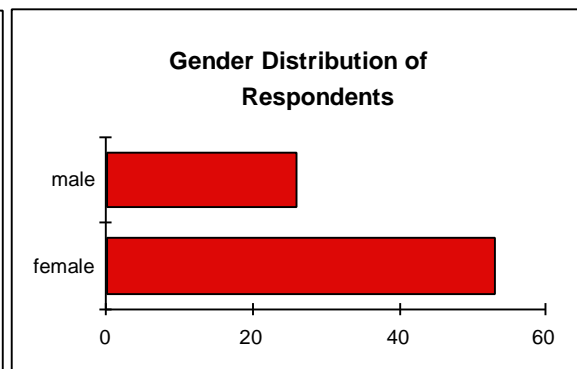
Figure 2. Age Distribution (N=85)   **Figure 3. Gender Distribution (N=79)**

There was a wide range of educational completion levels in the group, and most respondents expressed an interest in learning more about computers. The responses to computer usage questions illustrate the gap between interest to learn and actual experience. Figure 4 shows where respondents accessed computers. Ten participants did not respond to this question. As they were asked to check all that apply, the total number of responses was 117. It is interesting to note that 51 respondents reported using a computer at home (60%), a figure that exceeded national survey data on home computer penetration nationally at that time (i.e., 37% of U.S. homes had computers in early 1995, *Communications of the ACM*, May, 1995, p. 9). Forty respondents (47%) reported using computers at work. Twelve of the 18 (67%) high school student respondents reported using computers at school. Eighteen respondents (21%) reported using computers both at home and work; eleven (13%) reported using computers both at home and school; no respondent reported using computers at work and school; and five respondents (6%) reported using computers in all three locations.

Figure 5 depicts how often respondents use computers. Twenty-nine (35%) of the respondents reported using computers on a daily basis; 20 (24%) reported using computers several times a week; 15 (18%) reported using computers several times a month; 14 (16%) reported using computers once in a while, and seven (8%) did not respond to this question. Seventeen respondents selected reasons for not using computers, with access (6 respondents) and knowledge (5 respondents) the most frequent selections. Twenty-five (29%) of the respondents reported having access to some online service. These data suggest that computer usage was pervasive with well over half of the respondents in this survey using computers in a variety of settings multiple times per week and that almost one-third had some type of online access.

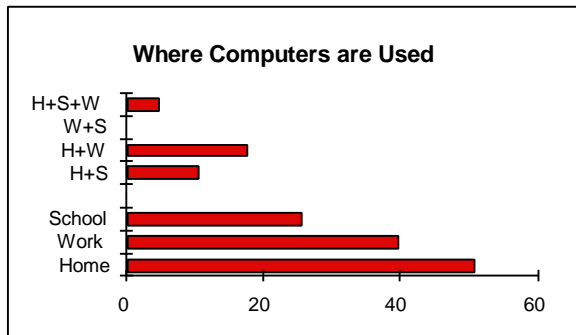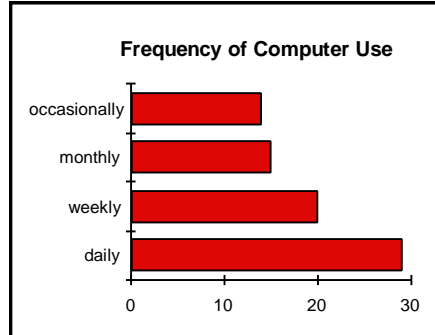**Figure 4. Where Computers are Used.**   **Figure 5. Frequency of Computer Use**

A very different portrait emerged from the responses related to the World Wide Web (WWW). Although 32 (38%) of the respondents had heard of the WWW, only nine (11%) had actually used it. Of those who did use it, only three were high school students so the adult users were beginning to use the WWW. Although this very low penetration of the WWW suggests that NDL access today would reach a small portion of these citizens, the facts that almost one-third now have some online access and over half are regular computer users suggest that WWW access will follow quickly as the NDL continues to evolve.

