DESIGN PRINCIPLES OF HELP SYSTEMS FOR DIGITAL LIBRARIES

Jose Perez-Carballo, California State University, Los Angeles
Iris Xie, University of Wisconsin-Milwaukee
Colleen Cool, Queens College/CUNY

ABSTRACT

In this paper we present and discuss a set of principles for the design of help systems for digital libraries. Previously a user study (Xie and Cool, 2009) was performed by two of the authors in order to identify a number of Help Seeking Situations encountered by novice users of digital libraries. In this paper we continue that work as we examine those help-seeking situations and, based on them, propose a set of principles and guidelines for the design of help systems for digital libraries. Some of the design principles discussed here may have been proposed by other researchers or may have been used in previous user-interfaces but the contribution of this paper is that it validates and connects the design principles with specific help-seeking situations encountered by users observed and categorized in our user study. Further user studies of other digital libraries as well as user interfaces that implement the proposed principles may provide evidence that could be used to modify and refine the design principles.

INTRODUCTION

The problem addressed by this project is that help-seeking situations are not well understood and at the same time, the design of help functionalities has proceeded without the benefit of such knowledge. The consequence is that the standard help features present on most digital library systems, are often not very effective (Xie, 2006; Xie & Cool, 2009) and particularly not helpful for novices.

The first steps of the project were to identify the types of problems novice digital library users experience for which they require help and to better understand the nature of these help-seeking situations. We conducted a user study (Xie & Cool, 2009) in order to identify the Help Seeking Situations (HSSs) encountered by novice users of digital libraries (see section “The User Study”). In this paper we examine the HSSs identified in the study (op. cit.) and propose a set of principles and guidelines for the design of help systems for a digital library (DL) that are intended to address those specific HSSs. Some
of the design principles that we discuss here have been proposed also by other researchers or used in previous user-interfaces. What this paper does is to validate and connect those principles with specific help-seeking situations encountered by users observed and categorized in our user study.

In order to design a usable help system it is necessary to understand what lead users to use help. The user study that we conducted (op. cit.) identifies the kinds of situation that cause users to consult an online help system. The same study also illustrates the cases in which users failed to find the help information they needed. Understanding of users' Help Seeking Situations allows the designer to put effort and resources in the areas that will result in the highest pay-off for users.

In our user study we identify 15 Help Seeking Situations (HSSs). The HSSs and the corresponding user interface features are listed below (see Table 1). For each situation in the list of HSS, the design features that would be relevant are listed.

**PREVIOUS WORK**

Previous research has evaluated a variety of help features in different types of Information Retrieval (IR) systems. Our approach is to study users first, in order to understand the help-seeking situations that give rise to help-seeking interactions and then, based on those situations, propose a set of design principles and guidelines to build online support systems.

Several collections of design guidelines and principles for the design of user interfaces have been proposed. For example, Shneiderman offers a list of principles in his book (Shneiderman, 1997). Other examples are the MITRE design guidelines (Smith & Mosier, 1986) and similar documents from Apple (Apple Inc., 2009), Microsoft (Microsoft n.d.), and Yahoo (Yahoo!, 2010). The latter is an example of a growing trend to publish user interface design patterns, in this case for web pages. These documents present recommendations to application developers. Their dual purpose is to make user interfaces more usable as well as the, sometimes conflicting, purpose of specifying the look and feel expected in the corresponding environment. The guidelines proposed by these documents may be the result of the experience, taste, and, in some cases, user studies performed by the corresponding organizations. The document from Apple emphasizes the need to "involve users in the design process", "analyze user tasks", "build prototypes" and "observe users". It encourages developers to "see where people have difficulty using [their] product and then use that information to improve [the] product". This is not unlike this paper, were we describe the specific problems that novice users of a digital library encounter when trying to use the help system, and then propose guidelines to design better help systems.
The principles we propose here are based on observations of how novice users behave when seeking help while using digital libraries. The user interface design principles proposed in this paper apply specifically to the design of the help system's user interface of a digital library.

<table>
<thead>
<tr>
<th>Table 2: DESIGN FEATURES FOR EACH HELP SEEKING SITUATION</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HSS (Help seeking situation)</strong></td>
</tr>
<tr>
<td>---------------------------------</td>
</tr>
<tr>
<td>Inability to get started</td>
</tr>
<tr>
<td>HSS1: unclear about the topic, HSS2: unfamiliar with digital libraries</td>
</tr>
<tr>
<td>Inability to identify collections</td>
</tr>
<tr>
<td>HSS3: inability to identify relevant collections</td>
</tr>
<tr>
<td>Inability to browse</td>
</tr>
<tr>
<td>HSS4: inability to browse information</td>
</tr>
<tr>
<td>Inability to create search statements</td>
</tr>
<tr>
<td>HSS5: inability to use appropriate query syntax</td>
</tr>
<tr>
<td>HSS6: inability to select appropriate query terms</td>
</tr>
<tr>
<td>HSS7: inability to limit searches to certain fields</td>
</tr>
<tr>
<td>Inability to refine searches</td>
</tr>
<tr>
<td>HSS8: inability to refine searches for different aspects of the search topic, HSS9: inability to identify other approaches for information, HSS10: inability to refine searches in general</td>
</tr>
<tr>
<td>Inability to monitor search</td>
</tr>
<tr>
<td>HSS11: inability to monitor search</td>
</tr>
<tr>
<td>Inability to evaluate results</td>
</tr>
<tr>
<td>HSS12: inability to identify relevant documents</td>
</tr>
<tr>
<td>HSS13: inability to identify specific information</td>
</tr>
<tr>
<td>HSS14: inability to compare items retrieved</td>
</tr>
<tr>
<td>HSS15: inability to verify authority and accuracy of retrieved documents</td>
</tr>
</tbody>
</table>
Some of the earliest help-system designers recognized the need to understand the help-seeking situations users encounter when using the system. In an early paper about help interfaces published in the first hypertext conference (1987) Walker (1987) described the Document Examiner (DE), a help system based on hypertexted documents. DE offered searching capabilities that were based on consideration of the strategies that people use in interacting with paper documents. As is often the case, Walker designed the interface providing features she thought would be useful to users and then ran some user surveys. One of the main drawbacks of the system from the point of view of the users was performance. Even though the documents took only 10MB (a significant size at the time) some lookup operations could take as long as 10 seconds. The system provided enough advantages that users were still willing to put up with a performance that today’s users would find intolerable. This fact also illustrates the fact that user behavior and user expectations may change significantly over time as the skills of the average user change, the technology itself changes, and the expectations of the average user change.