The questionnaire was also designed to determine the basic library use facility of the respondents. The results (see Figure 6) illustrate that most (50=59%) of the respondents use the library on occasion (less than several times a month). In response to the question about why they use libraries, responses varied across the categories, with school (40=47%) and reference (38=45%) garnering the most frequent usages (see Figure 7). It is interesting that 33 respondents (39%) selected leisure reading as a usage. Of all these usages, perhaps leisure reading is least likely to be affected by NDL availability.

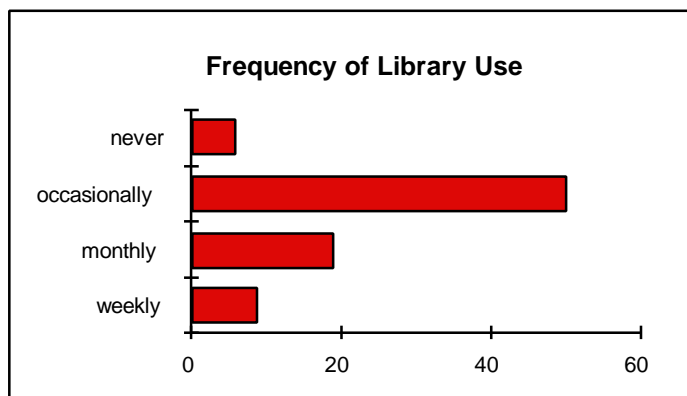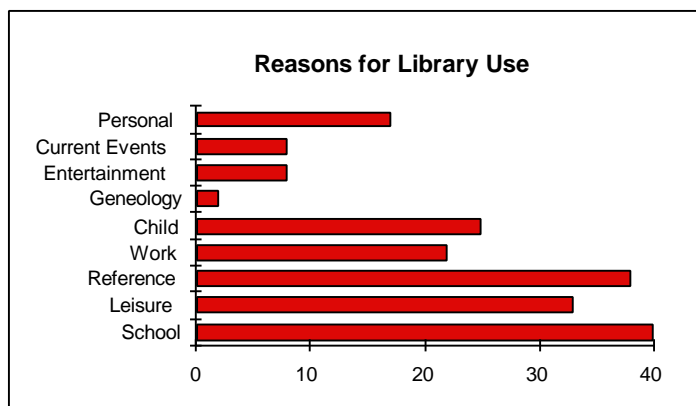**Figure 6. Frequency of Library Use (N=84)**

Figure 7. Reasons for Library Use.

Reasons for Library Use

Personal
Current Events
Entertainment
Geneology
Child
Work
Reference
Leisure
School

0   10   20   30   40

The main objective of the questionnaire was to determine what needs exist or may evolve for the NDL. This was also the most difficult objective to achieve since subjects were asked to articulate needs related to a new entity about which they knew very little if anything. Thirty-six (42%) of the respondents said they expected to use the NDL for the same reasons they currently use libraries, however, the fact that 41 (48%) selected don't know illustrates the lack of knowledge about the NDL. To address the difficulty of eliciting information about an unknown entity, the questionnaire was designed to be open ended in this section. Although, we recognized that this would make responses more difficult and many participants would skip these questions, we wanted to provide the broadest possible scope of responses. For the question about why subjects would like to use the NDL, most respondents (54=64%) wrote nothing or NA/none. Several of the respondents made generic comments about easy, fast access to information or about gaining more information or knowledge. A few made general comments about using it for school and a few noted the importance for children. Five respondents gave specific information topics or types of information (World War II, genealogy, history from a different perspective, educational materials for teachers, and maps and photographs). Two respondents noted that they would use the NDL to find information not available in their local libraries. Thus, only a small number of respondents were able to give any reasons beyond generic library use for the NDL.

A question about search strategy (title, author, subject, and browse) was not answered by 53 subjects (62%). Most who did respond noted subject (11), author (5), or some combination of subject/author/title (11 Thus, the respondents mainly expected to use the NDL as they do finding aids in physical libraries. To the question about other preferable ways to find information, 10 wrote something about computers or other electronic tools (CD-ROM) and three listed a librarian.

To the question asking if they thought they could find information easier if were available electronically, most respondents (84%) answered yes, showing confidence about their ability to use computers to find information and high expectations of the use of computers in searching for information.