Borgman (Borgman, 1996) argued that online catalogs are difficult to use because their design does not incorporate sufficient understanding of user searching behavior. In several papers (Borgman, 1987; Borgman, 1987b) she examined studies of information retrieval systems searching for their implications for online catalog design.

Cool and Xie (2004) reported that people tend to use help mechanisms infrequently because of inadequacies in the interactive capabilities of these systems in their support during help-seeking situations.

Dworman and Rosenbaum (2004) suggested that users failed to use even help systems that did contain useful and well-organized information because of how they perceived and accessed help systems. They listed several reasons for users' inability to use help:

Cognitive blind spots: users appear to simply not see help even when the delivery mechanism is right in front of them.
Distraction aversion: users often are unwilling to leave their current task to start a help-seeking task.
Fear: users may be afraid to suffer the potentially dire consequences of leaving their current task.
Refusal to admit defeat: users insist that they can figure out a solution and don't need to look it up.
The word “help” itself: users may access hints, tips, and quick-reference guides, but refuse to click on something called “help.”
It is argued that people seldom find the answers to their questions when they try to use the help of search systems (Haas, Brown, Denn, Locke, & Shneiderman, 2006). Not surprisingly, the panelists agreed that there are problems with help systems. Locke, one of the participants, said that before we can design a help system we need to know who our users are and what brings them to help. Fundamental divisions of the help system may start with the distinction between definitions and procedures that relate to the software interface and those that relate to the content the interface presents.

Shneiderman argued (op. cit.) for multi-layered interface designs which enable first-time and novice users to begin with a limited set of features at the first layer. Users can remain at the first layer, then move up to higher layers when needed or when they have time to learn further features. He pointed out that there are still interesting design problems in how to define the layers (consider, for example, layering by function or by application or task), it is a very promising strategy that needs further exploration and testing.

Denn (op. cit.) made the distinction between interface help and content help, defined as support for the use of the content and concepts contained within an information system. She discussed the methodological issues of designing studies to explore interactions between user and task characteristics. She pointed out the importance of defining the kinds of metrics that we can use to measure them, as well as how to apply the results to the design of integrated interface and content help.

According to Jansen (2005) searchers seldom make use of the advanced searching features that could improve the quality of the search process because they do not know these features exist, do not understand how to use them, or do not believe they are effective or efficient. His study suggests that search effectiveness could be improved by suggesting or implementing assistance automatically.

Krull, Friauf, Brown-Grant, & Eaton (2001) reported studies of help systems that never were overwhelmingly successful in enabling users to complete tasks. The two consistent problems they identified were that the organization of help information was not clear to users and that help topics were expressed in product-specific vocabulary rather than user vocabulary.

In the analysis of our user study (Xie & Cool, 2009) we find evidence of many of the problems mentioned above. In this paper we propose principles and mechanisms to address them.

THE USER STUDY

Two of the authors of this paper (Xie & Cool, 2009) address the following questions in a study of novice users of digital libraries:
What are the typical types of HSS experienced by novice users of digital libraries?

How do novice users use Help in digital libraries? Specifically, which Help features are used to address what types of problems?

One hundred and twenty novice participants from Milwaukee and New York representing general public with various ethnic backgrounds, education and literacy levels, computer skills, occupations, and other demographic characteristics were recruited to participate in the study. Multiple data collection methods were used for this project: pre-questionnaires, think aloud protocol, transaction logs, and post-questionnaires. First, participants were asked to fill out a pre-questionnaire about their demographic information as well as their experience in using different IR systems and help mechanisms.

Next, participants were instructed to conduct 3 tasks in two digital libraries respectively: American Memory (http://memory.loc.gov/ammem/index.html), this digital library will be referred as LOCDL in this document; New Zealand Digital Library (www.NZDL.org), which will be referred to as NZDL in this document. They were instructed to think aloud during their process. The entire search sessions were logged and recorded unobtrusively by using Morae™ software. The three types of tasks were:

A task that requires the user to explore each digital library. For example: “use three different approaches to find an 1895 map of Yellowstone National Park.”

A task that requires users to search for specific information. For example: “what is another name that was used for Bubonic Plague in the nineteenth century?”

A task that requires users to search for information that has a variety of characteristics related to content, format (audio clip, video clip, etc.), and search strategy required. For example: “identify at least four issues regarding immigration policy in the U.S., using as many sources from the digital library as you can. Each issue you identify should have a different source.”

Finally, participants were also asked to fill in post-questionnaires that solicited information related to their experience in using help mechanisms in accomplishing their tasks.
HELP SEEKING SITUATIONS (HSSS)

The following is the list of the 15 types of Help Seeking Situations (HSSs) identified in the user study that lead novice users to try to use help features of digital libraries. The principles are described in a section below. See Table 2 for a list of the HSSs and their corresponding design principles.

HSS1: inability to get started - unclear about the topic. Some participants could not get started because they felt they needed more domain knowledge. Therefore, they did not know how to start their searches. Some users tried to use the help system in order to find information about the domain. Example: some users tried to ask the help system “where is Yellowstone” when dealing with the task: “use three different approaches to find an 1895 map of Yellowstone National Park.”

HSS2: inability to get started - unfamiliar with digital libraries. Users try to get information about the kinds of searches they could do. Example: “find different search strategies supported by the DL (e.g. search by phrases).”

HSS3: inability to identify relevant collections. Users need help choosing which collection to search. Example: “which would be the better collection to search given that the question is about immigration?”

HSS4: inability to browse information. This type of help-seeking situations was mainly caused by lack of system knowledge, poor interface design as well as unintuitive browsing mechanisms.