Faster and easier access to information was given as an example in the questionnaire items on expected advantages and disadvantages of DLs[3]. Forty-seven (55%) of the respondents indicated one or both of these reasons. Two of those respondents also noted that there could be more information, another noted that it would be faster since they could send information to their local printer, and another noted that there could also be many more cross references. Twenty-two (26%) gave no response. Five respondents said there would be more information; six said they could do work from home or work; two noted that they could find information beyond that available in their local libraries; two thought that they could get more precise (exactly what I want) information; and one noted that it would have no advantages since the computers in libraries were too confusing to use already.

---

[3] In asking open-ended questions about phenomena that respondents may have little experience with, a questionnaire design decision is whether to give examples or not. Examples explicate and focus the question but also may lead and bias responses. We decided to use speed and ease attributes as examples of both possible advantages and disadvantages of DLs. Participants were asked to speculate on advantages and disadvantages that might be expected with NDL use.

The disadvantage question gave "slower more difficult access to information" as an example. Six respondents noted that access might be slower; seven said that it might be difficult to use the technology, three noted generic difficulty, and three noted difficulties related to finding information. Forty-five (53%) of the respondents did not answer or said no disadvantages. Two respondents were concerned with access to terminals; five noted potential problems with technology failure; three noted concerns about system overload (busy lines); one noted potential information overload; one was concerned that not everything needed would be available; three were concerned about costs; one noted that no librarian would be available to help; and one thought the entire idea was a waste of tax dollars.

The final open-ended question asked participants to list ways that they might use NDL materials. Thirty-eight (45%) of the respondents did not write anything or wrote NA. Fifteen (18%) specifically noted schoolwork or projects; seven (8%) noted to help children (one said younger brother); eight (9%) said reports, presentations, or projects; and four said for general reference. The remaining responses related to some sort of personal knowledge: five noted general personal information; five noted travel or vacation information; three noted general knowledge acquisition; two said knowledge of history; two said genealogy; and one each noted music/literature, multiple sclerosis, and art inspiration. One other respondent said for entertainment, and one said "just like a library."

Although it was difficult for the respondents to speculate about the pros and cons of the NDL they knew little about, the comments given do cover the commonly expressed advantages of digital libraries and provide a wide range of concerns about technical and intellectual challenges to the NDL.

Summary of Questionnaire Analyses. At the time of this assessment, the general populace showed interest in computers and the information resources networks offer but had limited first-hand experience with the WWW. Home computer use was high with some sort of online access becoming more common. Schools used a wide range of computers and many have some type of online capacity in the building but few had it in classrooms. The day care results show that few respondents gave specific reasons for using the NDL and cited no specific collections. The media specialists gave some specific reasons but no specific collections. The teacher respondents who had participated in LC training cited specific materials. These results reinforce the theme of diversity in users, needs, and computational settings and suggest that the NDL must be purposefully introduced to potential users.

LC Document Analysis

Reading Room Handouts. The first documents examined were those available in the reading rooms as handouts for visitors. There is a rich set of knowledge captured in these handouts and the NDL interface must provide some of their functionality. Some of the handouts related to physical space and working hours - functions less critical in the NDL but still needed in terms of services and the interplay between digital and physical collections. Others provided descriptions of the collection and policies for using the collection and are clearly applicable to the NDL. A challenge is finding ways to integrate the many different reading room descriptions into a common introduction for the integrated NDL. Documents that help users actually use the collection or conduct research in the field will also be as valuable for NDL users as they currently are for LC reading room users. Finally, the reading rooms have created specialized documents related to particular user needs or hot topics (e.g. current events, anniversaries). The existence of these documents illustrates the need for some similar service in the NDL interface.

All these documents reinforce the complex nature of the individual reading rooms, let alone the complexity of the aggregate LC. Preliminary versions of the interface can simply point to electronic versions for these documents, however, specialized electronic versions will emerge over time and be integrated into the system help and guided tour components.

LC User Studies. The two LC user studies we used offer a rich set of specific interface recommendations. Some of the impressions and commentary at the early American Memory test sites strictly reflect individual preferences, however some trends were common across many sites, for example:
- Students preferred buttons to pull-down menus,
- There were many instances of poor understanding of search strategy (e.g., misunderstanding Boolean AND, lack of systematic reformulations or use of results),
- Complaints about time to use the system (both physical access due to a single workstation and response time),

- Teachers noted that linkages between the primary material and their curricula would be beneficial (a common use for the content and system was as an enrichment adjunct rather than an integral part of the course), and several noted that it is essential that materials can be printed and/or saved to disk.

The Prints and Photographs WWW Public User Interviews identified a variety of navigational problems. For example the lack of basic graphical user interface literacy made it difficult to use a mouse, scroll, use pull-down menus. Moving among levels of representation was also difficult (e.g., between a full image, thumbnail, and brief display), and losing buttons after scrolling was frustrating.

The search strategy problems identified included:
- Confusion regarding relevance ranking of results,
- Need for easy-to-use alphabetical subject headings,
- Limiting hit lists to 100, or having too much to look through, and
- Need for posting data (frequency of occurrence) after all moves.

Both of these sets of results highlight challenges for the NDL interface. In addition to reinforcing the importance of computer experience, these findings illustrate: the need for crisp and intuitive dynamics between screens/windows so that users can focus on different levels of representation easily or possibly juxtapose them; judicious use of any scrolling functions; and the need for clear yet powerful search articulation and results display.

Email Messages. The sample of representative electronic mail messages answered by LC staff provides an interesting mosaic of users who are actually using the infant NDL and the types of information needs and problems they bring to the Library. Based upon a content analysis of user messages, staff responses, and a brief classificatory commentary provided by staff for the exchanges, the messages fall into three broad categories: system, user needs, and miscellaneous.

The "system" category includes the following types of messages:
- Praise (compliments about the NDL)
- Crashes (complaints or inquiries about server or telecommunications problems)
- Technical questions or suggestions (difficulty launching video players on client, suggestions for alternative data formats)
- Suggestions to add
  - Content (add specific collections)
  - Indexes or services (one interesting request was for a zipped index of the NDL that could be downloaded and used locally to save connect charges)
- Corrections (e.g., mislabeled photograph).

The "user needs" category includes the following types of messages:
- Reference questions (e.g., census data for NY state 1600-1900 and how to find it)
- Requests for special collections
- Requests for specialized services (e.g., photograph copy requests online)
- Requests for help (Some have done searches and want help in narrowing; others confused about what to ask, where the mail is going).

The "miscellaneous" category includes the following types of messages:
- Contributions of two types
  - Specific items (e.g., want to donate a photograph)
  - URLs (requesting LC to link to their web sites)
- Job seeking (to work on NDL)
- General communication (some users reply with thanks and commentary to a reply to their earlier message--want to maintain dialogue).

Taken together, these documents capture a wide range of knowledge about the LC, the users who use it, and the types of information needs they bring to it. The existing documents must be captured and integrated into the NDL. Analysis of future email will surely contribute important data as the system evolves and new, systematically

scheduled user studies are highly recommended.

User Type/Task Taxonomy.

A central theme emerging from this investigation is diversity. There were wide ranges in users, the tasks they bring to the NDL as manifestations of their needs, the technological settings they work in, and in the wide variety of content that makes up the NDL. To integrate the results from the different components of the investigation we revisited the nature of users and needs in light of the types of user communities identified before data collection began. It was clear that the three communities (LC staff, current scholars, and prospective NDL users) were too broad to capture the diversity or to fully guide the interface design.

A more fine-grained classification is possible by combining the analytical framework of users and tasks (Marchionini, 1995) with the empirical data from this investigation. The analytical framework crosses users, tasks, and the personal situation that motivates the search.

Users have a host of individual characteristics, preferences, and experiences, which are not orthogonal:
- Personal attributes (physical, cognitive, social).
- Experience in the domain of knowledge related to the information need.
- Experience in using library systems (including information technology) and research techniques.

Five non-orthogonal dimensions of the information-seeking task are:
- Complexity (both the number of facets involved and the level of abstraction).
- Specificity (how confident the user is to determine the accuracy and completeness of results).
- Quantity (the amount of information expected).
- Criticality (how important it is that the need be met).
- Timeliness (how much time the user is willing or expects to invest).