HSS5: inability to formulate query statements. Users had trouble constructing statements that could be submitted to the system as queries. This led them to try to find information about search features and query syntax from the help system. They tried to construct queries using syntax learned in other systems but which may not work in the digital libraries used in the study. In particular, they tried to use boolean operators, quotes for phrases, and + for the “must be present” feature.

HSS6: inability to select appropriate query terms. Users had trouble finding the right terms to use in their queries due to lack of domain knowledge.

HSS7: inability to limit searches to certain fields. Users do not know how to specify the facets of their search, for example: publication year, author, title, format of documents, etc. Problems in the search system design caused some of these situations. Example: “how to search just for maps”, “how to specify a date”.

HSS8: inability to refine searches for different aspects of the search topic. This situation was often caused because of lack of retrieval and domain
knowledge. Examples: “how to find more than one source for a single issue”, “having found different name for bubonic plague, tried to find countries and cities impacted by the epidemic.” “I am overwhelmed by the amount of the results. I don't want to read all the results to identify multiple issues of immigrant policy.”

HSS9: inability to identify other approaches for information. One of the study tasks asked participants to apply three different approaches to find the same information. This situation is often caused by lack of users' knowledge about the system as well as the domain. Examples: “Let me try browse through the collection [another approach] and see if I can find the map.” “I'm trying to think of another collection besides conservation or maps that would include a map of Yellowstone.”

HSS10: inability to refine searches in general. After their query fails, users need help to create a new one. Example: users tried to re-write their queries after the original query submitted did not retrieve any documents, retrieved too many documents, or all results seemed irrelevant to the user's question.

HSS11: inability to monitor search. Study participants sometimes were lost. Either they did not know where they were (they had lost track of how they had reached the state they were in and could not tell what to do next) or they wanted to return to a previous state (for example to start again) but they did not know how. Participants were often confused about how to move between the main digital library site and its help site as well as between the digital library site and other sites such as the home of the browser. Several participants tried to use the search mechanism from the help document describing it by trying to click on an image used as an example in the help document. Other users unsuccessfully tried to use the browser navigation buttons to navigate in the digital library site: “I click home, get out of the site to MSN. How do I go back to American Memory site.” “How do I get here? Is this the beginning the homepage. I'll try help.”

HSS12: inability to identify relevant documents. Participants were often unable to recognize relevant documents from the long lists of document surrogates presented by the search system. Example: “Many results, unable to identify the relevant answers”

HSS13: inability to identify specific information. This situation was observed when participants were asked to find specific answers to questions. For example: “find how many people have been affected by HIV in Uganda.” Users did not know which tools to use, or whether any was provided by
the system, in order to find specific information within the documents found. Example: “how do I search within one of the documents retrieved?”

HSS14: inability to compare items retrieved. This situation was observed when participants wanted to compare two of the documents retrieved. Some pairs of items made participants suspect the documents were the same, even though the surrogates presented by the system were slightly different, but they thought the only way to decide was to actually compare the documents. This situation was caused by several reasons: the surrogates presented by the digital library did not contain the information necessary, and the users lacked enough domain knowledge. Examples: “… this looks like the same thing if I take the time to read…” “are these two retrieved items actually the same?”

HSS15: inability to verify relevant documents. When participants found multiple answers or confusing results in topics about which they lacked enough domain knowledge, they did not know how to decide which answer was more authoritative and accurate than the others. Example: “There are many answers for the question HIV in Uganda, but I am not sure which result is correct.”

OVERVIEW OF THE DESIGN PRINCIPLES

In this section we suggest a list of principles for the design of help systems for digital libraries (DLs) based on the help seeking situations identified in the user study.

Support systems can be categorized as follows:

- online tutorials
- online documentation
- online help

The differences between the three are determined by the way in which they are intended to be used.

The three kinds of online support systems may contain some combination of descriptive text, examples, and interactive features such as hands-on experimentation.

Online tutorials are systems designed to teach users about features and tasks. Users may use tutorials to learn about the system not necessarily while they are using it but perhaps in preparation to use it. Tutorials represent an investment in time and effort that few casual users may be willing to spend. Online documentation is intended to be used more like a reference than a learning tool. Online help is meant to support users...
who must solve problems that they encounter while they are using the system. Online help should be fast, efficient and provide support within the context in which the Help Seeking Situation arises.

This paper deals mostly with online help but we propose that a full support system should include also what is called above “online documentation” and “online tutorials”. The study we conducted on which we base the design principles described here focused on users trying to solve problems they encountered while using digital libraries. Such users needed the kind of help we called “online help” above. It is interesting to note that some systems we have studied (such like the LA Public Library online system as of 2009) seem to offer “online documentation” in lieu of “online help”. The experience we gathered while conducting the user study suggests such help is probably not very effective.

What follows is the list of design principles for help systems. In the following sections each principle is described and, for each, the relevant HSSs are listed.

General principles: apply the same general design principles that would apply to any user interface
Reduce users' cognitive load, make all interfaces as simple as possible, consistent, easy to learn, and easy to use
Increase effectiveness
Reduce defects of the main search system that cause the user to seek help.
Search engine: better ranking algorithms, implicit and explicit feedback, and question answering.
Main search interface: better surrogates, browser independent navigation and search controls, browsing tools, avoid having different states, open document in most relevant section, and don’t make users choose between collections
Access to online help
Don't force users to interrupt their task to seek help
Provide multiple access methods to the different layers of the help system (menus, text search).
Integrate help and search system
Provide context sensitive help
Identify common Help Seeking situations and offer relevant help.
Help tools
Provide effective browsing tools
Provide interactive search agent
Do not force the user to learn a query language. Help users build queries instead of expecting them to learn any more or less formal query language.
Help information content and style
Help information should include information about the content of the digital library as well as how to use it.
Include definitions of all specialized terms.
Include examples, tutorials and demos.

GENERAL PRINCIPLES

The user interface for an online help system must follow the same design principles of any user interface. Several researchers have proposed rules and principles for the design of user interfaces. Shneiderman's golden rules for dialog (Shneiderman, 1997) are a good example:

Consistency. Consistent sequences of actions should be required in similar situations; identical terminology should be used in prompts, menus, and help screens; and consistent commands should be employed throughout.