Analytically, we would define scales for each dimension (e.g., low, average, high) and populate the resulting matrix with examples or cases. Assuming three points per dimension, we would have 27 cells for user characteristics and 243 cells for tasks. Crossing the two matrices would yield 6561 cells to fill. Adding the many types of motivations and situations that contextualize specific instances of information seeking similarly expands the theoretical possibilities. Clearly, this approach is a) impractical due to the large number of variations for which interface features are considered, and b) minimal and simplistic on the user characteristics side.

A more realistic approach is to collapse some of the dimensions in light of the empirical evidence collected in the reading rooms, surveys, and document examinations. Based on the evidence, factors that characterize users are:
- Motivation (the personal situation that brings one to the Library, including quantity, criticality, and timeliness)
- Domain knowledge (related to the particular need)
- Library system knowledge (including information technology)
- Focus (a combination of complexity, quantity, and specificity) and
- Time allocated (combining timeliness and criticality).

These factors were applied to the different types of users and user needs described in the reading room visits. Such an approach yielded nine user classes listed in Table 4. These classes are not exhaustive, nor are they mutually exclusive. Moreover, any individual belongs to a class for each information need (i.e., in different visits or sessions users may fall into different classes). It is highly likely that this taxonomy will change as the NDL evolves since it is rooted in data that come from a mix of current users of the physical LC, early adopters of the nascent implementation, and non-users asked to speculate about using it.

Table 4. User Taxonomy.

1. LC staff. high motivation, medium domain knowledge, high library system knowledge, high focus, and limited time allocations

2. Hobbyists (e.g., genealogy, Civil War, railroads, other examples). high motivation, typically high domain

knowledge, a range of library system knowledge, high focus, and high time allocations

3. Scholars (e.g., historians, sociologists, anthropologists, authors). high motivation, high domain knowledge, high library system knowledge, high focus, and high time allocations

4. Professional researchers (e.g., picture researchers).   high motivation, medium domain knowledge, average to high library system knowledge, very high focus, and medium time allocations,

5. Rummagers (browsers) (e.g., PhD students looking for topics; scholars looking for new directions, topics).   high motivation, medium domain knowledge, range of library system knowledge, low focus, and medium to high time allocations

6. Object seekers (e.g., some authors, CD-ROM/multimedia developers, TV/video producers, and instructional materials developers). high motivation, range of domain knowledge, low library system knowledge, high focus, and low to medium time allocations

7. Surfers (e.g., those who are curious, those who bump into the NDL, etc.). low motivation, low domain knowledge, low library system knowledge (but may be high computing system knowledge), low focus, and very low time allocations

8 Teachers K-16   medium motivation, medium to high domain knowledge, low to medium library system knowledge, medium focus, and low time allocations

9. Students K-16 low to medium motivation, low domain knowledge, low library system knowledge, low to medium focus, and low to medium time allocations.

This taxonomy was very useful to the prototyping phase of the project by providing detail to the general design vision and acting as a "reality check" for design discussions. Three additional uses are apparent for continued development of DL interfaces. First, it can guide the development of features that substitute for the reference interview. Since a human resource will not be readily available, the system must provide ways for users to articulate specific needs (e.g., queries) as well as contextual information (e.g., granularity and scope of need). A set of user templates varying according to the parameters above was sketched early in the prototyping phase of work. That work was not pursued as other features with higher priority were addressed. Second, the taxonomy can provide the basis for variations in discussions of interface features such as help, tours, and tutorials. Third, the taxonomy can provide the basis for testing prototypes with scenarios and representative tasks.

Design Implications.

This needs assessment identified many design challenges specific to LC. Table 5 summarizes these challenges. As the design process proceeded, these challenges were prioritized and specific prototype features were created. We chose to address the general challenges of content and strategies in the prototypes we implemented.