Enable frequent users to use shortcuts.
Offer informative feedback.
Design dialog to yield closure. Sequences of actions should be organized into groups with a beginning, middle, and end. The informative feedback at the completion of a group of actions gives users the satisfaction of accomplishment.
Offer simple error handling.
Permit easy reversal of actions.
Support internal locus of control. Experienced operators strongly desire the sense that they are in charge of the system and that the system responds to their actions.
Reduce short-term memory load.

We focus here in principles that facilitate the use of help systems. For example: reduce user's cognitive load, and make the user interface simple, consistent, easy to learn, easy to use, and effective.

The design principles described in the following sections include some of these general principles and also some new ones that are specific to online help systems. The principles proposed specifically for online help systems are based on the lessons learned in our user study.

An important design principle applicable to help system interfaces is the need to reduce the user’s cognitive load. In order to achieve that goal everything that is not essential to help users enter their query should be removed from all interfaces. If you are
designing a help user interface, consider each element of the main interface, if it doesn't directly help the user to enter a more effective query delete that element from the interface. Users are overwhelmed by too many choices or too much information. The main search page of the DL system must eliminate all unnecessary clutter. Any user interface element that does not directly contribute to help the user find the information they are looking for actually hurts the effectiveness of the system.

The following items represent some of the elements a help user interface should include. Anything else should be added only after careful consideration:

- a banner that identifies the website
- a prominently displayed text box where the users can enter a query or question (preferably on the top left of the page or in the center).
- text that invites users to enter a query or question: this can be a link next to the text box that offers users examples of queries that can be entered (addressing HSS2: users unfamiliar with digital libraries). This is also an example of applying context sensitive help.
- buttons for each of the formats (images, maps, documents, sound, etc.) offered by the DL
- a button labeled “Advanced Options” that takes the users to a query by example (QBE) page.
- a button labeled “about this Digital Library” that takes users to the help documents that describe the DL: its history, its contents, its authority, etc.
- a button labeled “Help” that takes the user to the main help system.

The LOCDL violates the simplicity principle in many ways. It offers “Collection Highlights” and “today in history” information on the main page (as of Jun 2, 2009). Actually less than half of the interface is actually about searching information. The rest is what would be considered “promotions” in a commercial site (e.g. “Teachers: use American Memory in the classroom”). This adds to the clutter on the page and does not help the user to search for information. It would be better to separate both aspects completely so users could find a page that has no other purpose but to help them find what they are looking for. All the “promotional” stuff could be in a different page containing information “about us”. The search page should be as simple as possible.

**REDUCE DEFECTS OF THE MAIN SEARCH SYSTEM THAT CAUSE THE USER TO SEEK HELP**

Several of the Help Seeking Situations described in the user study are the result of shortcomings or defects of the search system. The better the main search system is
designed the less help-seeking situations will occur. No redesign of the help system should be used (or can be used) to overcome what are really problems of the main search system. The help system cannot be used to make effective a search system that is not.

Listed below are several examples of problems or defects of the search systems used by the DLs used in our study that were found to cause the user to seek help.

Question answering. At least simple (and common) questions should be parsed automatically and transparently into good queries for the search system (see HSS1 (unclear about the topic), HSS2 (unfamiliar with digital libraries), HSS5 (inability to formulate query statements), HSS6 (inability to select appropriate query terms)). Google does this quite effectively. For instance if users want to find out the name of their representatives they can simple enter “who is my congressman and senator?” and find websites that contain the answers. This as opposed to trying to figure out what kind of documents might contain the desired answer.

Improve surrogate presentation. The document surrogates are difficult to evaluate (see HSS12 (inability to identify relevant documents)). It is not clear to the user which surrogates actually stand for the same document (see HSS14 (inability to compare items retrieved)).

Avoid having different states, e.g. the state where the user can enter queries, the state where the user can look at surrogates (see HSS11 (inability to monitor search)).

SEARCH ENGINE

Some defects of the search engine itself cause users to seek help. Some of the Help Seeking Situations observed in the study were caused by defects in the search engine. Some areas in particular that could be improved in order to reduce Help Seeking Situations are: the ranking algorithms, use feedback to reformulate queries, and question answering.

If the ranking algorithms are not very good, it will be harder for users to find relevant documents at the top of the list of hits retrieved. Users feel that they are overwhelmed by too many irrelevant documents (see HSS12 (inability to identify relevant documents)).

QUESTION ANSWERING

At least simple queries should be translated automatically, and transparently, into appropriate queries for the Digital Library. For example, queries such as “Where is
Yosemite?”, “What is Bubonic plague?”, “Who is my representative” should be “understood” by the system, i.e., translated into queries that are likely to return documents that contain the answer to the question. Several users in our study thought they needed to learn some specific facts about the domain. For example, in order to find a map of Yellowstone, as require by one of the study tasks, they tried first to find out where is Yellowstone from the DL. Some users did not know how to proceed and tried then to use the help system to answer such questions. It should be clear from the user interface and from the help available at the point where users must enter a query (context-sensitive help), that it is possible to enter simple questions as queries.

Users may have trouble figuring out how to start a search (HSS1 (unclear about the topic), HSS2 (unfamiliar with digital libraries)). From a “getting started” button (located right next to the search box) users should be able to get information about how to create a query. On the pop-up help window the user should be told that one possibility to start a query is simply to “ask a question”. It should not be necessary to learn a special syntax, nor should it be necessary to choose between collections. The system should be able to do at least some very simple parsing of users' queries in order to identify simple questions (such as “what is implicit feedback”, “where is Yellowstone”) and retrieve documents that may contain the answers.

The surrogates presented should be built in such a way that the answer is likely to be shown in the surrogate (see for example Figure 1). This might work simply showing the fragment of the document most relevant to the query.

The users’ experience would have been completely different if instead of using a DL they were using the Web for the following reasons:

- in the Web, users don't use a help system to ask questions about the domain, they consult the web itself
- often Web search engines, such as Google, are designed to parse simple questions such as “where is ...”, “what is...” When such queries are used, documents likely to be relevant will appear at the top
- the document surrogates of a Web search engine are very likely to actually contain the answer to the question, so the results are easy to evaluate. This is because the Web, due its size, contains many variations of answers to a large number of questions, and search engines, such as Google, do a good job building useful surrogates to present to users. See examples of surrogates in Figures 1, and 2.
Figure 1: information shown for first hit of query “where is Yellowstone” by Ask.com

Yellowstone National Park. Established in 1872, Yellowstone National Park is America's first national park. Located in Wyoming, Montana, and Idaho, it is home to a...More

Go To: Official Site, Map, Directions, Activities, Campgrounds
Search For: Camping Reservations Nearby Airports: Jackson Hole, Idaho Falls

Figure 2: surrogates shown for first hit of query “where is Yellowstone” by Google

Where is Yellowstone National Park? Yellowstone National Park... is in the northern Rocky Mountain Range. Located mostly in Wyoming, the park crosses into Montana to the north and Idaho to the ... www.yellowstonejobs.com/whereisyellowstone.html - 11k - Cached - Similar pages

On the other hand, in the LOCDL:

users think they may find information about the domain using the help system queries such as “where is ...” may not be parsed appropriately. For example the question “Where is Yellowstone?” in the LOCDL, retrieves documents that contain “Yellowstone” but that are not necessarily likely to contain information about its location (see Figure 3).

The LOCDL surrogates are not very good.

Figure 3: surrogates for query “where is Yellowstone” from LOCDL

1. Lamar Buffalo Ranch, East of Tower Roosevelt on Northeast Entrance Road, Canyon Village vicinity, Park County, WY Built in America

2. DCM 0876: Anonymous, North American Indian (Sioux) / Courting Flute (Vertical Whistle Flute) Miller Flutes


MAIN SEARCH INTERFACE: SURROGATES

Users need better surrogates (see HSS12 (inability to identify relevant documents)). The surrogates provided by the LOCDL (see figure 3) are particularly bad. LOCDL shows only the titles of the documents. Unfortunately the titles of government documents are infamously bad at providing clues about their content. The surrogates provided by the system should include the titles but also keywords in context, collection the document comes from, dates and authors. Some of the information provided in the
document surrogates should be made mouse sensitive so the user can reformulate the query by restricting the search to those dates, that collection, those authors, etc. Also the terms in context should be clickable so the user can jump directly to the relevant passages in the documents. The fact that the LOCDL surrogates are hard to evaluate cannot be solved by a better help system but by providing better surrogates.

Figure 3 shows the first few hits for the query “Where is Yellowstone?” as shown by the LOCDL. From the text shown to the user in the document surrogates it is hard for them to tell whether any of these documents may be relevant to their information problem.

Figures 1 and 2 show examples of surrogates generated by Ask.com and Google respectively in response to the same query. Google's surrogates include the title of the document as well as document text that is presented showing the keywords the user entered highlighted. This kind of surrogate makes it easier for the user to evaluate the relevance of the document without having to actually look at it. Better designed surrogates would address HSS12 (inability to identify relevant documents) and HSS13 (inability to identify specific information).

HSS13 (inability to identify specific information) could be addressed by presenting in the surrogate the text from the document that is most relevant to the query. See Figure 2 for an example of how Google does something similar.

HSS14 (inability to compare items retrieved) would be addressed by detecting duplicates automatically (as Google does) or other relationships between docs such as one being a superset of the other, etc. For example Google uses indentation to indicate that some surrogates belong to the same website.

HSS15 (inability to verify relevant documents) would be addressed by having the surrogate itself display information about the source such as name of the document, name of the collection, dates, and authors. Each of the document characteristics should be linked to more information about them that helps users evaluate the document. For example: information about the institution or publication sponsoring the document, information about the author, etc. From the surrogate it should be possible to open a window containing a different kind of surrogate: a summary built (automatically) from the contents of the document as well as the meta-information about the document. The summary would contain information about the genre of the document, the kind of information it has, etc. Summaries should be designed to help users evaluate documents without having to actually look at them. Some of the possible information for a summary could be used to answer questions such as “does it contain statistics?”, “does it contain information about numbers of casualties?”, “does it mention countries or geographical regions?”, etc.
BROWSER INDEPENDENT CONTROLS

In order to address HSS11 (inability to monitor search) the DL system should provide full navigation within the system without having to rely on browser's buttons. This means: (1) the designer has more control over the user experience, (2) the user experience depends less on the browser the user happens to be using, (3) the user is not confused (as happened to participants of the study) between the browser navigation controls and the system's.

Navigation controls should include a way to show users how they got to where they are as well as a way to go back to any point in the search process. This can be achieved by using breadcrumb devices (see Figure 4). The breadcrumb technique displays the search history on the same window and allows the user to jump to any previous stage. Another way, would be to present the search history on a different window (for an example of how Google does it see Figure 5).

<table>
<thead>
<tr>
<th>Figure 4: bread crumbs example from Altavista</th>
</tr>
</thead>
<tbody>
<tr>
<td>Directory &gt;Arts&gt; Design Arts &gt;Architecture &gt;Architects &gt;Masters &gt;Meier, Richard (b. 1934)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Figure 5: search history example from Google</th>
</tr>
</thead>
<tbody>
<tr>
<td>May 19, 2009,  5:38pm Searched for java programming example; Viewed 1 result</td>
</tr>
<tr>
<td>Java Programming Examples - idevelopment.info</td>
</tr>
<tr>
<td>5:28pm Searched for java tutorial sun Microsystems; Viewed 1 result</td>
</tr>
<tr>
<td>The Java Tutorials - sun.com</td>
</tr>
</tbody>
</table>

In order to address HSS13 (inability to identify specific information) provide a search mechanism within the DL site. The DL site should provide its own tools to search within documents. Most browsers offer this feature but, if it’s provided by the website itself, it will be homogeneous no matter what browser is being used. If the DL system provides the feature it can be made more powerful and friendly than what the browser uses.

BROWSING TOOLS

HSS6 (inability to select appropriate query terms) arises when users have trouble selecting appropriate query terms due to (1) lack of domain knowledge, (2) lack of experience using online search engines, and (3) lack of experience formulating queries for the search engine of that particular digital library. When users are not sure what terms to choose, or in the very common case when they are not quite sure what they need
to know from the help materials, browsing tools can help (Rice, McCreadie, & Chang, 2001).