The needs assessment identified a wide range of users and needs and corresponding interface design challenges. The interface must communicate to users what the NDL is and is not, i.e., is composed of multiple but not uniform data types (both formats and levels of representation) and is not the entire LC let alone the entire Internet. The interface must help users see that the NDL has multiple but not uniform access points (entry points are useful for some objects but not others). It must support a range of search strategies from hierarchical selections to formal and comprehensive queries so that novices and experts are all well served. It must provide help and guidance and emulate the reference interview.   It must be usable on a variety of platforms and sensitive to a variety of physical infrastructures.   The interface must allow users to meet their needs without the benefit of a human intermediary and for casual users it must do so without long processing delays.

These are strong requirements and the degree of success varied according to how prioritized the challenges. The general user needs identified in this assessment oriented the design team in understanding both the needs of potential users and the characteristics of the rich and varied collections of materials scheduled for the NDL. In addition to

contributing to the eventual prototypes, the multifaceted approach to assessing needs was applied and extended in other projects in different public access settings. Hert and Marchionini built upon these methods to develop a user-task taxonomy for federal government statistical web sites (Hert & Marchionini, 1998; http://ils.unc.edu/~march/blsreport/mainbls.html).

Table 5. Interface Challenges identified the users need assessment and to be addressed by the prototypes.

Content - interface challenges

General interface challenges across reading rooms

developing a framework characterizing for users the granularity, size, and nature of objects in the NDL across all the reading rooms
communicating to the user what items are NOT in the NDL
supporting users in the NDL without human intervention
creating an interface that is accessible to users with state-of-the-market technology
inventing new techniques to search for multimedia objects and to integrate those techniques into the interface (e.g., visual and audio query languages)

Specific interface challenges from different reading rooms

audio access (searching as well as download times)
access to multi-format items (i.e., a sound track and textual fieldnotes)
user specification of areas/regions on maps
place name ambiguity
identifying and representing links
distinguishing documents that help one do genealogical research from the primary materials
integrating full-text and controlled vocabulary searching both across and within collections
distinguishing finding aids levels and primary material
need for linkages from browsable covers to bibliographic records and microfilm text available at LC
image searching (including displaying series of related images and images in challenging formats such as panoramas and oversized posters)--possible use of P&P thesaurus;
enabling researchers to absorb enough of the context for historical images and captions (why images were made) to deal sensitively with content that might otherwise be deemed offensive.

Users and strategies - interface challenges

General interface challenges across reading rooms

serving a wide range of users
serving a wide variety of information needs
helping users distinguish primary and secondary materials (including multiple layers of each)
helping users make links among items across different collections and reading rooms
capturing the essential elements of the reference interview so that users can find what they need without human intervention

Specific interface challenges from different reading rooms

specifying geographical areas
potentially huge files to transfer and display
determining when to point users elsewhere
helping LC NDL users to quickly understand that few of the primary materials are online
helping LC NDL users to quickly understand that there are many levels of search to work through
identifying copyrighted materials (so patrons do not find pointers to them and expect they can come and copy them at LC)
supporting hot topics, specialized exhibits

overcoming patrons lack of knowledge about media (e.g., how pictures were produced at different times)

Conclusion

These case studies are meant to illustrate three different DL situations in which multifaceted and flexible approaches to human-centered DL design and evaluation were taken. They are juxtaposed to highlight some of the general approaches to DL design and evaluation while demonstrating the unique challenges of each setting. Practical design is a compromise between the top-down BITWC and bottom-up, organically grown system. Our central assumption is that practical design is guided by general design visions that are informed by multifaceted approaches to ongoing assessments of user needs and system impact

The Perseus DL evaluation illustrates how multifaceted approaches can over time yield evidence of significant effects across a broad community of users. In this case, the user needs were perceived by the DL designers, the primary user populations are university students and professors, and the primary design goals were to provide large quantities of primary text and image materials with little or no interpretation. The evaluation effort was begun at the inception of the project, used a multiplicity of methods to collect data on a large set of evaluation questions tied to a matrix of stakeholder, DL mission, and technical dimensions. Early results demonstrated typical effects of mechanical advantage, the difficulties of learning to teach with technology, and some examples of new kinds of learning but it was only through the longitudinal results that systemic change in larger sets of stakeholders (the field of classics) began to emerge. Without a steady and varied stream of data, these larger effects would likely not have been demonstrated except through some future historical efforts.