The case of browsing tools is also a good example of the case in which the help device will not be as effective if it is not integrated into the user's workflow. If the user has to consult a separate help system he may not even know how to look for relevant help. An example of this problem are Digital Libraries with interfaces similar to the one offered (as of Jan 2009) by the Los Angeles Library online catalogue. The only apparent help offered to the user is a button labelled “Help”. When the user clicks on the button a new window (at least it is a different window!) opens showing all the help available in the form of a manual of the system. The user would have to find in this document explanations (if any) about how to build queries and select terms. The Digital Library user will seldom have the patience or time to explore this manual.

A solution is to integrate browsing and information exploration tools into the search process. Browsing tools have been used by librarians even before the advent of Digital Libraries (see (Anderson & Perez-Carballo, 2005) chapter11: displayed indexes). A browsing tool is a tool that displays to the user some level of the structure of the topics covered by the documents. In the case of Digital libraries a browsing tool should be seen as an alternative to the search box, where the user just enters the query that she hopes will retrieve the documents relevant to her information problem. Instead of the user having to come up with appropriate query terms connected with the appropriate query operators, a browsing tool allows the user to navigate through the structure of topics offered by the system. This allows users to explore the contents of the Digital Library and discover solutions to their problems. There are numerous research and development opportunities in this area (see for example (White, Kules, & Drucker, 2006)).

An example of a browsing technique, deeply rooted in the best traditions of library science, is the use of facet classifications (Stoica, Hearst, & Richardson, 2007). In a facet classification each category groups one aspect of the topic. The use of facets helps the user explore the information contained in the Digital Library. Several studies (e.g. (Hearst, 2006), (Venkatsubramanyan & Perez-Carballo., 2007)) suggest that interfaces that present results organized into categories or faceted hierarchies meaningful to users may help them make sense of their information problem as well as the information system itself.

**AVOID HAVING DIFFERENT STATES**

HSS11 (inability to monitor search) describes situations in which the users got lost navigating the search system. Changing the users' environment to a different state contributes to them getting confused and lost. Whenever possible it is better to display help information on a pop up window than to change the state of the user to a help mode.
OPEN DOCUMENT IN MOST RELEVANT SECTION

HSS13 occurs when users have trouble finding specific information inside of the documents they retrieve. One possibility is to open the document in the area deemed the most relevant by the system (e.g. the fragment that contains the most terms from the query). It would also be possible to offer to the user several points of access to the document. One for each fragment that contains possible answers to the information problem represented by the query terms used by the user.

DON'T MAKE USERS CHOOSE BETWEEN COLLECTIONS

Do not force users to choose between collections or make other unnecessary choices at the top level. Offering the choice or even forcing the users to select a collection is completely unnecessary and reduces the ability of users to find relevant information. Choosing between collections may have made sense when collections were physically distinct in brick and mortar libraries. In a digital library, a user can search all collections at the same time effortlessly. Choosing between collections must be an “advanced” choice.

Design guidelines about choosing collections:

- Provide information in the help system about the content of the different collections.
- Parse the user's query and offer to user the possibility of learning about specific collections, restricting search to some collections.
- By default search all collections. Show surrogates indicating which collection they come from.

Relevant HSS: HSS3 (inability to identify relevant collections).

ACCESS TO ONLINE HELP

The purpose of the principles described in this section is to make help information easily accessible and highly visible. Help should be transparent and always available. It should not feel to the user as a separate system, and should never feel like an interruption or diversion.
DO NOT FORCE USERS TO INTERRUPT THEIR TASK

Requesting help should not put users in a different mode or otherwise interrupt the flow of their interaction with the search system. As much information as possible should be already on the main interface. Help screens should open on a different window so users can continue their search process at the same time that they consult the help system. Whenever possible the help system should offer help on topics that may be relevant to the current state of the users' interaction. This can be achieved by a help pane that displays one or two relevant help topics at any time. This device doesn't need to be cute, annoying or distracting like some previous implementations of this idea have proved to be (e.g. Microsoft's infamous animated paper clip). Another possibility is to have the special help topics pane appear also on the main help screen that opens when the users request help. Relevant HSS: HSS6 (inability to select appropriate query terms), HSS7 (inability to select appropriate query terms).

PROVIDE MULTIPLE ACCESS METHODS TO THE DIFFERENT LAYERS OF THE HELP SYSTEM

The user should be able to access the help system through a variety of methods: menu bar, tool bar, help buttons on dialogs, help messages triggered by hovering over certain elements, etc.

All elements on the interface should be set to provide help just by hovering on them or by clicking on a nearby link. For example: hovering on the “images” button tells users that selecting this option retrieves only images. Hovering over the text box describes its purpose and provides links to query examples. Each one of the elements of the query by example (QBE) interface should have its own direct access to the help system, both by hovering over the appropriate element, and/or by clicking on a nearby link that would lead to several layers of help information (from fast and concise to deeper information, examples, tutorials, etc). Help documents and messages should open on their own window as opposed to changing the mode on the user.

Whenever a dialog window is open there should be a help option on it. If the user clicks on it the system provides information about that specific window and its choices.

If help documents are organized as a hypertext collection of documents, it should contain a table of contents and a search function. There should be a FAQ page (frequently asked questions) with links to the more popular questions and the most useful answers.

It must be possible to jump directly to the relevant situation in the system described in the document. For example: when describing how to use a query by example interface, the user should be able to open it by clicking a button. In general,
moving from a search situation to the corresponding relevant document or from a
document to the corresponding relevant search situation should always be only one click
away.

One of the access methods may be a “help” button from the main search interface
that provides access to a general help system. Relevant HSS: HSS2 (unfamiliar with
digital libraries), HSS4 (inability to browse information).

INTEGRATION OF HELP AND SEARCH SYSTEM

Help should be integrated into the system as much as possible. The user should
not be expected to click on “Help” and navigate to a document that answers their
questions. If the user needs help searching the digital library the last thing they need is
the additional challenge of navigating its help system. Using the LOCDL interface as it
was as of June 2, 2009, the user would have to follow at least three clicks to reach help
topics about “Search All Collections” (Help - Search Help - Search All Collections) or
“How to Format Search Terms.” (Help - Search Help - How to Format Search Terms).

In the aforementioned LOC interface the search box is on the top right, perhaps
not as visible as it could be. Ideally the search box should be in the most visible place of
the page, on the left hand side. Right next to the search box there should be a link labeled
“examples of questions”. When users click there they are taken to a page showing
examples of questions that can be entered in the system.

The next paragraphs show suggestions of how the integration of help and search
system can be used to address specific help seeking situations.

In order to address HSS1 (unclear about the topic), and HSS2 (unfamiliar with
digital libraries) the help and search systems should be integrated. Buttons on different
points of the user interface should display appropriate help information to the user. There
should always be a link to appropriate help information from several points on the search
user interface. As the users read the help information displayed on the pop-up help
window they should be able to jump back to the search system easily. An example of
integration between help and search system can be seen in the “Spotlight” search feature
of Mac OS X. If users click, for example, on “Open Spotlight Preferences” on the help
page, the appropriate preferences window opens allowing users to set preferences as
explained on the help page they are reading.

HSS7 (inability to limit searches to certain fields): from the search interface users
should be able to go directly to help documents showing examples of queries. From the
help documents users should be able to click on the example and start a query like the
one being explained.

HSS3 (inability to identify relevant collections): topics covered by the DL should
be linked to the corresponding collections. If a user wants to choose a specific collection
they should be able to browse through their names and descriptions, or through the topics covered by the DL. Each collection should have a help button next to it that links to information about it, its description, and examples of questions that could be answered by it. For example: “if you are looking for ... you could try collection ...”

HSS3 (inability to identify relevant collections): help user choose collection. Given the query entered by the user, the system may help the user choose collections using a variety of methods. One possible method: each collection is associated with a set of words that represent their content (lets call this “signature”). Depending on which signature the query matches better, the user is offered a list of ranked collections to choose from. Another possible method: collections can be ranked according to the number of retrieved documents that belong to each collection. For example, the collection to which most documents retrieved belong to would be ranked first, etc. When results are presented the system would offer to restrict the search to the collection (or collections) it guessed may be most relevant to the users’ query.

HSS1 (unclear about the topic), HSS2 (unfamiliar with digital libraries): There should be a direct link to information about how to start. A “getting started” button takes user to examples of tasks and corresponding queries. There should be direct access to tasks such as “asking question”, “looking for a map”, “looking for a biography”, “looking for a specific answer”. For each case, help documents show tips and strategies to build an effective query.

As the user is typing in text in the search box the system displays information about it. Google exhibits a similar behavior. This is particularly useful on a small mobile device without a full size keyboard. There are many examples of search systems that do this. For example, while the user is typing “yellow” in the search window of the Safari browser, the system displays in a near window a list of clickable relevant queries under the label “Suggestions” such as: “yellow pages”, “yellow book”, “yellowstone national park”, “yellow freight”, “yellow”, “yellow fever”, “yellowstone volcano”, “yellowstone earthquakes”, “yellow freight tracking”, “yellow lyrics.” At any time the user can stop typing and select any of the choices offered by the system.

HSS6 (inability to select appropriate query terms): when users are creating a query they can click on the words they have entered and get suggestions about synonyms, more general terms, more specific terms, or related terms.

HSS7 (inability to limit searches to certain fields): provide direct connection from the point where the user needs to create a query to help documents that offer examples on how to limit queries to certain fields or facets.

HSS15 (inability to verify relevant documents): links or buttons on the appropriate places in or next to the surrogate display allows users to investigate details about the source of the information. This context sensitive link should take users directly to help information about how to evaluate the quality of retrieved documents and that
specific source. From there, users could find out more information about sources and how to evaluate them. Another link on the surrogate opens a window containing a summary of the document designed to help the user evaluate the document retrieved easily and efficiently (see section on surrogates for the desired characteristics of the summaries).

HELP SHOULD BE CONTEXT SENSITIVE

Helpful information should be incorporated into all screens and dialog boxes of the main search interface.

Each screen of the user interface should show to the user what functions are available. The functions available at any point during the search interaction should be visible to the user instead of hidden inside of menus. The user should be able to tell what functions are available just by scanning the search interface in front of him. This should allow the user to get help on what they need with a minimum number of clicks or navigation.

The system should present to users the operations they may want to use at each stage of the interaction. For example: when showing a document, the system identifies the user's browser and provides a reminder about how to search within the document (e.g. ctr-F in Internet Explorer under windows, command-f in safari under Mac OS). Also, offer appropriate points of entry into the help system to help user with the choices offered in this context. Relevant HSS: HSS7 (inability to limit searches to certain fields).

The main interface should make clear to the user what kinds of questions can be asked. The users in the study did not understand (see HSS1 (unclear about the topic), HSS2 (unfamiliar with digital libraries)) what kinds of questions/queries could be submitted to the search system (looking for content) and which to the support system (look for help information).

Help seeking situations HSS1 (unclear about the topic) and HSS2 (unfamiliar with digital libraries) are situations in which users had difficulty starting the search process because they felt they did not have enough information about the domain (HSS1 (unclear about the topic)) or about the kinds of questions they could submit to the system (HSS2 (unfamiliar with digital libraries)).

IDENTIFY COMMON HELP SEEKING SITUATIONS

The help system should be able to identify common Help Seeking Situations and offer relevant help. For example: if the number of retrieved items is too small or zero, the system may parse the query entered and suggest to the user changes. This is similar to the suggestions offered currently by Google in that situation:
Make sure all words are spelled correctly.
Try different keywords.
Try more general keywords.
Try fewer keywords.

For each situation there should be a point of entry to relevant documents in the help system. The help system documents could, for example: explain ways to make the query less specific, or explain different strategies to build a query. Relevant HSS: HSS6 (inability to select appropriate query terms).

HELP TOOLS

As mentioned before in the section about browsing tools for the general search system (section 7.7), browsing tools help users when they are not sure what they want. Since users may be even more confused about what they want or need when they are looking for help, this kind of device would be particularly useful in help seeking situations. Consequently, one of the ways to access help should be in the form of browsing tools similar to the ones recommended earlier for the full system.