The Baltimore Learning Community case illustrates the difficulty of influencing complex human behavior (teaching) through external interventions regardless of the informational or technical innovation. In this case, user needs were determined by state and local administrative curricula and assessment programs, the primary user population is middle school science and social studies teachers and their students. The primary design goals were to provide relatively small quantities of multimedia materials that teachers could map to well-defined curriculum guidelines in creating lesson plans. The experience of this case demonstrates that advanced technical solutions and high-quality content are not sufficient to initiate or sustain community in settings where day-to-day practice is strongly determined by personal, social and political constraints.

The Library of Congress NDL case details how one multifaceted user needs assessment was conducted and used to guide user interface prototypes. User needs were largely unknown and the first phase of the design project was devoted to investigating potential needs in a variety of settings. This DL aims to serve the entire US population and the design goals were to provide universal access to historical materials in a variety of formats and at the entire range of granularities from bibliographic records and finding aids to full manuscripts and images. The user needs assessment yielded a set of user types and interface design challenges that guided the eventual prototypes and may find use in other DL settings.

Although these three DLs aim to serve different user populations and had distinct design goals, there are several general design and evaluation principles that resonate across the cases. First, DL designers and evaluators must know the user(s). This general rule is followed through needs assessments and the different cases illustrate that needs may be externally or internally motivated, well-defined or ill-defined, and narrow or wide-ranging. The three cases demonstrate the benefits of using direct methods such as interviews and observations and indirect methods such as document analyses or interviews with intermediaries. It is only through understanding user characteristics and needs that DL designers can build tools to help users map information needs onto DL tasks and evaluators can develop a good set of questions to guide research.

Second, it should be clear that DL design and evaluation are processes that aim to create and understand complex products. As such, they proceed in stages over multiple iterations. DLs are constantly evolving to meet the needs of users who grow and learn and technologies that advance exponentially. Design and evaluation efforts must be embedded into the overall DL management culture and applied on an ongoing basis.

Third, in addition to the complexities of change above, DLs must serve people with a wide range of characteristics and needs. To do so requires that designers and evaluators create and use multiple and flexible systems and tools. Designers must create overviews that allow DL patrons to quickly understand what is and is not available in the

collection. We must provide alternative entry points and search tools so that diverse user needs and experiences can be accommodated. Evaluators must seek data from all fronts and create ways to integrate those data streams to make judgments about progress and next steps. Just as the blind men could only understand parts of the elephant they experienced first hand, our understanding of DL design and evaluation processes will require us to examine multiple views through multiple lens'.

References

Belkin, N. (1980). Anomalous states of knowledge as a basis for information retrieval. *Canadian Journal of Information Science*, 5, 133-143.

Braha, D. & Maimon, O. (1997). The design process: Properties, paradigms, and structure. *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, 27(2). 146-166.

Brooks, F. (1975). *The mythical man-month: Essays on software engineering*. Reading, MA: Addison-Wesley.

Dervin, B. & Nilan, M. (1986) Information needs and uses. In M. Williams (Ed.) *Annual Review of Information Science and Technology* (Vol 21, pp. 3-33), White Plains, NY: Knowledge Industries.

Ding, W., Marchionini, G., & Tse. T. (1997). Previewing video data: Browsing key frames at high rates using a video slide show interface. *Proceedings of the International Symposium on Research, Development, and Practice in Digital Libraries*, (Tsukuba, Japan) p. 151-158.

Enomoto, E., Nolet, V., & Marchionini, G. (1998). Video and dynamic query capability in schools: Implications for learning in a networked community. *Bulletin of Science, Technology, & Society*, 18(6), 432-440.

Flagg, B. (Ed.). 1990. *Formative evaluation for educational technologies*. Hillsdale, NY: Lawrence Erlbaum Associates.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligence*. New York: Basic Books.

Harter, S. & Hert, C. (1997). Evaluation of information retrieval systems: Approaches, issues, and methods. . In M. Williams (Ed.) *Annual Review of Information Science and Technology* (Vol 32, pp. 3-94), Medford, NY: Information Today Inc.