INTERACTIVE SEARCH AGENT

Allow the possibility to interact with the user through a dialog (HSS8 (inability to refine searches for different aspects), HSS9 (inability to identify other approaches), HSS10 (inability to refine searches in general)). It is important to implement this in as unobtrusively as possible. The agent should not interrupt the users' task or put them in a different state. On a different window the agent may offer context sensitive help depending on the results of the search and some characteristics of the query. For example, if the number of documents retrieved is too large, the agent may decide that it is because the query is too short (one or two words) or the query terms are too general. It may then suggest to users ways of making their searches more specific. In other cases the agent may decide the query is too specific (e.g. no documents retrieved) and may suggest ways of making the query more general.

DO NOT FORCE THE USER TO LEARN A QUERY LANGUAGE

The user should be able to get good results without having to learn the syntax of a special language. A possibility is a Query by Example interface (QBE). A QBE interface uses a query template or form that helps naive users create complex queries. In that sense
it is part of the help system. The main interface should assume the most useful defaults for query construction. The user should be asked only to enter a question or a few words. More advanced users should be able to enter directly complex queries. Most other user choices should be delegated to an “advanced page” which should be reachable from the main page with a single click. There, users can use a query by example form (QBE) in order to create more complex queries.

In order to help users get started with the search process a query by example QBE interface is always available. QBE guides the user suggesting sets of possible values for the different facets of the search. This can be done without changing modes, like Google, which offers QBE as an advanced option (this is relevant to HSS1 (unclear about the topic) and HSS2 (unfamiliar with digital libraries)).

A query by example interface facilitates the creation of the query as well as its modification and refinement (relevant to HSS: HSS2 (unfamiliar with digital libraries), HSS3 (inability to identify relevant collections), HSS5 (inability to formulate query statements)).

HELP INFORMATION CONTENT AND STYLE

This section discusses the help documents themselves: the different styles, organizations, and kinds of help documents. Help seeking situations HSS1 (unclear about the topic) and HSS2 (unfamiliar with digital libraries) have to do with the inability to get started. Users that found themselves in HSS1 (unclear about the topic) had trouble starting because they felt they needed more information about the domain. A help system must include information about the information contained in the digital library. In keeping with the principle of keeping the interface simple, users should be able to enter questions about the domain in the same text box they would use to enter questions about the interface.

DEFINITIONS, EXAMPLES, TUTORIALS AND DEMOS

Help documents should include an overview of search strategies implemented by the DL. For example: when to use browsing, when to use “keyword” search. The documents should be written using concrete examples of common user tasks. For example: “asking a specific question”, “if you know words that may appear in the documents you are looking for”, “exploring the topics covered by the DL”, “looking for a map”. From the explanation on the help documents it is always possible to go directly to performing the task (see Integration of help and search system) (relevant to HSS1 (unclear about the topic), HSS2 (unfamiliar with digital libraries)).
It would be useful to have at least a very short demo of how to use browsing tools. This can be in the form of a very short movie or animation. Must be on demand and it should be possible to interrupt it at any time. It should appear on its own pop-up window. It should be available at any point the user can browse (relevant to HSS4 (inability to browse information)).

Give examples of when browse helps and how to do it (relevant to HSS4 (inability to browse information)).

Information about how to assign credibility or authority to the different sources contained in the DL should be provided (relevant to HSS15 (inability to verify relevant documents)).

HELP DOCUMENT STYLE

The documents should be written by a professional writer in plain language, without obscure acronyms or technicisms. They should be concise, short, and provide links to other help documents when necessary. Most help documents should contain links to related topics, as well as more general, and more specific ones.

The documents should include definitions of all specialized terms. There should be examples of different kinds of queries and information problems. There should be tutorials and step by step examples for different kinds of users. There should be help documents at different levels of detail and complexity in order to address the needs of users at different levels of competency, from the most naive to experts.

CONCLUSIONS AND FUTURE WORK

Designing help systems that will effectively support people in their information seeking goals is not an easy task, for a variety of reasons. During the course of an information search session, people often encounter a variety of problems, or help-seeking situations, that help systems are not uniformly designed to address. Typically, system design has preceded user studies, leading to poorly designed help systems. In this paper we describe a program of research that is grounded in analysis of users first, the kinds of help they need in a variety of contexts, which we identify as help-seeking situations, and from this empirical grounding we have suggested design principles to respond to these HSSs. Many of these principles are novel but others have already been incorporated into the interfaces of some internet search engines or certain operating systems. Very few have found their way yet to the interface of Digital Libraries. The next step in our research is to implement these design principles into the interface of a Digital Library and evaluate the outcomes. Our premise in pursuing this course of research is based on a belief that users ought to be considered designers and actively brought into the design
stage of system development. In our own study, participants were not directed to comment on help features during their searches, although they offered many spontaneous comments. As a next step, we suggest that users be given an active role in participatory design of future help systems. Results of our work thus far suggest several fertile areas for follow up research. Some of the questions that warrant further investigation include the following:

Do HSSs identified by us vary across novice/expert differences?

To what extent are there gender differences in using system help, and at what point during the search. We have learned from studies of social behavior that males are more reluctant than females to ask for help from a stranger. Does this finding hold true within the context of human-computer interaction?

How do searchers respond to help mechanisms in different modes of presentation? That is, what differences in help-seeking behavior might we find under conditions in which help is offered by an anthropomorphic agent; or a speech activated device?

Quite importantly, we need to remember that our classification of HSSs is based on user experiences with two digital libraries. Clearly, advancement in this area of research will require replication of the experiment with other digital libraries.

While beyond the scope of this paper, a larger goal of our research is to identify prototypical patterns of help-seeking behaviors over the course of an entire search.

Finally, the high level goal of our research program is to develop help systems that can act as interaction partners with users. A next step in this direction is to identify interaction strategies that precede use of help and facilitate the implementation of automatic help assistance. This is a challenging problem which we look forward to addressing.

REFERENCES


Borgman, C. L. (1987b). Why do some people have more difficulty learning to use an information retrieval system than others? *SIGIR '87: Proceedings of the 10th annual international ACM SIGIR conference on research and development in information retrieval*, New York: ACM.