Hert, C. & Marchionini, G. (1998). Information seeking behavior on statistical websites: Theoretical and design implications. *Proceedings of the 61st Annual Meeting of the American Society for Information Science*. (Pittsburgh, PA, Oct. 25-29, 1998). P 303-314.

Hix, D. & Hartson, H. (1993). *Developing user interfaces: Ensuring usability through produce & process*. NY: John Wiley & Sons.

Komlodi, A. & Marchionini, G. (1998). Key frame preview techniques for video browsing. *In Proceedings of Digital Libraries '98: The third ACM Conference on Digital Libraries* (Pittsburgh, PA, June 23-26, 1998). NY: ACM Press. 118-125.

Krikelas, J. (1983). Information-seeking behavior: Patterns and concepts. Drexel *Library Quarterly*, 19(11), 5-20.

Library of Congress. (1995). *The National Digital Library Program-A Library for All Americans*. Washington, DC. February, 1995.

Library of Congress. (1995). *P&P World Wide Web Public User Interviews*. Washington, DC. Prints and Photographs staff, May 10-23, 1995.

Library of Congress. (1993). *American Memory User Evaluation 1991-1993: Site Summaries. An Addendum to the Main Report*. Prepared by the American Memory User Evaluation Team. Washington, DC December, 15, 1993.

Marchant, M. (1991). What motivates adult users of public libraries? *Library and Information Science Research*, 13(3), 201-235.

Marchionini, G. (1995). I*nformation seeking in electronic environments*. NY: Cambridge Univ. Press.

Marchionini, G. (1998). Research and development in digital libraries. Allen Kent (Ed*.) Encyclopedia of Library and Information Science*, Vol. 63. 259-279.

Marchionini, G. (1995). The costs of educational technology: A framework for assessing change. *Proceedings of ED-MEDIA 95 World Conference on Educational Multimedia and Hypermedia* (Graz, Austria, June 17-21, 1995). Charlottesville, VA: AACE. 33-38.

Marchionini, G. & Crane, G. (1994). Evaluating hypermedia and learning: Methods and results from the Perseus Project. *ACM Transactions on Information Systems*, 12(1), 5-34.

Marchionini, G. & Fox, E. (in press). Progress toward digital libraries: Augmentation through integration. Guest Editors' introduction. *Information Processing & Management*.

Marchionini, G., Plaisant, C., & Komlodi, A. (1998). Interfaces and tools for the Library of Congress National Digital Library Program. *Information Processing & Management*, 34(5), 535-555.

Nielsen, J. (1993). *Usability engineering*. Boston: AP Professional: Academic Press.

Norman, D. (1988). *The design of everyday things*. NY: Doubleday.

Paisley, W. (1980). Information work. *Progress in Communication Sciences*, 2, 113-165.

Petroski. (1996). *Invention by design: How engineers get from thought to thing*. Cambridge, MA: Harvard U. Press.

Plaisant, C., Marchionini, G., Bruns, T., Komlodi, A., Campbell, L., Bringing Treasures to the Surface: Iterative design for the Library of Congress National Digital Library Program , *CHI 97 Proceedings*, Atlanta GA, 22-27 March 1997, ACM New York, 518-525

Shneiderman, B. (1998). *Designing the user interface: Strategies for effective human---computer interaction* (3nd ed.). Reading, MA: Addison-Wesley.

Simon, H. (1996). *The sciences of the artificial (3rd Ed.)*. Cambridge, MA: MIT Press.

Slaughter, L., Shneiderman, B., & Marchionini, G. (1997). Comprehension and object recognition capabilities for presentations of simultaneous video key frame surrogates*. Proceedings of Research and Advanced Technology for Digital Libraries, First European Conference* (Pisa, Italy, September 1-3, 1997), p. 41-54.

Sternberg, R. (1985). *Beyond IQ: A triarchic theory of human intelligence*. New York: Cambridge U. Press.

Suchman, E. (1967). *Evaluative research: Principles and practice in public service and social action programs*. NY: Russell Sage Foundation.

Taylor, R. (1962). The process of asking questions. *American Documentation*, October, 391-397.

Wilson, T. (1981). On user studies and information needs. *Journal of Documentation*, 37(1), 3-15.